Market Concentration and Aggregate Productivity: The Role of Demand

Jeremy Pearce[*] Liangjie Wu[†] [‡§]

March 26, 2026

Abstract

We study the interaction of concentration and productivity when firms invest in customer capital as a key dimension of firm heterogeneity alongside productivity. We start by showing that firms invest more in customer capital (sales and marketing expenses) than in R&D, and customer capital covaries more strongly with firm size and is more persistent than productivity. We build a dynamic model of granular firms that invest in customer capital both to complement their productivity and to amplify market power. This second force creates persistent mismatch between demand and productivity. Calibrating the model to product-level price and sales data, we find that endogenous marketing more than doubles the time it takes for productivity shocks to transmit to the aggregate. Standard policies that aim to correct static markup distortions can exacerbate the dynamic misallocation between demand and productivity.

Key Words: Firm Dynamics, Productivity, Demand, Customer Capital, Market Concentration, Competition, Innovation.

JEL Code: O31, O32, O34, O41, D22, D43, L11, L13, L22

[*]Federal Reserve Bank of New York, jeremy.pearce@ny.frb.org;

[†]Einaudi Institute for Economics and Finance, liangjie.wu@eief.it;

[‡]Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

[§]We thank Tincho Almuzara, Mary Amiti, and Simone Lenzu for helpful comments. The views expressed here are the authors' and do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

How does market concentration affect aggregate productivity? This question has received growing attention in policy circles and the economics 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, a central driver of the firm-size distribution (Hottman et al., 2016; Foster et al., 2016; Einav et al., 2021), has been largely ignored. This paper introduces customer capital[1], the accumulated stock of marketing expenditure, as a second dimension of firm heterogeneity alongside productivity (Gourio et al., 2014; Afrouzi et al., 2023).

We start with three facts from Compustat and Capital IQ data on publicly traded U.S. firms. First, firms invest more in customer capital than in R&D: the average marketing-to-revenue ratio is 4.1%, compared to 2.4% for R&D. Second, customer capital covaries more strongly with firm size than productivity does: the within-industry correlation of revenue with customer capital is 0.73, compared to 0.53 for revenue per worker. Third, customer capital is substantially more persistent than revenue per worker (AR(1) coefficients of 0.98 vs. 0.86), and the two evolve largely independently. These facts challenge models that attribute firm size to a single productivity dimension and raise the question: is demand-driven concentration efficient or distortionary?

We address this by building a dynamic model of oligopolistic firms that invest in customer capital. Firms invest in demand for two reasons: to complement their productivity (the *productivity incentive*) or to amplify their market power through variable markups (the *size incentive*). Customer capital is purely allocative: among equally productive firms, marketing only redistributes attention, so it increases aggregate productivity only when matched to more productive firms. We provide a micro-foundation based on rational inattention.

This model is challenging to analyze because firms are granular: firm-level shocks are aggregate shocks. A computational contribution of the paper is a tractable algorithm based on a quadratic approximation of equilibrium profits around the point of equal customer capital, preserving nonlinearity in the productivity dimension. The resulting linear-quadratic dynamic game admits closed-form equilibrium strategies.

In equilibrium, a novel endogenous path dependency arises. Because markups are

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

increasing in market shares, larger firms have a stronger incentive to invest in customer capital regardless of their productivity, the size incentive. Firms also invest more to *escape competition* when rivals are close in size. The initial distribution of customer capital thus has a long-lasting impact on its evolution.

We then take the model to detailed product-level data. Using NielsenIQ scanner data, which provides prices and quantities at the barcode level within narrowly defined product groups, we structurally decompose market shares into demand, productivity, and markup components following Hottman et al. (2016). The residual demand component from the scanner data behaves similarly to the customer capital constructed from marketing expenditures: it covaries strongly with firm size, is more persistent than productivity, and the two are positively correlated. The product-level data also enables us to measure the gross and net flows of customer capital, through product entry, continuation, and exit, which identify the two core parameters of the model: the decay rate of customer capital ($\delta = 0.13$) and the cost of marketing ($\kappa = 1.10$).

We highlight two quantitative findings. First, the median half-life of customer capital convergence in the calibrated model exceeds 12 years, compared to 5 years implied by exogenous decay alone. The size incentive more than doubles the persistence of demand heterogeneity. Second, this endogenous path dependency weakens the correlation between productivity and demand: removing size incentives raises the predicted correlation by 15%.

The path dependency also affects the transmission of firm-level productivity shocks to the aggregate. A firm with high initial customer capital that receives a positive productivity shock accumulates *even more* customer capital over time than an equivalent firm starting with low customer capital, amplifying the aggregate effect by 7% after 10 years.

Turning to welfare, we find that endogenous marketing is welfare-enhancing on net: a marketing ban reduces welfare by 7%, as the productivity loss from eliminating the productivity incentive dominates the markup and cost savings. Underneath this overall gain, the size incentive alone reduces welfare by 4%.

The interaction of static markups and dynamic misallocation has novel policy implications. A social planner who corrects both markup distortions and the business-stealing externality achieves welfare gains of 41%. A production subsidy, which resolves markup distortions when firms differ in a single dimension (Edmond et al., 2023), can exacerbate

the dynamic misallocation: the static markup gain of 12% is partially offset by a 1.4% dynamic loss from amplified size incentives.

The remainder of this section reviews the literature. Section 2 documents the marketing facts. Section 3 develops the model. Section 4 takes the model to NielsenIQ scanner data and estimates the dynamic parameters. Section 5 calibrates the model. Section 6 presents the quantitative analysis and policy implications.

Related Literature. A growing literature studies demand as a driver of firm size, treating it as a slow-moving state variable that depends on firm investment (Dinlersoz and Yorukoglu, 2012; Gourio and Rudanko, 2014; Ignaszak and Sedlácek, 2022; Cavenaile et al., 2023; He et al., 2024). Several papers study endogenous demand in strategic settings (Cavenaile and Roldan-Blanco, 2021; Afrouzi et al., 2023; Bornstein and Peter, 2022), while others document that residual demand explains the largest share of sales variation (Hottman et al., 2016; Eslava et al., 2024) and that product creation is central to maintaining demand (Argente et al., 2020; Einav et al., 2021). Our contribution is to connect the dynamics of demand accumulation with the strategic incentives of oligopolistic firms and quantify the resulting mismatch between customer capital and productivity.

We build on a literature connecting firm heterogeneity to aggregate outcomes (Jovanovic, 1982; Hopenhayn, 1992; Klette and Kortum, 2004; Baqaee and Farhi, 2019) and on work studying how market power interacts with concentration (Akcigit and Ates, 2023; De Loecker and Eeckhout, 2018; Gabaix, 2011). We extend the insight from Atkeson and Burstein (2008) that variable markups generate misallocation into a dynamic setting, connecting to the static misallocation literature (Hsieh and Klenow, 2009; Edmond et al., 2023; David and Venkateswaran, 2019) and to the role of intangible assets in firm value (Haskel and Westlake, 2017; Crouzet and Eberly, 2019; Pearce and Wu, 2022).

2 Marketing Facts

What drives differences in firm size within an industry? A natural candidate is productivity. In this section, we document three facts suggesting that a different margin, accumulated marketing expenditure or customer capital, is quantitatively important: firms invest more in customer capital than in R&D, customer capital predicts market share more strongly than productivity does, and customer capital is substantially more

persistent. These facts motivate the dynamic model in Section 3.

2.1 Data and Measurement

We draw on Capital IQ matched to Compustat for firm-level marketing expenditures, revenue, and employment of publicly traded U.S. firms. Capital IQ reports sales and marketing expenditures separately from SG&A, following the classification in He et al. (2024). The sample covers 2007 to 2022. All dollar variables are deflated to 2017 real dollars.

We construct a stock of customer capital via the perpetual inventory method:

$$B_{it} = (1 - \delta)\, B_{i,t-1} + M_{it}, \tag{1}$$

where M_{it} is real marketing expenditure and $\delta = 0.15$ is the annual depreciation rate.[2] We initialize the stock in 2007, the first year of Capital IQ marketing coverage. We measure productivity as log real revenue per worker. The stock B_{it} is the empirical counterpart of the demand state variable b_{it} in the model of Section 3.

In Section 4, we turn to NielsenIQ Retail Measurement Services (RMS) scanner data, which provides prices and quantities at the barcode level within narrowly defined product groups. This enables a structural decomposition of market shares into demand, productivity, and markups following Hottman et al. (2016), and identifies the dynamic parameters governing customer capital accumulation and decay.

