Concentrating on Customers *

Alex Blumenfeld
UC Berkeley

Christina Patterson
Chicago Booth & NBER

Joseph Vavra
Chicago Booth & NBER

November 2026

Abstract

A growing literature documents rising sales concentration at "superstar" firms, but aggregate market share need not measure how important a firm is to its customers. Using trillions of dollars of card transaction data covering the near universe of firms in a broad set of U.S. consumer-facing retail and service categories, we measure each firm's "effective" market share, defined as its share of category spending among the particular customers it serves. We find that effective shares rise only weakly with conventional market shares. As firms grow, they deepen relationships with existing customers but also reach more marginal customers with lower effective shares. This distinction between conventional and effective shares is economically meaningful: effective shares better predict the persistence of customer relationships and switching to new competitors following entry. Prominent sources of recent firm growth tilt toward customer reach rather than depth, so aggregate sales concentration has risen substantially over the last decade without a comparable rise in effective shares. Interpreted through the lens of a variable-markup model, this divergence suggests that rising concentration has not led to a similar rise in size-based product-market power.

*We thank Violet Hamlin, Zifan Huang and Luisa Pomarici for excellent research assistance, and seminar participants at Chicago Booth, Columbia, UIC, UVA, University of Minnesota, Ohio State, NBER Summer Institute Macro-Productivity, the Empirical Macro Workshop-Phoenix, the Federal Reserve Banks of Philadelphia and San Francisco, and the BLS for helpful comments.

1 Introduction

A growing literature documents rising sales concentration at "superstar" firms. But aggregate market share does not necessarily measure how important a firm is to its customers. Consider two firms that each account for 10 percent of category sales. One could capture all spending from 10 percent of customers, while the other captures 10 percent of spending from all customers. The two firms have identical market shares, but very different relationships with their customers. More generally, a firm could grow large by reaching many customers without becoming especially important to any of them, while a firm with small total sales might still account for a large share of its regular customers' spending.

This paper takes this distinction to the data using transaction-level card data covering the near-universe of firms in a broad set of consumer-facing retail and service categories. We construct what we term "effective market shares," defined as a firm's sales-weighted average share of category spending among the customers it serves. Effective shares thus capture information about a firm's importance to its customers, and comparing effective shares with conventional market shares reveals a sharp distinction between growing large and becoming more important to customers.

We document this through four findings. First, differences in firm size overstate differences in firms' importance to customers: large differences in conventional market shares translate into much smaller differences in effective shares. Second, effective shares capture durable customer relationships: effective shares measured in one year predict subsequent customer retention and switching following competitor entry substantially better than conventional market shares do. Third, the weak relationship between effective and conventional market shares reflects a tension between customer reach and depth as firms grow: deepening relationships with existing customers raises effective shares, whereas acquiring new customers lowers them because new customers typically have shallower relationships with the firm. Fourth, the distinction between conventional and effective shares matters for concentration trends. Prominent sources of firm growth, including geographic expansion and e-commerce, disproportionately expand customer reach rather than depth. Reflecting these patterns, the largest firms have substantially increased their conventional market shares over the last decade without comparable increases in their effective shares.

How can we interpret the difference between conventional and effective shares? In variable-markup models with representative households such as Atkeson and Burstein (2008), larger firms face less elastic demand because of their higher aggregate market shares. We show that when these same models are generalized to incorporate persistent customer heterogeneity, the size-dependent component of demand elasticity depends on the effective rather than conven-

tional market share.[1] Thus, although our empirical findings do not rely on any particular model, effective shares have a natural interpretation as measures of size-based market power.[2] Taken together, our findings suggest that conventional market shares overstate both differences in size-based market power across firms and the growth of market power over time.

Measuring effective shares requires transaction data that link customers to firms and cover a large enough share of firms to observe customers' spending baskets.[3] We use credit- and debit-card transaction data from a large card-payment processor, covering trillions of dollars of U.S. spending each year. We focus on consumer-facing retail and service categories, where we observe spending for the near universe of U.S. firms and establishments. Transactions include the dollar amount, anonymized card identifiers, merchant identifiers, establishment location for in-person transactions, and a card-present flag that we use to distinguish online from in-person transactions. Beginning in 2021, we link cards to TransUnion records, which allows us to aggregate spending across multiple cards per individual and provides demographics like age, income and address. Because we observe spending at particular merchants but not on the specific products or UPCs they sell, our results are informative about customer demand and size-based market power at the merchant level rather than about demand for specific products or manufacturers within a store.

Our first finding compares effective shares to conventional market shares. We begin nationally, since national market shares have received substantial attention in past empirical work. Effective shares are substantially compressed relative to national market shares: large differences in conventional market shares correspond to much smaller differences in effective shares, so the implied firm "size" distribution is much less skewed when measured using effective shares.[4] Compression also holds when we compare effective shares to local market shares: smaller firms tend to have effective shares above their local market shares, while larger firms have effective shares below them. This reflects two forces: (1) within-market segmentation, which raises effective shares for smaller firms (e.g., a neighborhood coffee shop capturing a large share of its regular customers' spending), and (2) out-of-market spending, which lowers effective shares for larger firms whose customers shop beyond local boundaries (e.g., purchas-

[1] This same result holds in a nested-logit discrete-choice model building on Mongey and Waugh (2025), and the qualitative logic extends to broader classes of demand systems exhibiting Marshall's second law.

[2] Effective shares capture the margin through which size affects demand elasticity in these models, but they do not in general identify elasticities in levels, as firms with identical effective shares may face very different substitution environments.

[3] To reduce the role of discrete choice and infrequent purchases, we pool each customer's transactions over a year and also show similar results using "synthetic customers" constructed by pooling demographically similar individuals living in the same small geographic area.

[4] This is relevant for work on markup distortions and misallocation, in which the firm size distribution is a key input (e.g., Baqaee and Farhi (2020), Edmond, Midrigan and Xu (2023), Afrouzi, Drenik and Kim (2025)).

ing online). Robustness to geographic aggregation is important because conventional market shares and their trends depend meaningfully on the level at which markets are defined (Autor et al. (2020), Rossi-Hansberg, Sarte and Trachter (2021), Autor, Patterson and Van Reenen (2023), Smith and Ocampo (2025)).[5] More broadly, effective shares sidestep the need to impose any market geography at all, instead allowing geographic scope to be revealed by customers' actual shopping behavior.[6]

Our second finding is that *current* effective shares predict *future* customer behavior, and do so better than conventional market shares. Effective shares mechanically reflect realized spending concentration within a measurement year, but we provide two pieces of evidence that effective shares predict behavior beyond that year. First, customers with higher past spending shares at a firm are substantially more likely to continue shopping there in future years. Second, customers with higher effective shares are substantially less likely to switch to a new entrant after it appears. These patterns indicate that effective shares capture persistent customer-firm relationships that continue to matter as new alternatives become available, rather than firm-level scale or transitory variation in realized purchases.[7]

These patterns are also consistent with the market-power interpretation discussed above, in which higher effective shares correspond to lower demand elasticities.[8] External evidence points in the same direction: Einav, Guido and Klenow (2026) show that bilateral customer-overlap measures related to effective shares closely track diversion ratios estimated using standard IO approaches. Argente, M'endez and Van Patten (2026) find that a customer-defined measure of retailer size close in spirit to effective shares predicts markup responses to an exogenous reduction in import competition.

Our third set of results shows that changes in customer relationships as firms grow help explain the weak relationship between effective and conventional shares. Conventional market shares reflect how much a firm sells, but effective shares depend on who those sales come from. As firms grow, they both deepen relationships with existing customers and expand their

[5]Becker et al. (2026) show that trade cost reductions in a spatial model can increase national concentration while reducing local concentration and markups. Our results emphasize a complementary point: even within a given local market definition, conventional shares need not capture firms' importance to the customers they serve.