2.2 Customer Capital and Firm Size

We document three facts about the relationship between customer capital, productivity, and firm size. Table 1 reports the key moments for two samples: firms with observed Capital IQ marketing data, and an expanded sample where we impute marketing for firms that report SG&A but not marketing, using the industry $\times$ year median marketing/SG&A ratio.[3] The patterns are consistent across both.

[2] We follow **?** for the construction of marketing expenditures and use the depreciation rate from purchase price allocation useful-life estimates. Our results are robust to alternative depreciation rates $\delta \in \{0.05, 0.10, 0.20, 0.25, 0.30\}$.

[3] For firms without observed marketing, we compute the NAICS-3 $\times$ year median ratio of marketing to SG&A among firms reporting both, then multiply by the non-reporting firm's SG&A to impute marketing expenditure. Customer capital is then accumulated via the same perpetual inventory method.

Fact 1: The average marketing-to-revenue ratio is 4.1% (He et al., 2024), compared to 2.4% for R&D.

Fact 2: The within-industry correlation of customer capital with log revenue is 0.73–0.83, versus 0.50–0.53 for revenue per worker (Table 1).

Fact 3: Industry-demeaned AR(1) coefficients are 0.98 for customer capital versus 0.86 for revenue per worker; within-firm AR(1) gives 0.81 versus 0.49 (Table 1).

Table 1: Customer Capital, Productivity, and Employment

	Customer Capital (b)	Productivity (a)	Employment
Panel A: Marketing Sample ($N = 30{,}816$ firm-years)			
Corr(log revenue, $\cdot$)	0.73	0.53	0.93
AR(1) coefficient	0.98	0.86	0.99
Correlation (x_t, x_{t-1})	0.99	0.88	0.99
AR(1), firm FE	0.81	0.49	0.79
Panel B: Imputed Sample ($N = 56{,}349$ firm-years)			
Corr(log revenue, $\cdot$)	0.83	0.50	0.92
AR(1) coefficient	0.98	0.88	0.99
Correlation (x_t, x_{t-1})	0.99	0.90	0.98
AR(1), firm FE	0.89	0.50	0.78

Notes: Correlations and AR(1) coefficients computed within NAICS-3 $\times$ year cells. Marketing sample: customer capital is log B_{it} from capitalized marketing expenditure ($\delta = 0.15$). Imputed sample adds firms where marketing is imputed from the NAICS-3 $\times$ year median marketing/SG&A ratio. Productivity is log revenue per worker; employment is log headcount. AR(1) from OLS with robust standard errors. Firm FE: within-firm AR(1) with year dummies. Sample: publicly traded U.S. firms, 2007–2022.

Across every row of Table 1, the customer capital column exceeds the productivity column: customer capital correlates more strongly with revenue and is more persistent, both across and within firms. These facts have direct implications for measured misallocation. Since market shares depend on both productivity and customer capital, what Hsieh and Klenow (2009) interpret as allocative inefficiency partly reflects accumulated demand heterogeneity. Whether this demand-driven concentration is efficient or distortionary depends on the mechanism through which customer capital accumulates. The model of Section 3 formalizes two forces: a *productivity incentive* that sorts demand toward productive firms, and a *size incentive* where variable markups reward large firms regardless of productivity. When the size incentive dominates, customer

capital entrenches incumbents and slows the transmission of productivity shocks to the aggregate. Section 4 quantifies these forces using scanner data and shows that standard policies targeting static markup distortions can amplify the dynamic misallocation between demand and productivity.

3 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 6.

Section 3.1 starts by describing the environment of households and firms. Section 3.2 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 3.4 then examines the path dependency of demand with an illustrative example. Finally, Section 3.5 discusses aggregation and welfare.

3.1 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

$$\max_{c_{it},L_t} \int_0^\infty e^{-rt} \left(\ln \frac{C_t}{B_t} - L_t \right) dt, \tag{2}$$

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.[4]

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 (CES) across I firms with

$$C_t = \left[\sum_{i=1}^I e^{b_{it}/\sigma} c_{it}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \qquad B_t = \left(\sum_{i=1}^I e^{b_{it}} \right)^{\frac{\sigma}{\sigma-1}}. \tag{3}$$

$\sigma > 1$ 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. Customer capital is purely allocative: among equally productive firms, any configuration of demand leads to the same social surplus.

[4] 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.

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.[5]

To map the productivity process directly into our empirical analysis, we assume that the (log-)productivities of firms follow independent Ornstein-Uhlenbeck process: $da_{it} = -\rho_A a_{it} dt + \nu_A dW_{it}^A$, where ρ_A measures the strength of mean-reversion of the productivity, ν_A is the firm-level volatility of productivity, and W_{it}^A is a standard Brownian motion.

The customer capital of firm i evolves both endogenously and exogenously. Firm i can grow its customer capital with marketing investment η_{it}. Marketing investment is costly. We model this as a quadratic labor cost in advertising $\kappa\frac{\eta_{it}^2}{2}$. Exogenously, customer capital also fluctuates. Mathematically, the law of motion for b_{it} is

$$db_{it} = (\eta_{it} - \delta b_{it})dt + \nu_B dW_{it}^B,\ b_{i0} = 0, \tag{4}$$

where δ measures the decay rate of customer capital, ν_B is the volatility of customer capital, and W_{it}^B is a standard Brownian motion.

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.

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 compete by choosing quantity, the Cournot competition, statically.

[5]For expositional simplicity, we scaled the labor productivity into the unit of market shares, so that a one percent difference in productivity between two firms maps directly into one percent difference in their market shares. This scaling will be treated explicitly in our quantitative analysis.

3.2 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:

$$s_{it} = \frac{\exp(a_{it} + b_{it})(1 - s_{it})^{\sigma-1}}{\sum_{j=1}^{I} \exp(a_{jt} + b_{jt})(1 - s_{jt})^{\sigma-1}}. \tag{5}$$

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 $\mathbf{a}$ 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 $\mathbf{a}$ and the customer capital $\mathbf{b}$.

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

$$V_i(\alpha, \mathbf{b}) = \max_{\eta_i \geq 0} \mathbb{E} \int_0^\infty e^{-rt} \left(\pi_i(\alpha_t, \mathbf{b}_t) - \frac{\kappa}{2} \eta_i(\alpha_t, \mathbf{b}_t)^2 \right) dt \tag{6}$$

s.t. $\forall\, (j, t)$

$$da_{jt} = -\rho_A a_{jt} dt + \nu_A dW_{jt}^A,\ a_{j0} = 0,$$

and

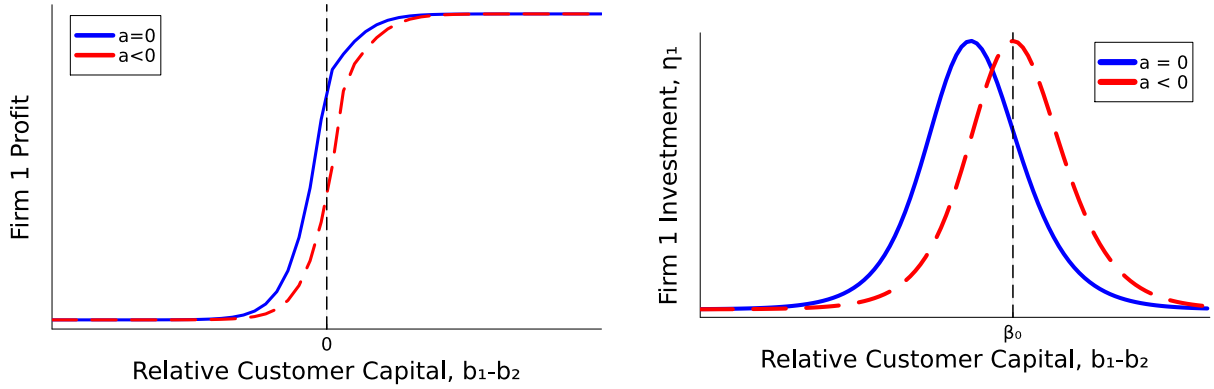
$$db_{jt} = (\eta_{jt} - \delta b_{jt}) dt + \nu_B dW_{jt}^B,\ b_{j0} = 0$$

There are three noteworthy features of the equation (6). First, we impose that all firms follow a Markov strategy: marketing investment $\eta_i(\mathbf{a}_t, \mathbf{b}_t)$ is only a function of the current payoff-relevant state $(\mathbf{a}_t, \mathbf{b}_t)$ 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* $\{\eta_i(\mathbf{a}_t, \mathbf{b}_t)\}_{i=1}^I$*, where the* $\eta_i(\mathbf{a}_t, \mathbf{b}_t)$ *solves equation* (6) *for each firm* i*, given the other firms' strategies* $\{\eta_j(\mathbf{a}_t, \mathbf{b}_t)\}_{j \neq i}$*.*

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

Figure 1: The Incentives for Customer Capital Investment



(a) Profit Function (b) Firm Investment Intensity

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 build intuition, consider the limit $r \to \infty$ where firms maximize static profits. The optimal marketing investment is

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

where $\zeta_j = \left(\frac{1}{\bar{s}_j} + (\sigma-1)\frac{1}{1-\bar{s}_j}\right)^{-1}$. The share-customer response approaches zero when the focal firm's share approaches 0 (the *discouragement effect*) or 1 (the *cannibalization effect*). Firms invest most when they are similar in size to competitors, to *escape competition* (Aghion et al., 2005). Figure 1 illustrates: the sigmoid profit function (panel a) implies that investment incentives vary with both productivity and existing customer capital, and granular firms internalize that larger market share raises both scale and markups, creating the size incentive (panel b).