[6]Even flow-based approaches such as Batch et al. (2024) assign a common geographic scope to all firms in a location-sector cell. In practice, the relevant scope may differ across firms and shoppers in the same location, as when two nearby grocery stores draw customers from different catchment areas or when customers near a boundary shop across it.

[7]In a discrete-choice environment with a single realization, for example, all firms would have effective shares of one regardless of underlying consumer preferences.

[8]We do not have the price and cost data necessary for traditional demand estimation. Moreover, markup estimates require assumptions about production technologies, demand, and cost measurement, and have generated substantial debate; see, e.g., Steve Bond (2021), De Loecker, Eeckhout and Unger (2020), Traina (2018), Syverson (2019), Benkard, Miller and Yurukoglu (2025), and Foster, Haltiwanger and Tuttle (2022).

reach to new customers, with opposing implications for effective shares. Deepening existing relationships raises effective shares by construction. In contrast, customer acquisition tends to lower effective shares because new customers are more marginal: on average, their effective shares are about 35% lower than those of incumbent customers. The overall weak relationship between effective and conventional shares reflects these two opposing forces.[9]

Our fourth set of results shows that this distinction matters for interpreting recent trends in firm size and the rise of "superstar" firms. We first show that two prominent growth channels in recent years tilt toward reach rather than depth and are therefore associated with particularly muted changes in effective shares. Geographic expansion by multi-establishment firms (Hsieh and Rossi-Hansberg (2023)) raises national market share by attracting customers in new markets who previously lacked access to the firm. We find that these "new-market" customers have lower effective shares than the firm's customers in established markets. Growth in online sales expands firm reach but also exposes customers to a broader choice set: customers who purchase only online have lower effective shares than customers who purchase in person at the same firm, and firms with a higher share of online shoppers see a weaker relationship between growth in conventional and effective shares.

We then show that the distinction between effective and conventional market shares matters more broadly for longer-run trends.[10] Over the last decade, superstar firms gained substantial aggregate market share, but the share of their customers' spending that they receive grew much less. Interpreted through the lens of the variable-markup framework discussed above, this pattern suggests that rising concentration has not led to a comparable rise in size-based product-market power.[11]

Income might affect both the concentration of customers' spending across firms and their average price sensitivity, so the two forms of heterogeneity may be related in principle, but

Our paper is closely related to a recent literature emphasizing the role of customer heterogeneity in firm growth and market power. The theoretical motivation builds on Mongey and Waugh (2025), who show that customer heterogeneity changes the mapping between firm size and market power because firms' demand elasticities depend on the customers they serve. Their focus is on how changes in income heterogeneity over the business cycle shape firms' price sensitivity. Our focus is complementary: we measure persistent heterogeneity in shop-

[9]We find that customer growth is more important than sales per customer in explaining total firm sales growth, consistent with other recent research we discuss below. However, incumbent customers outnumber new customers at any point in time, and their higher-intensity relationships receive more weight in the effective-share calculation, so the net relationship is positive but small.

[10]Most of our analysis uses the linked TransUnion sample, which has the highest-quality cross-sectional coverage, but begins in 2021. To study longer trends, we complement it with a historical, less granular version of the transaction data available back to 2014.

[11]Size may matter beyond the product market, for example through monopsony or lobbying power.

ping patterns across firms, which can arise from a variety of factors such as geography, tastes and shopping technology. The two margins may be related, since income might affect both the concentration of customers' spending across firms and their average price sensitivity, but empirically our main patterns are robust to controlling for income and other demographic characteristics, and to applying elasticity adjustments motivated by their framework. This suggests that persistent customer-firm shopping relationships are an important source of heterogeneity in their own right. Despite focusing on different margins, we share the conclusion that customer heterogeneity, rather than firm size alone, is central to understanding variation in demand across firms and over time.

Our growth results connect to work emphasizing the customer-acquisition margin. Consistent with Klenow et al. (2022), Argente et al. (2025), and Afrouzi, Drenik and Kim (2025), we find that sales growth is driven in large part by customer acquisition. Our results also complement the finding of Afrouzi, Drenik and Kim (2025) that markups are more closely related to sales per customer than to customer counts. In our framework, sales per customer is not a competing mechanism but an aggregate proxy for customer depth: effective shares measure this depth more directly by using the firm's share of each customer's category spending and the distribution of those shares across customers. Empirically, our measures of customer retention and switching after entry are more strongly related to effective shares than to either customer counts or sales per customer. Our trend results are also consistent with Hubmer and Nord (2026), who show theoretically that increased investment in acquiring customers can generate rising concentration without increasing markups, because intensified competition for customers raises demand elasticities.

Argente, M'endez and Van Patten (2026) construct retailer-specific catchment areas and use Costa Rica's elimination of its de minimis exemption as a shock to import competition. Their customer-defined size measure is close in spirit to effective shares, and their finding that larger catchment-area shares are associated with stronger markup responses provides complementary support for the size-based market-power interpretation of our measure. We differ in two respects: we measure each firm's importance to its customers directly using individual customers' spending baskets rather than geographically defined catchment areas, and we use effective shares to reassess broad cross-sectional and time-series patterns in firm size and concentration rather than to study markup responses in a particular policy setting.

Finally, our measurement is closely related to the customer-overlap statistic in Einav, Guido and Klenow (2026), who use the same transaction data to measure bilateral overlap in customer spending across firms. Their measure captures where a firm's customers shop when they purchase from competitors, providing a scalable way to measure pairwise diversion for applications such as merger analysis. Effective shares instead capture how much of its customers'

spending a firm receives, making them useful for studying customer attachment across many sectors, locations, and periods. We do not view effective shares as a substitute for demand estimation in specific markets where sufficiently rich data are available. Rather, for macro and sector-level questions where systematically estimating demand across many markets and over long periods is typically infeasible, effective shares provide a scalable, customer-based alternative to conventional market shares as measures of size-based market power.

The remainder of the paper proceeds as follows. Section 2 lays out a theoretical framework that motivates our measurement by relating effective shares to size-based market power under customer heterogeneity. Section 3 describes the transaction and credit-record data and our construction of effective shares. Section 4 documents the cross-sectional relationship between effective and conventional shares. Section 5 shows that effective shares better predict customer retention relative to traditional market shares. Sections 6 and 7 study the dynamics of effective size and explore the implications of these patterns for understanding the rise in market concentration. Section 8 concludes.

2 Motivating theoretical framework

This section motivates our effective share measurement by extending the variable-markup mechanism in Atkeson and Burstein (2008) to an environment with persistent customer heterogeneity. In their representative-household nested-CES model, firms with larger aggregate shares of category spending face less elastic demand, because a larger share of substitution away from them occurs through the relatively inelastic between-category margin rather than the more elastic within-category margin. With persistent heterogeneity in customers' relative demand across firms, the same mechanism operates customer by customer: demand is less elastic among customers for whom the firm accounts for a larger share of category spending. Aggregating these customer-level elasticities implies that the size-dependent component of a firm's demand elasticity depends on its effective share.

2.1 Customer heterogeneity and effective shares

Consider an economy with differentiated firms grouped into categories. Customers substitute across firms within a category with elasticity η and across categories with elasticity θ, where $\eta > \theta$. We focus on one category and suppress its index, indexing customers by i and firms by j. Relative to the standard nested-CES benchmark, we allow persistent customer–firm-specific demand shifters, a_{ij}, which may reflect geography, tastes, product availability, shopping technology, or other sources of persistent heterogeneity in where customers shop.