These strategic incentives can sustain suboptimal equilibria: when an unproductive

firm has sufficient initial customer capital, the size incentive can lead it to invest more than the productive firm, maintaining market leadership. Once $\delta > 0$, stability is restored, but the intuition carries over to the transition dynamics. We now turn to the full framework with many firms and finite r.

Linear-quadratic Approximation. Analyzing the non-linear game with more than two firms is challenging. To gain tractability in our dynamic game, we perform a second-order Taylor expansion of firm i's profit function, $\pi_i(\mathbf{a}, \mathbf{b})$ near the point $(\mathbf{0}, \mathbf{0})$, for every i. In the Appendix, we derive such an approximation analytically. It suffices to consider only the gap between an individual firm's productivity and customer capital to the group-level averages: $\alpha_i \equiv a_i - \frac{1}{I}\sum_j a_j$ and $\beta_i \equiv b_i - \frac{1}{I}\sum_j b_j$.

Lemma 1 (Quadratic approximation of profits) *Consider the symmetric allocation* $(\mathbf{a}, \mathbf{b}) = (\mathbf{0}, \mathbf{0})$ *and define firm-level deviations from the cross-sectional means*

$$\alpha_i \equiv a_i - \frac{1}{I}\sum_j a_j, \qquad \beta_i \equiv b_i - \frac{1}{I}\sum_j b_j.$$

Let

$$\bar{\zeta} \equiv \left.\frac{\partial \log s_i}{\partial (a_i + b_i)}\right|_{(\mathbf{a},\mathbf{b})=(\mathbf{0},\mathbf{0})} \quad \textit{and} \quad \bar{\nu} \equiv -\left.\frac{\partial^2 \log s_i}{\partial (a_i + b_i)^2}\right|_{(\mathbf{a},\mathbf{b})=(\mathbf{0},\mathbf{0})}.$$

Then, in a neighborhood of $(\mathbf{0}, \mathbf{0})$*, firm i's profit can be approximated by*

$$\pi_i(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \bar{\pi} + p\,(\alpha_i + \beta_i) + q\,(\alpha_i + \beta_i)^2 + \phi \sum_{j=1}^{I} (\alpha_j + \beta_j)^2 + o\big(\|(\boldsymbol{\alpha}, \boldsymbol{\beta})\|^2\big), \tag{7}$$

where

$$p = \left(\frac{1}{\sigma} + 2\frac{\sigma - 1}{\sigma}\frac{1}{I}\right)\bar{\zeta},$$

$$q = \frac{1}{2}\left[-\bar{\nu}\bar{\zeta}^3\left(\frac{1}{\sigma} + 2\frac{\sigma - 1}{\sigma}\frac{1}{I}\right) + 2\bar{\zeta}^2\frac{\sigma - 1}{\sigma}\right],$$

$$\phi = \frac{1}{2}\bar{\nu}\bar{\zeta}^3\left(\frac{1}{\sigma} + 2\frac{\sigma - 1}{\sigma}\frac{1}{I}\right).$$

The elasticity $\bar{\zeta}$ measures the percentage change in market share from a one-percent change in composite productivity at the symmetric point. The parameter ϕ captures the size incentive: the marginal return to customer capital is increasing in a firm's existing

composite productivity. This is the force that generates endogenous path dependence. The approximation preserves the nonlinearity in the productivity dimension, as we expand around each $(\alpha, \mathbf{0})$. The closed-form derivatives of market shares with respect to customer capital, and the resulting non-monotonic relationship between firm size and responsiveness to customer capital investment, are derived in Appendix A.

Equilibrium Marketing Investment. The sufficiency of the productivity gap and the customer capital gap also manifests in the dynamic marketing investments. Specifically, we prove in the Appendix that the equilibrium marketing investment of firm i only depends on its gap to the averages. The result is summarized in the following Lemma:

Lemma 2 *In the marketing equilibrium with approximated profits, firm i's marketing investment is*

$$\eta_i(\alpha_i, \beta_i) = \bar{\gamma} + \epsilon_A \alpha_i + \epsilon_B \beta_i$$

where

$$\bar{\gamma} = \frac{p}{\kappa(r - \epsilon_B)}$$

$$\epsilon_B = \left(\frac{r}{2} + \delta\right) - \sqrt{\left(\frac{r}{2} + \delta\right)^2 - \frac{2}{\kappa}(q + \phi)}$$

$$\epsilon_A = \frac{2(q + \phi)}{(r + \rho_A + \delta - \epsilon_B)\kappa}$$

Firms' optimal investment implies that the group-level customer capital follows a VAR representation:

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

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,

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

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*: more productive firms invest more in customer acquisition due to complementarity. $\mathcal{E}(\alpha)$ captures the *size incentive*: firms with larger customer capital reinforce their advantage through variable markups. The evolution of firm-level

customer capital is driven by both incentives. 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.

3.3 Joint Distribution between Productivity and Demand

We now explicitly characterize the joint distribution between firm-level productivity and customer capital. Given the linear responses of firms to the gap between their productivity/demand and the group averages, the full dynamics of a product group are summarized by two vectors of Ornstein-Uhlenbeck processes. More precisely

$$d\mathbf{a} = -\rho_A \mathbf{a}\, dt + \nu_A d\mathbf{W}_A$$

$$d\mathbf{b} = (-\delta + \epsilon_B)\mathbf{M}\,\mathbf{b}\, dt + \epsilon_A \mathbf{M}\,\mathbf{a}\, dt + \nu_A d\mathbf{W}_B$$

where $\mathbf{M} = \mathbf{I} - \mathbf{R}$, with $\mathbf{I}$ is a I-dimensional identity matrix and $\mathbf{R}$ is a matrix where every element is $\frac{1}{I}$. $\mathbf{M}$ takes the productivity vector $\mathbf{a}$ and demand vector $\mathbf{b}$ and convert them into deviations to the mean. Within each individual product group, the realizations of $\mathbf{a}$ and $\mathbf{b}$ never settle, given the granular nature of the shocks. When aggregated across product groups, the economy has a stationary distribution for $\mathbf{a}$ and $\mathbf{b}$ that is normal, $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ with

$$\mathbf{\Sigma} = \begin{bmatrix} \bar{v}_a \mathbf{I} & \bar{v}_{ab}\mathbf{M} \\ \bar{v}_{ab}\mathbf{M} & \hat{v}_b \mathbf{R}T + \bar{v}_b \mathbf{M} \end{bmatrix},$$

where $\bar{v}_a = \frac{\nu_A^2}{2\rho_A}$, $\bar{v}_{ab} = \frac{\varepsilon_A \nu_A^2}{2\rho_A(\rho_A+\delta-\varepsilon_B)}$, $\hat{v}_b = \frac{\nu_B^2}{2\delta}$, and $\bar{v}_b = \frac{\nu_B^2}{2(\delta-\varepsilon_B)} + \frac{\varepsilon_A^2 \nu_A^2}{2\rho_A(\delta-\varepsilon_B)(\rho_A+\delta-\varepsilon_B)}$. We are concerned about the group-level dispersion of productivity and demand, as well as their covariance. Given the realized value $\mathbf{a}$ and $\mathbf{b}$, the variance across firms within the product group is

$$V_{\mathbf{a}} = \frac{1}{I-1}\mathbf{a}^T \mathbf{M}\,\mathbf{a} \sim \frac{\bar{v}_a}{I-1}\chi^2_{I-1}$$