Let ω_{ij} denote firm j's share of customer i's spending in the category:

$$\omega_{ij} \equiv \frac{p_j q_{ij}}{\sum_{k \in \mathcal{F}} p_k q_{ik}} = \frac{a_{ij} p_j^{1-\eta}}{\sum_{k \in \mathcal{F}} a_{ik} p_k^{1-\eta}}, \tag{1}$$

where $\mathcal{F}$ denotes the set of firms in the category. The second equality follows from the nested-CES demand system derived in Appendix A2. Customer-specific demand shifters therefore generate persistent differences in the shares of category spending that customers allocate to particular firms. The distinction between conventional and effective shares arises only when aggregating these customer-level spending shares across customers.

In the same appendix, we show that customer i's elasticity of demand for firm j is given by

$$\varepsilon_{ij} = \eta - (\eta - \theta)\omega_{ij}. \tag{2}$$

Equation (2) is the customer-level analogue of the standard Atkeson–Burstein mechanism. When firm j accounts for a small share of customer i's category spending, substitution away from the firm occurs primarily toward other firms within the category, so demand is relatively elastic. When firm j accounts for a large share of customer i's category spending, substitution away from the firm places greater weight on the less elastic between-category margin. Demand from that customer is therefore less elastic.

To aggregate across customers, let

$$\alpha_{ji} \equiv \frac{p_j q_{ij}}{\sum_{i'} p_j q_{i'j}} \tag{3}$$

denote customer i's share of firm j's sales. Because firm j charges a common price across customers, its demand elasticity is the sales-weighted average of customer-level elasticities. We therefore obtain

$$\varepsilon_j = \eta - (\eta - \theta) \sum_i \alpha_{ji} \omega_{ij}. \tag{4}$$

Define firm j's effective share as its sales-weighted average share of category spending among the customers it serves:

$$\omega_j^{\text{eff}} \equiv \sum_i \alpha_{ji} \omega_{ij}. \tag{5}$$

Equation (4) can therefore be written as

$$\varepsilon_j = \eta - (\eta - \theta)\omega_j^{\text{eff}}. \tag{6}$$

For given η and θ, a firm's demand elasticity declines with its effective share. Under Bertrand pricing, this implies that markups increase with effective share.

Importantly, the standard aggregate-market-share result in Atkeson and Burstein (2008) is a special case of this framework. If customers do not differ in their relative demand for firms (e.g., $a_{ij} = a_j$ for all i), then[12]

$$\omega_{ij} = \Omega_j \qquad \forall i, \tag{7}$$

where

$$\Omega_j \equiv \frac{\sum_i p_j q_{ij}}{\sum_i \sum_{k \in \mathcal{F}} p_k q_{ik}} \tag{8}$$

is firm j's aggregate share of category spending. Since α_{ji} sums to one across customers, it follows that in this special case,

$$\omega_j^{\text{eff}} = \sum_i \alpha_{ji} \omega_{ij} = \Omega_j. \tag{9}$$

In this representative-household case, equation (6) reduces to the familiar Atkeson and Burstein (2008) formula:

$$\varepsilon_j = \eta - (\eta - \theta)\Omega_j. \tag{10}$$

However, in the presence of customer heterogeneity, aggregate market shares are not the correct measure of size for determining this channel. Our empirical analysis focuses on measuring the wedge between ω_j^{eff} and Ω_j.

Although nested CES provides a particularly simple benchmark, the distinction between aggregate market share and customer-level shares is not specific to this setting. In Appendix A2.2, we derive the same dependence of demand elasticity on effective shares in a parsimonious nested-logit discrete-choice model building on Mongey and Waugh (2025), and discuss how related logic extends to broader demand environments. Effective shares thus have a natural interpretation: in standard variable-markup frameworks extended to incorporate customer heterogeneity, they are the relevant notion of firm size that governs demand elasticity and, hence, markups.[13] These are the same frameworks used to interpret rising aggregate concentration as rising market power, but within them aggregate market share is the relevant measure of size only if customers do not differ in their relative demand. Section 5 provides empirical evidence consistent with interpreting effective shares as a measure of size-based market power. At the same time, our empirical results do not depend on this interpretation: they measure and explain

[12]The same aggregate-market-share formula obtains if customers differ in their total demand but not their relative demand (i.e., $a_{ij} = b_i a_j$).

[13]Mapping spending shares to elasticities necessarily relies on assumptions about demand, whether those shares are conventional market shares or effective shares. Relative to the conventional-market-share approach, effective shares relax the additional representative-customer restriction.

variation in effective shares and do not structurally estimate demand elasticities.

3 Data Description and Measurement Details

3.1 Data Description

Our main data sample begins from the universe of U.S. credit and debit card transactions of a major card-payment processor from 2018-2025. The data cover spending at nearly all firms in the consumer-facing retail and service categories that we study, but it captures spending only of this processors' cards. The scope of the data is large: it covers more than 500 million cards and nearly 3 trillion dollars of spending each year. This corresponds to roughly one-fifth of aggregate annual personal consumption expenditures (PCE), excluding housing services.

The unit of observation is an individual transaction between an (anonymized) credit card and a merchant. Each transaction records the exact dollar amount, the time and date, a card identifier, the merchant, and whether the transaction was in-person (card present) or remote (card not present). The firm identifier constructed by the data provider generally links franchised branches with common names (e.g. McDonalds), but does not link all subsidiaries within parent companies (e.g. Whole Foods is separate from Amazon). For in-person transactions, the data record the 5-digit zip code of the store (and typically the exact address), which allows us to identify purchases at different establishments of the same firm. Because nearby establishments of the same firm are close substitutes, we aggregate all establishments of a firm within a ZIP3 and refer to the resulting unit as an "establishment".[14] For remote purchases (e.g. online, phone, recurring purchases), the data include the firm identifier but no physical location. For many remote transactions we further observe if it was an online purchase, a recurring payment initiated by the firm (e.g. monthly subscription fees), or a purchase made over the phone. In the analysis that follows, we refer to all transactions without cards present as "online" transactions.[15]

Each transaction is assigned a four-digit merchant category code (MCC), which captures the type of good or service that the firm provides. The code is assigned when a merchant registers to accept card payments and determines its fee structure and associated card rewards, so is a fundamental characteristic of each transaction. We combine very similar MCC codes into a single MCC category and restrict to customer-facing categories with good coverage of

[14]While we could measure market shares at finer geographies (e.g. ZIP5), it is not clear this is economically relevant, since some firms have many establishments even in narrow geographic areas. Aggregation also avoids concerns about inconsistent measurement of exact addresses.

[15]We use this broad definition since some transactions lack the breakdown and because phone and recurring payments are not obviously economically distinct from online payments.

individual firms, which leaves 33 MCC categories in our analysis sample.[16] Some merchants have multiple MCC codes corresponding to establishments that provide different products (e.g. Costco wholesale vs. Costco gas stations). We assign all of a firm's transactions to its MCC code with the most transactions. The largest categories in our sample are General Merchandise Stores (27% of our sample), Grocery Stores (16%), and Restaurants (9%) (see Appendix Table A4 for details).

Due to the nature of the transaction data, our unit of analysis is the firm rather than the individual product.[17] We therefore interpret consumer choice as occurring over firms within a category. This is a data limitation, but firms are also an economically relevant margin in many retail and service markets. To address the concern that firm-level shares abstract from product-level pricing decisions, we show that our results are robust in categories such as Auto Repair, Health and Beauty Spas, and Restaurants where the firm itself produces and prices what it sells, so there is less role for separate manufacturer markup decisions (see Appendix Table A10).

The raw transactions data are a card level panel with no links across cards owned by the same customer. Beginning in 2021, the panel can be linked to TransUnion credit records, allowing us to group multiple cards belonging to the same person and providing key consumer characteristics including home location, age, and estimated income. The linked sample covers around 135 million individuals and $5.5 trillion of spending, appears broadly representative of the adult U.S. population on a number of demographic characteristics and includes 35% of the spending in the full transaction dataset. Credit records do not capture debit-cards, which is the main source of dropped spending.[18] Our conclusions are robust to constructing effective shares using card-level instead of person-level spending. Because card-level analysis does not require the TransUnion link, it also lets us verify that patterns are very similar when we expand the sample to include debit cards and credit cards that are not linked.[19] In the linked sample, the average person has 1.6 credit cards per year and 2.4 cards over the sample from 2021-2025, while only about 16% of people have at least 3 cards within a year.[20]

[16] We drop categories such as airlines and hotels, where a large fractions of transactions cannot be linked to specific firms because they are made through intermediary websites. See Appendix table A13 for a detailed list of categories.

[17] Neiman and Vavra (2023) document related but distinct patterns at the UPC level for grocery-store products in Nielsen data: household concentration diverges from aggregate concentration, although aggregate product concentration declines as new UPCs are introduced.

[18] Of the cards not linked, 82% are debit cards. The remainder arise because not all credit-issuers report to TransUnion.

[19] Specifically, the relationship between firm-level market shares and effective shares is very similar whether effective shares are calculated using all cards or only credit cards. See Appendix Table A10 for the details.

[20] Individuals with multiple cards may use certain cards for certain types of spending due to rewards incentives. Rewards incentives vary across MCC codes but not usually across merchants within MCC codes. Since our analysis focuses on spending concentration within categories as measured by MCC codes, cross-MCC card substitution is unlikely to matter for our conclusions.

Starting from the linked sample, we impose six restrictions that focus our sample on consumer cross-firm substitution patterns. First, we drop business cards. Second, we restrict the sample to transactions that are either online or occur within 100 miles of the consumer's home ZIP code, which excludes idiosyncratic travel spending (e.g., vacation purchases) and focuses on stores plausibly relevant for regular shopping.[21] Third, we remove transactions routed through payment facilitators, intermediaries, or peer-to-peer systems (e.g., PayPal, Toast, Venmo), because these records often obscure the underlying merchant identity.[22] Fourth, we drop customer-years for customers whose home address in TransUnion changes ZIP5 within the year. Fifth, we keep only category-ZIP5 pairs with at least $450K of total spending in the linked sample summed over 2021-2025. Sixth, to limit potential sampling error in measured effective shares, we keep only firms with at least 200 customers nationally in our linked subsample in a year.[23] We refer to the resulting sample as our "Primary" sample.

Figure 1 shows that card spending in our data aligns well with Census firm-level data, suggesting that our sample captures the distribution of spending across firms. In particular, the national sales share of the top 4 firms (CR4) in each category is close to the same measure for similar categories in the Census.[24] This holds in the raw card data as well as in our more restricted Primary sample. The most notable deviation is for "Fast-Food Restaurants", because the Census treats franchises of the same chain as separate firms while our data combines all franchises of a firm. Our measure is likely more relevant for understanding customer demand and size-based market power, since consumer shopping choices, pricing and product assortment are often determined at the firm and not at the franchise level.

In addition to our primary dataset, we also utilize a less granular "historical panel" that goes back to 2014 to study trends in spending concentration. Its main benefit is longer time coverage, since our primary sample starts in 2021. This historical panel has two important limitations. First, it cannot be linked to TransUnion, so the unit of observation is the card and not the person. Second, it records each card's total spending in each zip code but disaggregates this spending by firm only for a subset of relatively large "named firms". Spending at named firms is identified separately, while the remainder is grouped together as spending at "other

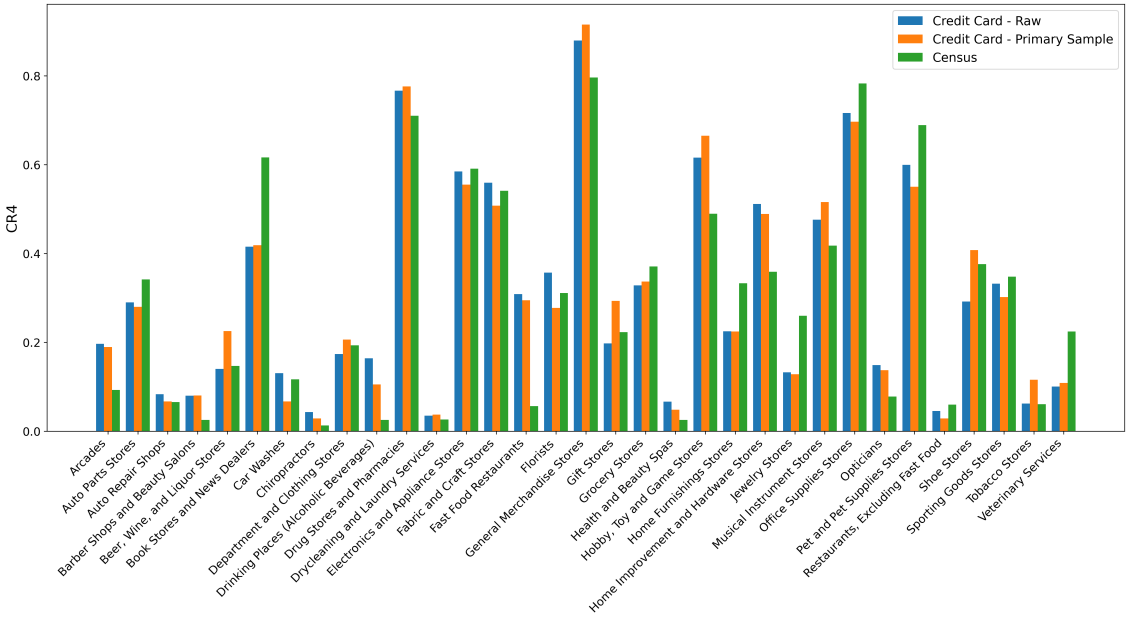
[21] Since this is a small share of spending, this restriction is unimportant in practice.

[22] Transactions made through Apple and Google Pay are included, because we can see the firm identifier for these payments. In our final subsample of 33 categories, 8.1% of all transactions are at payment facilitators, a number that rose from around 4% in 2018 to 8% in 2022. Payment facilitators are most prevalent in barber shops and restaurants. Excluding these transactions biases our results only to the extent that these transactions are unevenly distributed across firms or locations.

[23] For establishment level analysis we require 200 customers at the establishment for inclusion.

[24] Appendix Table A13 shows the crosswalk from NAICS codes to MCC categories. NAICS and MCC-based categorizations do not perfectly align, so we should not expect an exact match even if spending for every individual firm matches exactly.

Figure 1: Benchmarking Card Data to Census Data



Notes: This figure plots the 2022 four-firm concentration ratio (CR4), the sum of the sales shares of the four largest firms, by category in three data sources: the raw card transactions sample, the Primary sample (credit card data linked to TransUnion), and U.S. Census data.

unnamed firms". We thus see a card's *total* spending at unnamed firms but not its spending at any specific unnamed firm. This means that the data is well-suited for studying spending trends at larger, named firms but not for studying smaller firms.[25] We restrict our analysis to categories with at least 10 named firms and where at least twenty five percent of total spending occurs at named firms in every year. The number of named firms varies substantially across categories, but the median category has 69, while we only need to reliably capture only the top four firms. In the years where the two panels overlap, effective shares and market shares are similar in our primary and historical panels, which suggests that the historical panel captures the same underlying patterns as the primary data. We discuss this benchmarking further in Section 7.