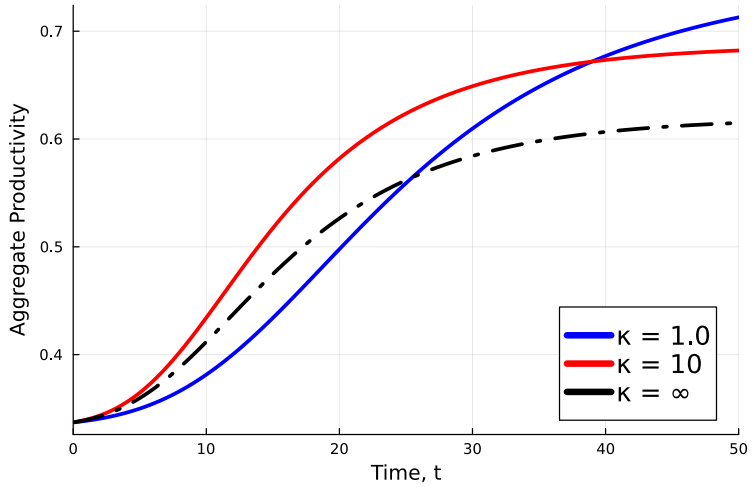
$$V_{\mathbf{b}} = \frac{1}{I-1}\mathbf{b}^T \mathbf{M}\,\mathbf{b} \sim \frac{\bar{v}_b}{I-1}\chi^2_{I-1}$$

The distribution of the covariance between productivity and demand follows a variance-gamma distribution, with an average $\frac{I-1}{I}v_{ab}$ and variance $\frac{I-1}{I^2}(v_a v_b + v_{ab}^2)$.

3.4 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 2: Aggregate Productivity w/ 2 Firms and no Shocks



The system's evolution follows $b_t = b_0 e^{-(\delta-\phi)t} + \frac{\gamma}{\delta-\phi}(1 - e^{-(\delta-\phi)t})$, where customer capital is initially distorted but converges to fundamentals in the long run. Without size incentives, the half-life of convergence is $\frac{\ln 2}{\delta}$. With size incentives, it becomes $\frac{\ln 2}{\delta-\phi} > \frac{\ln 2}{\delta}$, where ϕ captures the strength of the size incentive. As $\phi \to \delta$, convergence slows to infinity and initial conditions permanently determine market structure.

Figure 2 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 2, 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.

3.5 Aggregation and Welfare

We aggregate across product groups using a Cobb-Douglas specification, which yields linear welfare aggregation and a well-defined stationary distribution. The full household problem is in Appendix A.6. The key result is the following welfare decomposition.

Lemma 3 (Aggregation) *In a steady-state equilibrium, the discounted utility of the household is given by:*

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

where $A(\alpha,\mathbf{b}) = \left(\frac{\sum_{i=1}^{I} e^{a_i+b_i}}{\sum_{i=1}^{I} e^{b_i}}\right)^{\frac{1}{\sigma-1}}$ *is group-level labor productivity,* $M(\alpha,\mathbf{b}) = \left(\frac{\sum_i^I e^{a_i+b_i}\mu_i^{1-\sigma}}{\sum_i^I e^{a_i+b_i}}\right)^{\frac{1}{1-\sigma}}$ *is the aggregate markup, and* $D(\alpha,\mathbf{b}) = \frac{d_0}{2}\sum_i^I \eta_i^2$ *is the aggregate marketing cost.*

Two externalities arise. First, markup dispersion misallocates labor across firms. Second, firms do not internalize the benefit of matching customer capital to productive

firms and may sustain mismatch if it builds market power. This second externality is novel to our setting. A policy that aims to improve efficiency should induce firms to sort customer capital to more productive firms and reduce markups; we formalize this in Section 6.

For the empirical analysis, we extend the model to multi-product firms with decreasing returns to scale (as in Hottman et al., 2016), where $\sigma_u > \sigma$ implies constant markups across product lines within a firm. Details are in Appendix A.6.

4 Data and Empirical Analysis

Section 2 established that customer capital is large, covaries strongly with firm size, and is substantially more persistent than productivity. In this section, we use NielsenIQ Retail Measurement Services (RMS) scanner data, which provides prices and quantities at the barcode level within narrowly defined product groups, to structurally decompose market shares into demand, cost productivity, and markups following Hottman et al. (2016). The demand component maps directly to the theoretical concept of customer capital from Section 3. We document two sets of results. First, we confirm the Compustat patterns in product-level data: residual demand explains approximately four times as much as productivity in the cross-sectional variation of firm size, and demand is more persistent than productivity. Second, we construct measures of the gross and net flows of customer capital, through product entry, continuation, and exit, which we use to calibrate the decay rate (δ) and marketing cost (κ) of the model.

4.1 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 2 reports the number of firms, the market shares of top firms, and the number of products in our sample.

Table 2: Summary Statistics

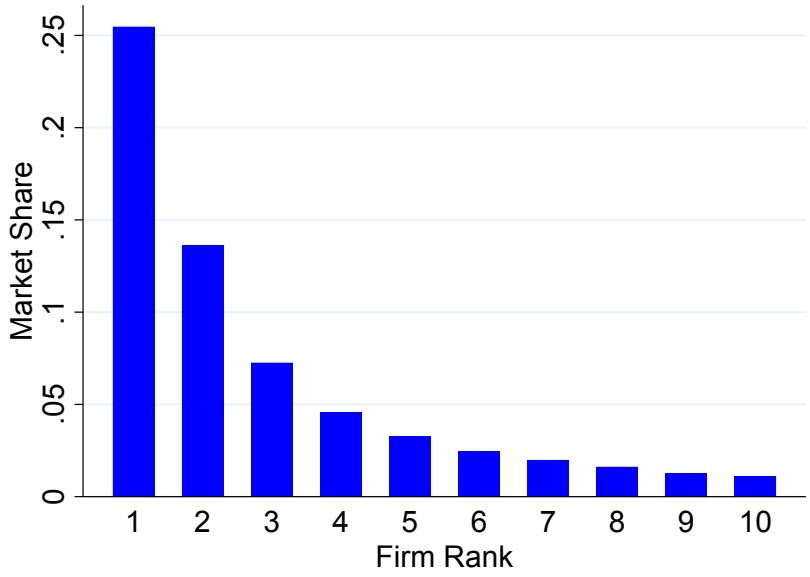
Panel A: Concentration		
	Mean	Median
Number of Firms	423	314
Average Sales (Thou)	5946	12
HHI	0.13	0.10
Top Firm Share	0.27	0.24
Top 3 Firm Share	0.50	0.46
Top 10 Firm Share	0.68	0.68
Panel B: Persistence		
	All Firms	Top 10 Firms
Sales	0.97	0.95
Price	0.76	0.43
Leadership	1.00	1.00

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

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 3 shows the firm size distribution across groups weighted by group size.

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 distribu-

Figure 3: Firm Size Distribution



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

tion. 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.

4.2 Relative Persistence of Demand and Productivity

We adapt the framework of Hottman et al. (2016) for measuring elasticities in a static setting to a dynamic setting. In the reduced form, share growth is more persistent than price growth ($\hat{\phi}_s > \hat{\phi}_p$), suggesting that demand advantages last longer than productivity advantages.

Identification. To discipline the structural elasticities, we combine the static slope conditions from the CES demand system with dynamic orthogonality restrictions in the spirit of Arellano and Bond (1991). We construct observable proxies for productivity and demand changes from price and share growth, and require that their quasi-differenced innovations are orthogonal to lagged information. This yields five moment conditions that identify $(\sigma, \omega, \rho_{aa}, \rho_{ba}, \rho_{bb})$. The key departure from Hottman et al. (2016) is that

if demand and productivity do not follow random walks, the dynamics must be incorporated to estimate σ and ω. The full moment conditions are in Appendix B.2. The resulting estimates are reported in Table 3.