3.2 Measuring Effective Shares

Using the Primary sample, we construct empirical analogues of the spending shares defined in Section 2. Because those definitions are static, we construct each customer's spending shares

[25] As defined in Section 3.2, the effective share of customer i at firm j depends on i's spending at j and i's *total* category spending at all other firms but not on how that other spending is split across specific firms.

over a calendar year. Suppressing the year index, let x_{ij} denote customer i's total spending at firm j during the year. We define firm j's share of customer i's category spending as

$$\omega_{ij} \equiv \frac{x_{ij}}{\sum_{k \in \mathcal{F}} x_{ik}}, \tag{11}$$

where $\mathcal{F}$ denotes the set of firms in the category. Customer i's share of firm j's sales is

$$\alpha_{ji} \equiv \frac{x_{ij}}{\sum_{i'} x_{i'j}}. \tag{12}$$

Firm j's effective share is then

$$\omega_j^{\text{eff}} \equiv \sum_i \alpha_{ji} \omega_{ij}. \tag{13}$$

Thus ω_{ij} measures the firm's importance within a customer's spending basket, while α_{ji} weights customers by their importance to the firm's sales. Because effective shares are built from customers' own spending baskets within an MCC category, they do not require choosing geographic market boundaries: we do not need to decide whether the relevant market is a ZIP code, an MSA, or the nation as a whole because the geographic scope of each firm is revealed by where its customers actually shop. Online spending is similarly straightforward, since online purchases enter customers' spending baskets regardless of where the seller is physically located. In contrast, conventional market shares require choosing a geographic partition and have no natural way to assign online spending to a firm location.

Although transaction data record spending directly, we observe only a finite number of purchases. This means that measured spending shares might reflect random noise rather than persistent customer–firm relationships. In the extreme, a customer who makes only a single purchase in a category has a share of one at the chosen firm and zero elsewhere, regardless of their underlying preferences. Pooling transactions within the year attenuates the role of these idiosyncratic realizations while preserving the individual-level heterogeneity that distinguishes effective shares from conventional market shares. To attenuate them further, we restrict attention in each category to customers who make at least 5 transactions in that category during the year.[26] The main concern with this filter is selection, since it drops the least frequent shoppers. However, customers with few transactions also tend to have low α_{ji}, so they have little influ-

[26] In Appendix A3 we provide evidence that with a 5+ transaction filter, we can reliably distinguish true variation in ω_j^{eff} from variation generated by random discrete realizations. We also discuss a specific small-sample bias correction.

ence on effective shares even if included. Our conclusions are similar if we restrict to categories where this restriction rarely binds.[27]

We also construct two alternative measures that pool transactions more extensively, which reduces both measurement error and potential selection induced by the minimum-transaction filter. However, each comes with its own tradeoffs. The first pools each individual's transactions over the full five-year period from 2021 to 2025.[28] This gives us many more transactions per person, but it assumes that customer-firm relationships are stable over a long period and precludes annual panel analysis. The second approach constructs "synthetic customers" by pooling transactions across demographically similar individuals in the same location, specifically individuals in the same census block group with similar age, income, and (inferred) car-ownership status. Synthetic customers have many more transactions, but this approach averages over some genuine heterogeneity in customer-firm relationships and it precludes individual-level analysis.[29] In practice, these measures are very similar: the firm-level effective shares calculated using 5-year pooled samples or using synthetic customers have a correlation with our primary measure of 0.95 and 0.92, respectively. Table A1 provides summary statistics for these three samples as well as for an even larger sample that measures spending at the card level and does not TransUnion links. We show below that our main results are robust across these alternative constructions.

4 The Difference between Effective Shares and Market Shares

Our first empirical finding is that differences in conventional market shares substantially overstate how much firms differ in their importance to customers: large differences in conventional market shares correspond to much smaller differences in effective shares. This compression holds whether conventional market shares are measured nationally or locally. We begin with national comparisons.

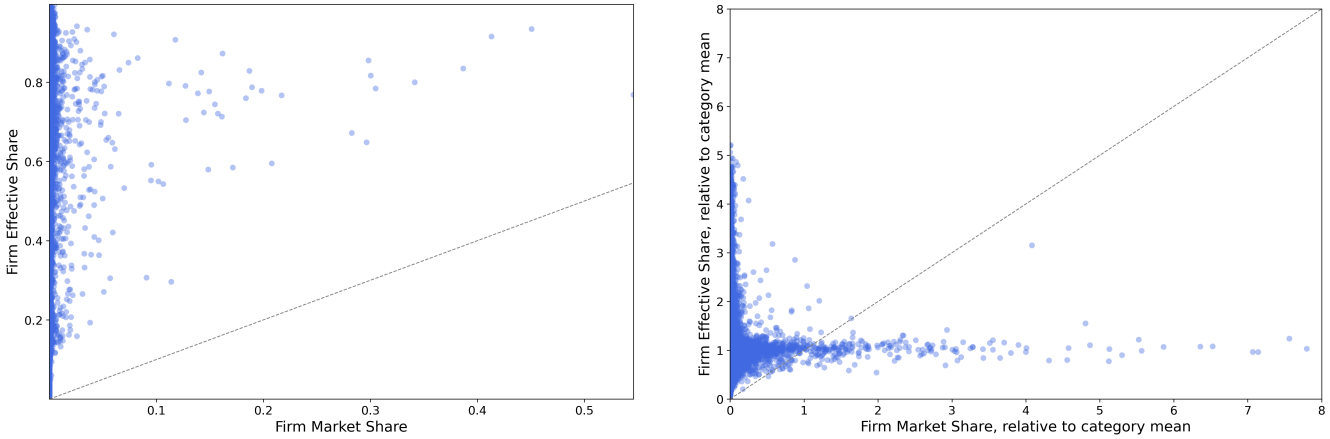
Panel (a) of Figure 2 plots firms' effective shares against their conventional national market shares in 2025. All points lie above the dashed 45-degree line by construction: a firm must have at least as large a share of category spending among its own customers as it does among all customers nationally. The gap is large for most firms, and the relationship between conven-

[27] Appendix Table A3 shows average transactions by category and Table A2 shows the fraction of customers and spending dropped in each category. The restriction drops only 0.3 percent of spending in General Merchandise Stores but 67 percent of spending in Jewelry Stores.

[28] This reduces concerns about selection, since requiring five transactions over five years is a less restrictive filter than requiring five transactions within a year.

[29] With synthetic customers, we no longer impose the five-transaction minimum for individuals, which further reduces concerns about selection.

Figure 2: Effective shares differ from national market shares



(a) Overall

(b) Within Category

Notes: Each dot represents one firm in 2025. In panel (b), each value is divided by the sales-weighted average value in the category, and both axes are truncated at 8. Both shares are defined using annual customer spending in our Primary sample.

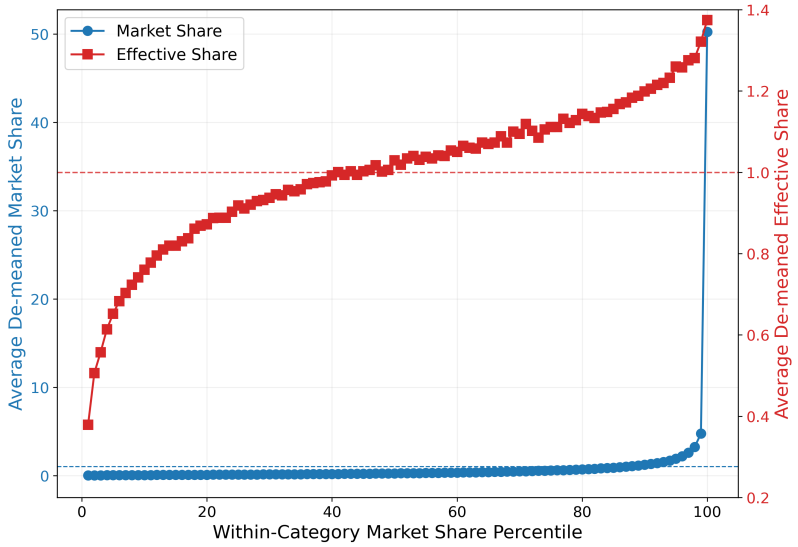
tional and effective shares is weak, with substantial heterogeneity in effective shares among firms with similar national market shares.

Panel (a) pools firms across all categories, but categories differ substantially in their market-share distributions, which might obscure within-category patterns of greater interest.[30] In panel (b), we therefore divide each variable by its sales-weighted category average, so the axes show each firm's conventional market share and effective share relative to the category average. After this normalization, the relationship is even weaker: firms that are large outliers in conventional market share relative to others in their category are generally not large outliers in effective share.

Because the distribution of firm size is highly skewed, most firms have small national market shares and cluster near zero in both panels of Figure 2, which makes patterns among most firms hard to see. Figure 3 provides an alternative visualization of how conventional and effective shares vary across the entire distribution. We rank firms by market shares within their categories and plot, for each percentile, the average market share relative to the category average (left axis) and the average effective share relative to the category average (right axis). Since percentiles are defined using conventional market shares, the market share series is weakling increasing by construction and so the comparison of interest is the comparison with effective

[30]For example, firms in "General Merchandise" and in "Drug Stores and Pharmacies" tend to have much higher national market shares than firms in "Restaurants - Excluding Fast Food".

Figure 3: Effective shares are more uniformly distributed than market shares



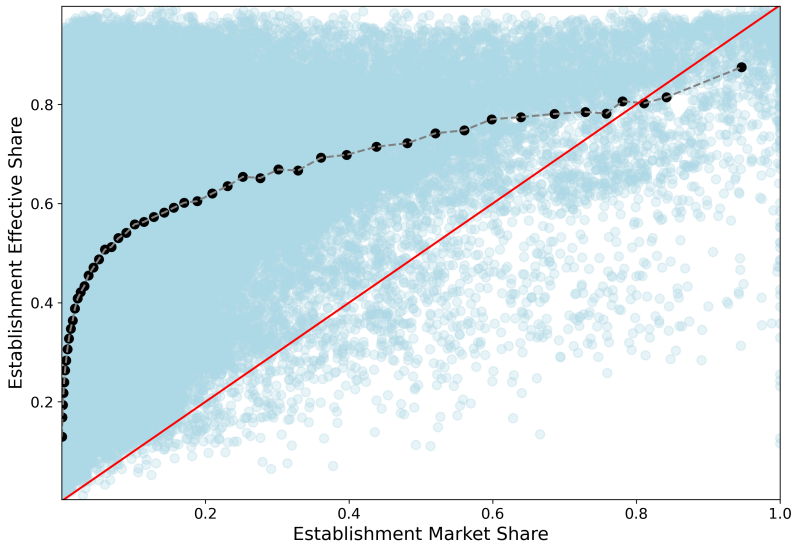
Notes: The x-axis reports within-category market share percentiles in 2025, constructed by ranking firms within each category by their market share and assigning them to up to 100 percentile groups. The left axis (blue) plots the average ratio of market share within each percentile group to the category average market share, so a value of 1 indicates being at the category average. The right axis (red) plots the average effective share within each percentile group relative to the category average effective share. Horizontal dashed lines at 1 show the category mean. The figure aggregates observations across categories after within-category normalization, and firms receive equal weight within percentile groups.

shares. Conventional market shares are extremely right-skewed: the curve is nearly flat through most of the distribution and then rises sharply at the top percentiles, with the largest firms having market shares nearly 50 times their category average. In contrast, effective shares are more compressed and rise more smoothly. The largest firms by market share have effective shares only about 30 percent above the category average.

A natural explanation for the gap between national market shares and effective shares is geographic aggregation. Many consumer markets in sectors like restaurants and grocery stores are local rather than national, so national market shares may be the wrong benchmark. If the gap is mainly a geographic market-definition issue, effective shares should line up much more closely with conventional shares once the latter are computed in the local markets where firms actually compete. We test this by moving to the local level, where an observation is an establishment (a firm–ZIP3 pair, as defined in Section 3.1). For establishment k of firm j, the local market share is the sales at that establishment divided by total category sales in its ZIP3.

Establishment effective shares require no new customer-level object. Because we treat all of a firm's establishments and its online channel as equivalent for a given customer, there is

Figure 4: Effective shares differ from local market shares



Notes: Each blue dot is an establishment (firm-ZIP3 combination) in 2025. The black dots show binscatter relationships with 50 bins, where the bins are constructed using establishment sales weights. For each bin, we plot the mean effective share against the mean market share. The red 45 degree line represents equal conventional and effective shares.

no customer-establishment spending share ω_{ijk}, only the customer-firm share ω_{ij} defined in Section 3.2.[31] We construct establishment k's effective share by re-weighting these customer-firm shares by each customer's share of the establishment's sales.[32] Establishments of the same firm thus differ in their effective shares only through the customers they serve (e.g., the Target in Houston serves different customers than the Target in Chicago).

Figure 4 plots the effective share of each establishment against its local market share. Three patterns stand out. First, effective shares and market shares are much closer in levels than in the national comparison, mainly because local market shares are much higher than national market shares. Second, the relationship remains weak: establishments with similar market shares often differ substantially in their effective shares. Third, the compression documented at the national level remains clearly visible at the local level, appearing as a strong concave relationship. Establishments with small local market shares have effective shares far above their market shares because they capture a much larger share of their customers' spending than of

[31] Results are robust to instead treating each establishment as a distinct entity in customers' baskets, i.e., defining ω_{ijk} as establishment k's own share of customer i's category spending and aggregating with establishment-level sales weights.

[32] Spending at online-only firms is included in customers' total category spending and therefore lowers ω_{ij} at brick-and-mortar firms. Online-only firms have no physical location over which to compute conventional local market shares, so they are not observations in the establishment-level analysis.

the local market as a whole. At the other extreme, establishments with very large local market shares have effective shares somewhat *below* their market shares. A store can appear dominant according to local market shares while accounting for a more modest share of its customers' category spending once purchases outside the local market are counted (either in-person or online).[33]