Table 3: Estimated Elasticities σ and ω

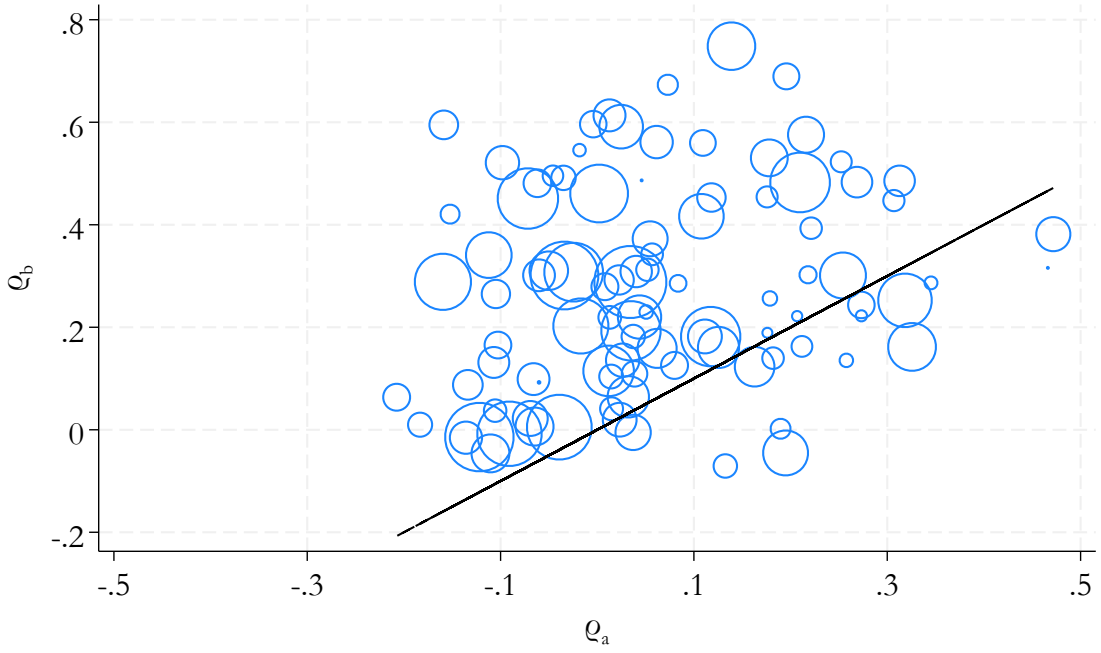
We now leverage the estimated values for σ and ω to study the persistence of demand and productivity.

Dynamics of Demand and Productivity. Given $(\hat{\sigma}, \hat{\omega})$ and the reduced-form persistence $(\hat{\phi}_s, \hat{\phi}_p)$, we can map into the structural persistence of demand and productivity:

$$\rho_b = \frac{(1+\omega)\phi_s + \omega(\sigma-1)\phi_p}{1+\sigma\omega}, \qquad \rho_a = \frac{\omega(\sigma-1)\phi_s + (1+\omega)\phi_p}{1+\sigma\omega}.$$

When $\phi_s > \phi_p$ and $\omega(\sigma-2) < 1$, demand shocks are more persistent than productivity shocks. Empirically this condition is satisfied, and we find $\hat{\rho}_b > \hat{\rho}_a$: demand advantages decay more slowly than productivity.

Figure 4: Persistence of Demand and Productivity



Notes: Demand and productivity persistence by group. Size of bubble represents average sales by group. Line represents 45-degree line which indicates equal persistence. Source: USPTO/RMS NielsenIQ.

4.3 The Dynamics of Demand

We now decompose the forces driving demand fluctuations at the firm level. Firm-level demand changes through three channels: product entry, continuation, and exit. We decompose both sales growth and residual demand into these margins. Table 4 reports the sales decomposition.

Table 4: 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.

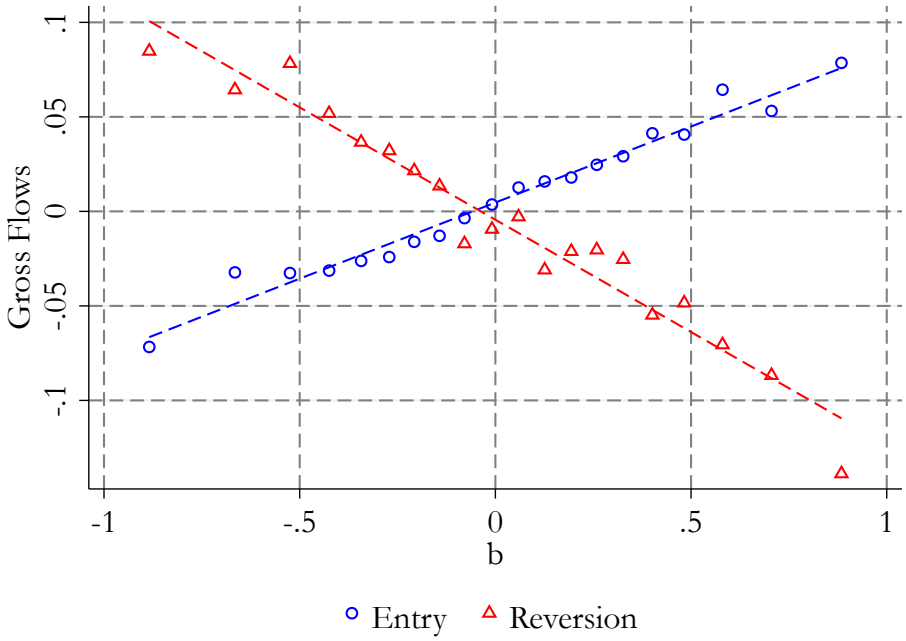
New product entry explains approximately 23% of the variation in aggregate sales, consistent with Argente et al. (2020). In the model, the two core drivers of demand changes are endogenous investment (new product entry) and natural decay (continuing and exiting products). We group the latter two together and decompose residual demand as

$$\Delta b_{ikt} = \underbrace{\frac{N_{entry}}{N_{ikt}} \cdot b_{entry}}_{\text{entry}} + \underbrace{\frac{N_{cont}}{N_{ikt}} \cdot (b_{cont,t} - b_{cont,t-1}) - \frac{N_{exit}}{N_{ik,t-1}} \cdot b_{exit}}_{\text{decay}}. \tag{10}$$

New products require active investment by firms. Given limited consumer attention, entry crowds out existing products. Figure 5 plots the creation and decay flows of customer capital against lagged customer capital for the top 10 firms within each product group. The *slope* identifies the relationship between firm customer capital and the gross flows, which we use in Section 5 to calibrate the decay rate and marketing cost.

Larger firms have higher marketing intensity: a firm with one log point higher customer capital introduces 0.07 log points more new product demand. These firms also face higher decay: 0.13 log points per log point of customer capital. The decay slope directly identifies δ; the creation slope serves as an out-of-sample test for the model's endogenous investment mechanism.

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 (10). Authors' calculations.

These findings confirm that demand heterogeneity is central to firm size and persistent, but the welfare implications remain ambiguous without a structural model. We turn to this next, using the calibrated model to decompose the dynamic manifestation of these static patterns.

5 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 4, 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.

5.1 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, $r = 0.03$. 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,

$$a_{i,t+h} = \alpha_0 + \alpha_{1,h} a_{it} + \alpha_2 a_{it-1} + \varepsilon_{it}, \tag{11}$$

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 4.3. 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 $\delta = 0.13$. 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.

Table 5: 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	δ	0.13	Empirical decay
Demand - Marketing Cost	κ	1.10	Average marketing cost share 4.1%

Notes: Parameters estimated separately and jointly. Source: RMS NielsenIQ and author calculations.

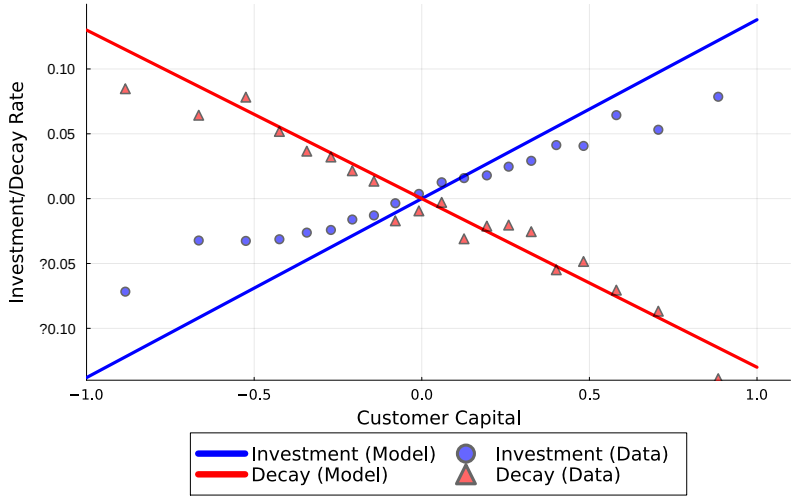
5.2 Goodness-of-Fit

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, the red line represents the targeted decay rate, matched directly by δ. The blue line is an untargeted prediction: the model's endogenous investment in customer capital, calibrated only through average marketing costs κ. The close alignment between model and data for the creation flow validates the model mechanism.