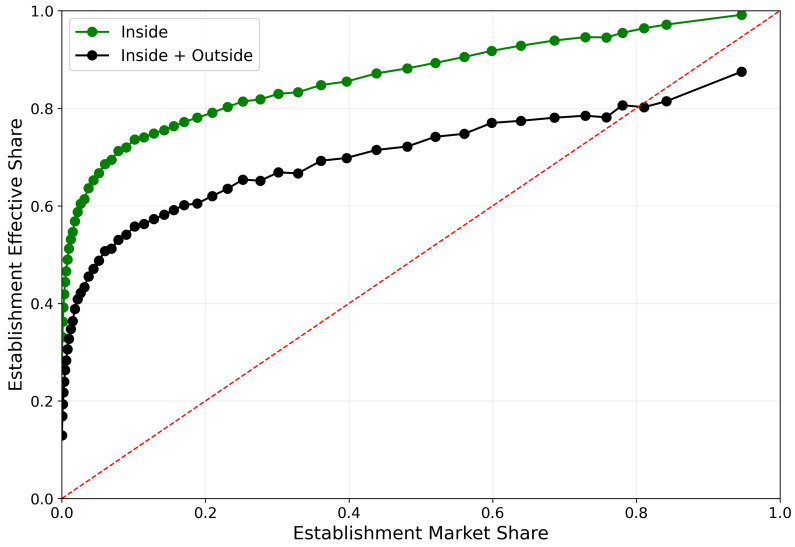
The divergence between conventional and effective shares at the ZIP3 level does not reflect this particular choice of local geography: it persists under alternative definitions of local markets (Appendix Figure A1). Defining conventional market shares using 5-digit ZIP codes raises them relative to ZIP3, and defining them using states lowers them, but in either case effective and conventional shares differ markedly. We also implement a category-specific approach that selects, for each category, the geographic definition among state, ZIP3, and ZIP5 that maximizes the R^2 between effective shares and local market shares. Even under this flexible specification, the two measures diverge substantially. More fundamentally, even establishments in a similar location and category often draw customers from different geographic areas. For example, some stores have parking while others do not, and different stores have different online footprints. A common geographic partition may therefore be too broad for some establishments and too narrow for others. Some definition may align the two measures closely for a particular establishment or subset of establishments, but no common partition is likely to work well for all of them.

To understand the wedge between effective shares and market shares, we decompose it into two forces in Figure 5. First, for each establishment, we recalculate effective shares after restricting each customer's spending basket to purchases made within the establishment's ZIP3. The green series plots these within-ZIP3 effective shares. They necessarily lie above the red 45-degree line, for the same reason that national effective shares do: once the ZIP3 is treated as a closed market, an establishment must account for at least as large a share of spending among its own customers as among all customers in the market. The vertical gap between the green series and the 45-degree line therefore captures within-market segmentation. This force is especially pronounced for small and mid-sized establishments, which often capture a substantial share of their customers' within-ZIP3 spending even when they account for only a small share of local sales.

The second force is spending that customers do outside the ZIP3, either online or in person. The black series in Figure 5 reincorporates this out-of-ZIP3 spending and therefore reproduces the black series from Figure 4. Relative to the green series, the curve shifts downward almost uniformly. The direction of the average shift is unsurprising: when customers shop outside

[33] Appendix Figure A2 shows that the same pattern holds under the alternative pooling choices discussed in Section 3.2.

Figure 5: Understanding the difference between market and effective shares



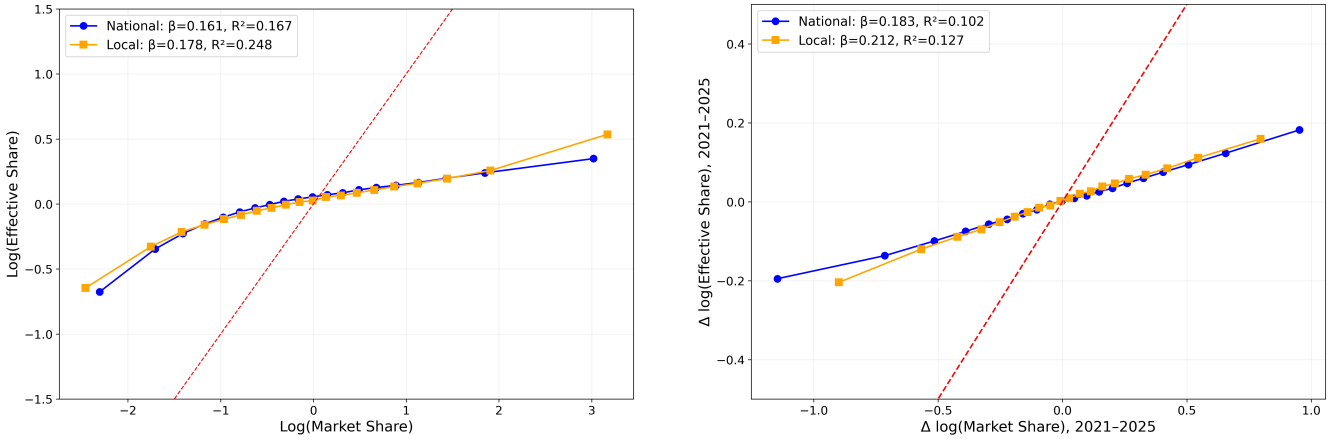
Notes: The black line shows the same binscatter as in Figure 4, where each bin has the same total sales. The green dots show the analogous binscatter but calculating effective shares only using the spending that is done within the ZIP3.

their ZIP3, they often do so at a different set of firms. More notable is that the shift is similar across the local market share distribution.[34] Since out-of-ZIP3 spending reduces effective shares fairly uniformly, the pronounced concavity in the relationship between effective and local market shares arises primarily from within-market segmentation.

The concave relationship between effective and conventional market shares in levels suggests that the relationship may be closer to linear in logs. Panel (a) of Figure 6 shows that this is the case. It plots binscatters of log effective shares against log market shares in 2025. The blue circles show relationships with national market shares for firms (as in Figure 2) and the orange squares show the relationship with local market shares for establishments (as in Figure 4). In both cases, the relationship is close to linear but substantially attenuated relative to the dashed 45-degree line, with a slope of 0.16 nationally and 0.18 locally. A 10 percent higher market share is thus associated with only about a 1.6 percent higher effective share among firms in the same category and a 1.8 percent higher effective share among establishments in the same category and ZIP3. Panel (b) shows that this same pattern holds for within-firm and establishment growth between 2021-2025. Appendix Table A10 shows that cross-sectional compression remains strong when weighting firms by their sales, with a nearly identical local slope and an even smaller national slope. It also shows that slopes are slightly smaller in services and larger

[34]For single-ZIP3 firms with no online presence, reincorporating out-of-ZIP3 spending can only reduce effective shares, since the firm captures none of its customers' spending outside the ZIP3. For firms with sales outside the ZIP3, the effect is ambiguous in principle, because customers may shop at the same firm in other ZIP3s or online.

Figure 6: Relationship between effective shares and market shares in logs



(a) 2025 Levels

(b) Changes from 2021-2025

Notes: This figure presents binscatters of log effective share against log market share, using national data (circles) and establishment data (squares). Panel (a) plots 2025 log levels and panel (b) plots log changes between 2021 and 2025. Variables are residualized on category fixed effects in the national sample and category × ZIP3 fixed effects in the local sample, and points are averages within 20 bins of the residualized log market share (or of its change). In panel (b), the top and bottom 2.5% of changes in log market share are trimmed. In both panels, the coefficients and R^2 of the regressions on the underlying data are reported in the legend, and the dashed 45-degree line is shown for comparison.

in retail.[35]

The results above show that conventional market shares are a poor summary of firms' importance to their customers. This raises two related questions about whether effective shares can be recovered from other observables. First, are high effective shares primarily explained by observable differences in the types of customers firms serve? Appendix Table A10 shows that controlling for the average customer characteristics of the firm such as income, age, and transportation constraints explains little of the variation in effective shares and has little effect on the relationship between conventional and effective shares.[36] Thus, effective shares are not simply a proxy for the income heterogeneity emphasized by Mongey and Waugh (2025). Instead, the importance of a firm within the spending baskets of its customers is a distinct dimension of customer heterogeneity.

Second, when the linked customer-firm spending data needed to construct effective shares

[35]ppendix Figure A3 shows the relationship between log establishment effective shares and the log of establishment market shares separately by category. The attenuated relationship is pervasive, ranging from 0.25 in General Merchandise to under 0.05 in Jewelry Stores and for Chiropractors.