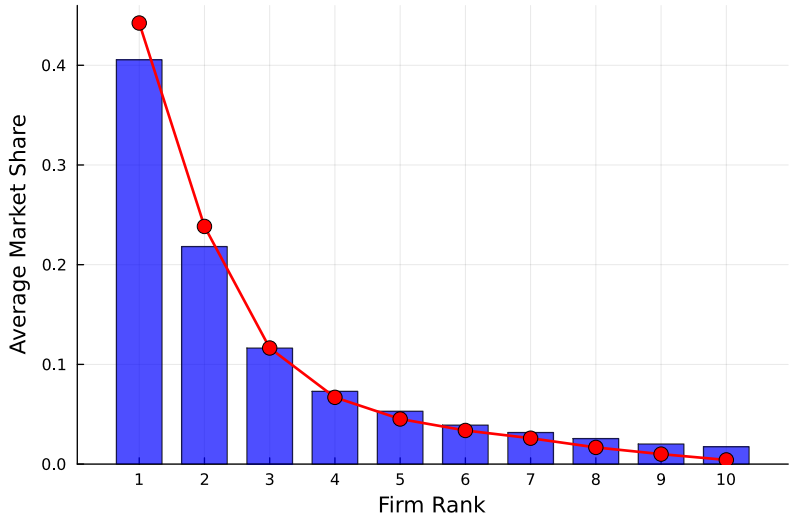
Figure 7 shows the model's predicted firm size distribution (red line) against NielsenIQ

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.

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.

data (blue bars). The model captures the skewness: the leader commands approximately 40% of the market (conditional on the top 10 firms), nearly double the second-ranked firm. This fit is untargeted: we calibrated only to customer capital dynamics, not the full size distribution.

We finally turn to moments on persistence and the correlation between productivity and customer capital in Table 6.

Table 6: 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 matches market leadership persistence (0.92 vs. 0.96 in data) exceeding productivity leadership persistence (0.90 vs. 0.88), consistent with Bartelsman et al. (2013). Customer capital leadership is more persistent than productivity (0.95 vs. 0.90 in model), and the model generates a positive endogenous correlation between productivity and customer capital (0.55 vs. 0.71 in data).[6]

6 Quantitative Analysis

This section answers the following questions. Section 6.1 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 6.2 addresses how customer capital affects the transmission of firm-level productivity shocks to aggregate productivity and, thus, aggregate allocative efficiency. Section 6.3 asks, how do these findings change our understanding of policies that aim to alleviate the distortions due to oligopoly?

6.1 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

[6]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. Details are in Appendix B.7.

of group-level productivities, we can decompose the change in aggregate productivity into four elements: (1) the change due to marketing investments increasing productivity, *productivity-enhancing investment*, $\Delta^{+}\mathbf{A}$; (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, $\Delta^{\delta}\mathbf{A}$; and (4) the change due to the exogenous shifts in productivities, $\Delta^{a}\mathbf{A}$:

$$\Delta\mathbf{A} = \Delta^{+}\mathbf{A} + \Delta^{-}\mathbf{A} + \Delta^{\delta}\mathbf{A} + \Delta^{a}\mathbf{A} \tag{12}$$

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 (12) 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: Distribution of Productivity and Customer Capital Correlation

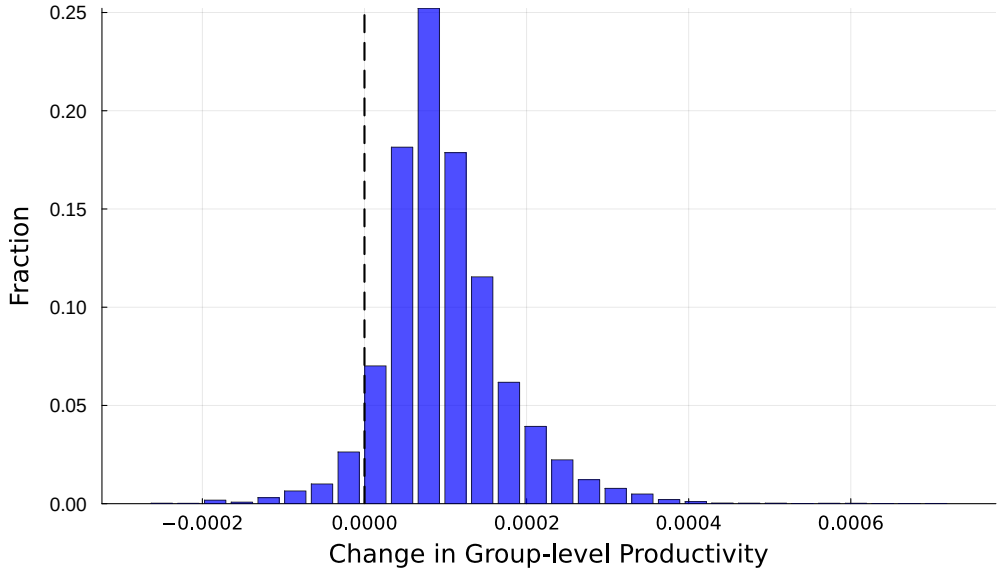
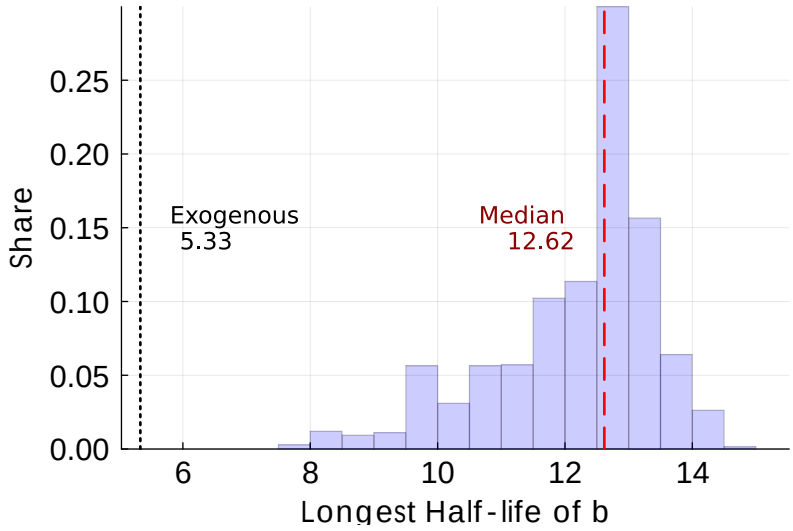


Figure 8 illustrates two central components of the drivers of the correlation between customer capital and productivity. First, the distribution reveals substantial heterogeneity in these correlations, ranging from slightly negative values to nearly perfect positive correlations, with most firms exhibiting 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.

Figure 9: Distribution of Half-Lives

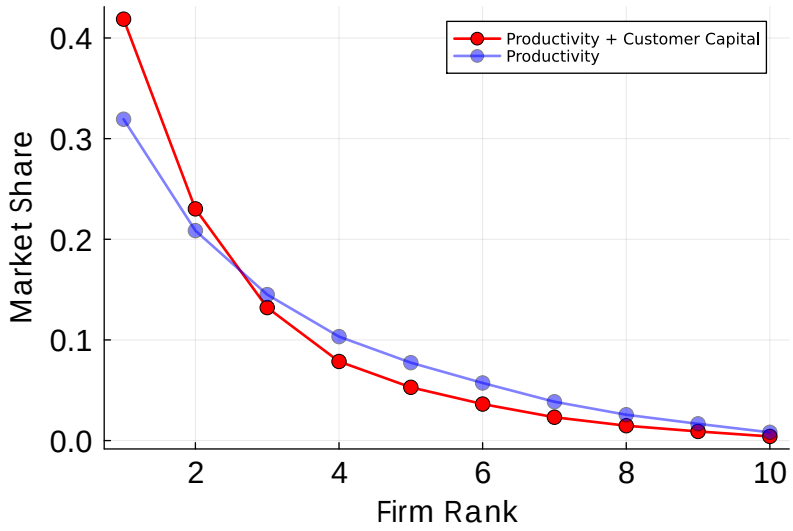


The median half-life in our estimated model is 12.62 periods (indicated by the ver-

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.

Figure 10: Decomposition of Concentration



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 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.

6.2 Productivity: Micro-to-Macro

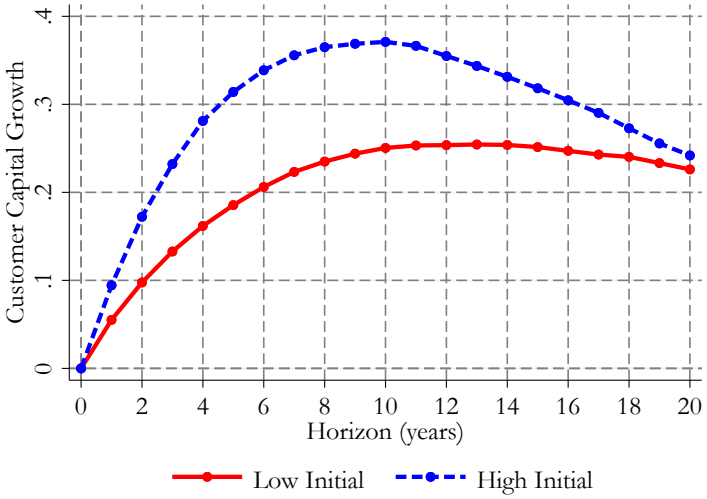
In this section, we study the relationship between productivity shocks, customer capital, and aggregate dynamics. We start by discussing responses at the firm level to shocks and then discuss the transmission to the aggregate. Finally, we compare this framework to existing frameworks of firm and aggregate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

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

Figure 11: Response to Productivity Shock by Initial Customer Capital



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 7: 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

Table 7 quantifies the productivity implications across scenarios. Without customer capital, aggregate productivity is only 0.07. Allocating all customer capital to the productivity leader yields 0.59, over eight times higher. Our calibrated model achieves 0.16, slightly above the no-size-incentive benchmark of 0.15: the productivity-enhancing and distortionary incentives roughly offset. However, this balance tilts in more concentrated markets where the size incentive dominates.

Two features of the model are central to these results. First, slow-moving customer capital impedes the transmission of productivity shocks to the aggregate; costly acquisition and strategic distortions both slow reallocation. Second, granularity matters: unlike monopolistic competition models where firm-level shocks wash out, the oligopolistic framework generates rich micro-to-macro transmission through endogenous market

share reallocation.

6.3 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 3. 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 8: 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 8.

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 3.

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

$$\mathbf{V}^* = \int_0^1 V_k dk, \tag{13}$$

where W_k is the expected utility from product group k:

$$V_k = \max_{\eta_{it}(\alpha_t, \mathbf{b}_t)} \mathbb{E}_0 \int_0^\infty e^{-rt} \left[\log A_{kt}(\alpha_t, \mathbf{b}_t) - 1 - \tau \sum_i^I \frac{\eta_i^2}{2}(\alpha_t, \mathbf{b}_t) \right] dt \tag{14}$$

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, $(\alpha, \mathbf{0})$.

$$v^*(\alpha, \mathbf{b}) = \frac{1}{\sigma - 1} \sum_i \left(\bar{s}_i(\alpha) - \frac{1}{I} \right) b_i + \frac{1}{\sigma - 1} \log \frac{\sum_i \exp(a_i)}{I}, \tag{15}$$

where $\bar{s}_i^*(\alpha) = \frac{\exp(a_i)}{\sum_j \exp(a_j)}$. In the planner's payoff as in Equation (15), 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 $\bar{s}_i^* > \frac{1}{I}$) is positive, 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 8. 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 8 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 interact with firm size.[7]

7 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 ques-

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

tion. 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.

References

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.

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

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., 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.

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.

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

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.

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.

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: 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.

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.

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.

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

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*.

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

Appendix

The Appendix contains two sections. Appendix A discusses the theoretical proofs, aggregation, and the firm's dynamic problem. Appendix B discusses the estimation, moment conditions, 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

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

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

$$\hat{\pi}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \bar{\pi} + p(\alpha_i + \beta_i) + q(\alpha_i + \beta_i)^2 + \phi \sum_j (\alpha_j + \beta_j)^2$$

We start by re-writing firm i's dynamic problem, given the competitors follow the strategy $\eta_j(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \bar{\eta} + \epsilon_a \alpha_j + \epsilon_b \beta_j$, for $j \neq i$.

$$\max_{\eta_i(\boldsymbol{\alpha}, \boldsymbol{\beta})} \mathbb{E}_0 \int_0^\infty e^{-rt} \left(\hat{\pi}(\boldsymbol{\alpha}, \boldsymbol{\beta}) - \frac{\kappa}{2} \eta_i(\boldsymbol{\alpha}, \boldsymbol{\beta})^2 \right) dt, \tag{A1}$$

s.t.

$$d\beta_i(\boldsymbol{\alpha},\boldsymbol{\beta}) = \left(\frac{I-1}{I}\eta_i(\boldsymbol{\alpha},\boldsymbol{\beta}) - \frac{I-1}{I}\bar{\eta} + \frac{1}{I}\epsilon_a\alpha_i + \frac{1}{I}\epsilon_b\beta_i - \delta\beta_i\right)dt + \sqrt{\frac{(I-1)}{I}}\nu_b dW^B,$$

and

$$d\alpha_i = -\rho_a\alpha_i dt + \sqrt{\frac{(I-1)}{I}}\nu_a dW^A.$$

From the law of motion for levels, we derive the law of motion for the gaps using the definition:

$$\begin{aligned}
d\beta_i(\boldsymbol{\alpha},\boldsymbol{\beta}) &= db_i - \frac{1}{I}\sum_{j=1}^{I} db_j \\
&= \left(\frac{I-1}{I}\eta_i(\boldsymbol{\alpha},\boldsymbol{\beta}) - \frac{1}{I}\sum_{j\neq i}\left(\bar{\eta} + \epsilon_a\alpha_j + \epsilon_b\beta_j\right)\right)dt + \sqrt{\frac{(I-1)}{I}}\nu_b dW^B \\
&= \left(\frac{I-1}{I}\eta_i(\boldsymbol{\alpha},\boldsymbol{\beta}) - \frac{1}{I}\left((I-1)\bar{\eta} + \epsilon_a\sum_{j\neq i}\alpha_j + \epsilon_b\sum_{j\neq i}\beta_j\right)\right)dt + \sqrt{\frac{(I-1)}{I}}\nu_b dW^B \\
&= \left(\frac{I-1}{I}\eta_i(\boldsymbol{\alpha},\boldsymbol{\beta}) - \frac{1}{I}\left((I-1)\bar{\eta} - \epsilon_a\alpha_i - \epsilon_b\beta_i\right)\right)dt + \sqrt{\frac{(I-1)}{I}}\nu_b dW^B
\end{aligned}$$

where in the first equation we use the definition of gaps; in the second equation we plug in the law of motions for competitors; in the third equation we rearrange summations; and in the last equation we use $\alpha_i + \sum_{j\neq i}\alpha_j = 0$.

We write out the HJB equation from this problem. For notation simplicity, we will drop the state variables from the equation since all functions depend on the same values of state variables.

$$\begin{aligned}
rV = \max_{\eta}\hat{\pi} - \kappa\frac{\eta^2}{2} + \frac{\partial V}{\partial\beta}\left(\frac{I-1}{I}\eta - \frac{I-1}{I}\bar{\eta} + \frac{1}{I}\epsilon_a\alpha + \frac{1}{I}\epsilon_b\beta - \delta\beta\right) - \frac{\partial V}{\partial\alpha}\rho_a\alpha \\
+ \frac{(\nu^b)^2}{2}\frac{I-1}{I}\frac{\partial^2 V}{\partial\beta^2} + \frac{(\nu^a)^2}{2}\frac{I-1}{I}\frac{\partial^2 V}{\partial\alpha^2}.
\end{aligned}$$

Take the first-order condition:

$$\eta = \frac{I-1}{I\kappa}\frac{\partial V}{\partial\beta}$$

We guess $V = \frac{\kappa I}{I-1}\left(\bar{V} + \bar{\eta}b\beta + \bar{\gamma}\alpha + \epsilon_b\beta^2 + \epsilon_a\alpha\beta + \tilde{\epsilon}_a\alpha^2\right)$ We now guess and verify that

the value function follows the quadratic form. More precisely, we guess that,

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

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

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:

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

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:

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

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

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

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:

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

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

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

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

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

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:

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

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

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

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

This equation has real solutions when

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

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

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

and

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

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:

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

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

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 pro-

ductive 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, $(\alpha, \mathbf{0})$:

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

A.6 Aggregation and Household Problem

The household makes static consumption and labor decisions and a dynamic savings decision. The household has flow utility $\log \mathbf{C}_t - \mathbf{L}_t$ and holds assets ω_t, with the budget constraint

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

where aggregate consumption aggregates across product groups via Cobb-Douglas: $\log \mathbf{C_t} = \int_0^1 \log \frac{C_{kt}}{B_{kt}} dk$, with C_{kt} as defined in equations (2) and (3). The wage is normalized to 1.

Market Share Derivatives. 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 coefficient ζ_i captures how responsive a firm's market share is to changes in its customer capital. This responsiveness varies with the baseline market share $\bar{s}_i(\alpha)$ and the elasticity of substitution σ. The sigmoid function underlying market shares yields an inverse U-shaped relationship: firms with intermediate market shares exhibit the highest responsiveness to customer capital investments, while both dominant and marginal firms show diminished responsiveness.

Multi-product Extension. For the empirical analysis, we model multi-product firms with constant elasticity of substitution across products and decreasing returns in production (as in Hottman et al., 2016). Firm-level consumption is

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

where $\sigma_u > \sigma$ implies constant markups across product lines, and production at the firm-group level exhibits decreasing returns to scale: $c_{ikt} = e^{a_{ikt}/(\sigma-1)} l_{ikt}^{\omega_k}$.

Aggregate Productivity. From the equilibrium characterization, $(\mathbf{a}, \mathbf{b})$ follows a joint Normal distribution. Pooling across industries, aggregate productivity can be written as:

$$\log \mathbf{A} \approx \frac{1}{2}\sigma^2 + \log I - \frac{1}{2}\log(1+R)$$

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 $(\sigma_{aT}^2, \sigma_{bT}^2, \sigma_{aP}^2, \sigma_{aP}^2)$. 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^a_{jt} is independent of past values. We denote the stationary value of the persistent variance as $Var(a^P)$. Plugging the stationary variance, we have

$$\Sigma^a_0 = \frac{2}{1+\rho_a}\sigma^2_{aP} + 2\sigma^2_{aT}$$

Now we move on to $k = 1$:

$$\begin{aligned}
\Sigma^a_0 =& Cov\left(a^P_{jt} - a^P_{jt-1} + \zeta^a_{jt} - \zeta^a_{jt-1}, a^P_{jt-1} - a^P_{jt-2} + \zeta^a_{jt-1} - \zeta^a_{jt-2}\right) \\
=& Cov\left(a^P_{jt} - a^P_{jt-1}, a^P_{jt-1} - a^P_{jt-2}\right) + Cov\left(\zeta^a_{jt} - \zeta^a_{jt-1}, \zeta^a_{jt-1} - \zeta^a_{jt-2}\right) \\
=& Cov\left(a^P_{jt} - a^P_{jt-1}, a^P_{jt-1} - a^P_{jt-2}\right) - \sigma^2_{aT} \\
=& Cov\left(a^P_{jt}, a^P_{jt-1}\right) - Var\left(a^P\right) - Cov\left(a^P_{jt}, a^P_{jt-2}\right) + Cov\left(a^P_{jt-1}, a^P_{jt-2}\right) - \sigma^2_{aT}
\end{aligned}$$

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

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

$$Cov\left(a^P_{jt}, a^P_{jt-2}\right) = \rho^2_{aa} Var(a^P_{jt-2}) = \rho^2_{aa} Var(a)$$

Plugging these values we have

$$\Sigma^a_1 = -\frac{1-\rho_a}{1+\rho_a}\sigma^2_{aP} - \sigma^2_{aT}$$

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.

$k = 0$

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

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 $t-k$.

$k = 1$

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

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

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

B.2 Moment Conditions for Elasticity Estimation

Let

$$a_{it}^*(\omega) = \Delta p_{it} - \frac{\omega}{1+\omega} \Delta s_{it}, \qquad b_{it}^*(\sigma) = \Delta s_{it} + (\sigma - 1) \Delta p_{it},$$

and define innovation residuals

$$\nu_{a,it} = a_{it}^* - \rho_{aa} \, a_{i,t-1}^*, \qquad \nu_{b,it} = b_{it}^* - \rho_{ba} \, a_{i,t-1}^* - \rho_{bb} \, b_{i,t-1}^*.$$

Let $Z_{i,t-1}$ denote instruments measurable at $t-1$ (e.g. $L^2\Delta p_{it}, L^2\Delta s_{it}, L^3\Delta p_{it}, L^3\Delta s_{it}$). The dynamic restrictions require

$$\mathbb{E}[Z_{i,t-1} \, \nu_{a,it}] = 0, \tag{A2}$$

$$\mathbb{E}[Z_{i,t-1} \, \nu_{b,it}] = 0, \tag{A3}$$

$$\mathbb{E}[\nu_{a,it} \, \nu_{b,it}] = 0. \tag{A4}$$

Equations (A2)–(A3) enforce that the quasi-differenced demand and cost shocks are orthogonal to lagged information, while (A4) imposes no contemporaneous correlation between the structural innovations. Together with the variance restrictions

$$\mathbb{E}\left[(a_{it}^*)^2\right] = \sigma_a^2, \tag{A5}$$

$$\mathbb{E}\left[(b_{it}^*)^2\right] = \sigma_b^2, \tag{A6}$$

(with σ_a^2 and σ_b^2 as nuisance parameters), this yields five moments to identify $(\sigma, \omega, \rho_{aa}, \rho_{ba}, \rho_{bb})$.

B.3 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.3.1 Productivity Persistence

For productivity, we estimate:

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

where $\Delta \epsilon_{jt} = \Delta \nu^a_{jt} + \Delta \zeta^a_{jt}$. The moment conditions are:

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

B.3.2 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 $\Delta \eta_{jt} = \Delta \nu^b_{jt} + \Delta \zeta^b_{jt}$. The moment conditions are:

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

B.3.3 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:

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

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:

$$\text{Baseline:} \quad a_{jt} = \rho_{aa} a_{jt-1} + \zeta_j + \nu_{ajt} \tag{A7}$$

$$\text{Alternative:} \quad a_{jt} = \rho_{aa} a_{jt-1} + \rho_{ab} b_{jt-1} + \zeta_j + \nu_{ajt} \tag{A8}$$

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:

$$\text{First Stage:} \quad \Delta a_{jt} = \pi_1 Z_{jt} + \eta_{jt} \tag{A9}$$

$$\text{Second Stage:} \quad \Delta b_{jt} = \rho_{ba} \Delta \hat{a} jt - 1 + \rho bb \Delta b_{jt-1} + \epsilon_{jt} \tag{A10}$$

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:

$$\Delta b_{jt} = \rho_{ba}(\phi(s_{jt-1}))\Delta a_{jt-1} + \rho_{bb}(\phi(s_{jt-1}))\Delta b_{jt-1} + \Delta \nu_{bjt} \tag{A11}$$

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.4 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.4.1 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:

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

B.4.2 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:

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

- Treatment of initial conditions

B.4.3 Results

Our second-stage estimates are:

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

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.5 Identification Discussion

The separate identification of variance components comes from a few forces. First, transitory shocks $(\sigma_{aT}^2, \sigma_{bT}^2)$ primarily identified by $k = 1$ autocovariances. Second, persistent shock variances $(\sigma_{aP}^2, \sigma_{bP}^2)$ identified by decay rates at higher lags. Third, cross-

persistence ρ_{ba} helps identify relative contribution of productivity versus customer capital shocks.

B.6 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

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).

MORE ON AB:

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

$$\Delta a_{jt} = \rho_{aa}\Delta a_{jt-1} + \Delta \nu_{ajt} \tag{A12}$$

$$\Delta b_{jt} = \rho_{ba}\Delta a_{jt-1} + \rho_{bb}\Delta b_{jt-1} + \Delta \nu_{bjt} \tag{A13}$$

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:

$$E[a_{jt-s}\Delta\nu_{ajt}] = 0 \quad \text{for } s \geq 2 \tag{A14}$$

$$E[b_{jt-s}\Delta\nu_{bjt}] = 0 \quad \text{for } s \geq 2 \tag{A15}$$

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,

$$E[\nu_{ajt}|a_{j1}, ..., a_{jt-1}, \zeta_j] = 0 \tag{A16}$$

$$E[\nu_{bjt}|b_{j1}, ..., b_{jt-1}, a_{j1}, ..., a_{jt-1}, \xi_j] = 0, \tag{A17}$$

and no serial correlation in the error terms:

$$E[\nu_{ajt}\nu_{ajt-s}] = 0 \quad \text{for } s \geq 1 \tag{A18}$$

$$E[\nu_{bjt}\nu_{bjt-s}] = 0 \quad \text{for } s \geq 1. \tag{A19}$$

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.7 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.