[36]We find a similarly small role for demographics when we control for them at the individual level before aggregating to the firm.

directly are unavailable, can simpler aggregate statistics capture similar information about customer depth? One natural candidate is sales per customer, the depth measure emphasized by Afrouzi, Drenik and Kim (2025). It requires only total firm sales and customer counts, not the distribution of spending across customer-firm pairs, and it is naturally related to effective shares because a customer shifting category spending towards a firm will raise both its effective share and its sales per customer. The two measures are nevertheless distinct: sales per customer can rise without a rise in effective share if it comes from serving customers with larger category spending rather than from capturing a larger share of their spending. In addition, because effective share is an average of customer-level spending shares weighted by each customer's share of the firm's sales, it depends not only on average sales per customer but also on how the firm's sales are distributed across customers.

Appendix Table A5 shows regressions between these variables. Within category, effective shares co-vary strongly with sales per customer: across specifications a 1 percent increase in sales per customer is associated with a 0.5 to 0.6 percent increase in effective shares. However, the partial R^2 from these regressions is around 0.5, so much of the variation effective shares is unrelated to sales per customer. Sales per customer is thus a useful but far from perfect proxy for a firm's effective share.

5 Effective Shares Predict Future Customer Behavior

We next show that customers with higher effective shares at a firm are more likely to continue shopping there in future years and less likely to switch to a new entrant when one appears. Effective shares therefore contain economically meaningful information about customer-firm relationships that persist beyond the year in which the shares are measured and that conventional measures of firm size do not capture.

5.1 Customer Retention

We begin by studying customer retention. Our baseline analysis is at the establishment level, where an establishment k is a firm-by-ZIP3 pair and $j(k)$ denotes the firm that owns it. We focus on establishments because local market shares are closer in level to effective shares than national market shares are, but similar results also hold for firm-level relationships. Among customers who shopped at establishment k in year $t-1$, we define retention as an indicator for whether customer i shops at k again in year t:

$$\text{Retention}_{i,k,t} = \mathbb{1}\{i \text{ shops at } k \text{ in } t\}.$$

Using data from 2021–2025 we then estimate the relationship between retention and the customer-level spending share that enters firm effective-share calculations:

$$\text{Retention}_{i,k,t} = \beta_1 \log \omega_{i,j(k),t-1} + \beta_2 \log \Omega_{k,t-1} + \gamma_{\text{market},t} + \epsilon_{i,k,t}.$$

Here $\omega_{i,j(k),t-1}$ is the share of customer $i's$ category spending at firm $j(k)$ and $\Omega_{k,t-1}$ is establishment k's conventional ZIP3 market share.[37] To simplify exposition, we refer to the customer-level share $\omega_{i,j(k),t-1}$ in these specifications as customer effective shares, since firm-level effective shares are constructed by aggregating these customer shares across the firm's customers. All specifications include market-by-year fixed effects, defined as category-by-ZIP3-by-year fixed effects. These fixed effects absorb differences in average retention across categories, local markets, and years, so the coefficients are identified from variation across customer–firm relationships within the same local market and year. We report within-R^2 relative to these fixed effects.

We use the binary retention indicator as our outcome because it directly measures whether a customer relationship continues.[38] We standardize both $\log \omega_{i,j(k),t-1}$ and $\log \Omega_{k,t-1}$ within category so that coefficients β_1 and β_2 are similarly scaled and correspond to the effect of a one-standard-deviation increase in the respective variable.[39]

This specification distinguishes economically meaningful persistence in customer-firm relationships from two less interesting possibilities. First, if customer effective shares mostly reflect random variation in realized purchases, they should have little predictive power for future behavior. Second, even if customers shopped completely at random, retention might be higher at establishments with larger conventional market shares simply because customers are more likely to encounter and purchase from them again, not because those establishments are important to the specific customers they serve. The key question is therefore whether customer effective shares predict retention after controlling for the establishment's conventional market share. If they do, this indicates that customer effective shares contain information about persistent relationships beyond establishment scale.

Table 1 reports the results. Column (1) relates retention to effective shares alone: customers with higher effective shares in one year are substantially more likely to continue shopping at

[37] Our primary measures of effective shares are defined at the firm level and not at the establishment level. A customer who shops at several establishments of the same firm therefore enters the sample once per establishment, with the same effective share but establishment-specific conventional market shares. Online transactions are excluded before constructing these measures, so online-only firms are not in this analysis.

[38] A natural alternative is to use $\omega_{i,j(k),t}$ itself as the outcome. However, this outcome is harder to interpret since a customer's relative spending shares must sum to one across firms, and so a change in any customer-firm share pair mechanically requires offsetting changes in other customer-firm shares. Nevertheless, Appendix Table A8 shows that results are similar if we instead study this outcome, or if we measure retention over longer horizons.

[39] The log specification provides a parsimonious way to compare differential effects of the highly skewed share variables, and is consistent with the approximately log-linear relationship we showed in Figure 6.

Table 1: Effective Shares and Customer Retention

	(1)	(2)	(3)	(4)	(5)
$\log \omega_{i,j(k),-1}$	0.037***		0.034***	0.039***	0.038***
	(0.001)		(0.001)	(0.001)	(0.002)
$\log \Omega_{k,-1}$		0.039***	0.032***		
		(0.002)	(0.002)		
Market FE	X	X	X	X	X
Establishment-by-Year FE				X	X
Drop in $\log \Omega$					X
Observations	19,005,540	19,005,540	19,005,540	19,004,762	2,468,369
R^2	0.117	0.114	0.118	0.170	0.174
Within R^2	0.006	0.003	0.007	0.006	0.007

Notes: This table reports regressions of retention, an indicator for shopping at the establishment in year t, on a customer's log effective share at the firm ($\log \omega_{i,j(k),-1}$) and the establishment's log conventional ZIP3 market share ($\log \Omega_{k,-1}$), both measured in year $t-1$ and standardized within category. The sample is customer–establishment pairs in which the customer shopped at the establishment in $t-1$ and was active in the category in t. The analysis sample focuses on a 1% random sample of approximately 22,000 establishments for computational efficiency. Market FE are category-by-ZIP3-by-year fixed effects. Column (4) adds establishment-by-year fixed effects and Column (5) additionally restricts the sample to establishments whose decline in log market share between $t-1$ and t exceeds the median decline among establishments with falling market share. These establishment-by-year FE in Columns (4) and (5) absorb $\log \Omega_{j,-1}$. Within-R^2 is relative to all included fixed effects. Observations are weighted equally, and standard errors are clustered at the category-by-ZIP3-by-year level.

the establishment in the following year. A one-standard-deviation increase in a customer's log effective share is associated with a 3.7 percentage point higher probability of retention, relative to a sample mean retention rate of 36 percent. Column (2) shows that conventional market shares are also positively related to retention, consistent with customers being more likely to return to establishments with a larger market presence. However, Column (3) includes both variables and shows that conventional market shares add little once customer effective shares are included. The coefficient on effective shares is essentially unchanged, and adding market shares raises the within-R^2 only from 0.006 to 0.007, while adding effective shares to market shares raises it from 0.003 to 0.007. Column (4) includes establishment-by-year fixed effects, identifying the relationship by comparing customers who shopped at the same establishment in the same year. This specification asks whether, among customers of the same establishment, those who devote a larger share of their spending there are more likely to keep shopping there in the future. This is indeed the case.

Table 1 weights all customer-establishment observations equally, so it describes retention for the average customer-establishment relationship. Appendix Table A6 weights observations by establishment sales, which describes retention for the average dollar of spending. The relationship between effective shares and retention is even stronger in that case: the coefficient on effective shares is 0.083 (versus 0.037 in the equal-weighted case), and the within-R^2 is 0.017 (versus 0.006). Appendix Table A7 shows that similar patterns also emerge when we aggregate to the establishment (firm-by-ZIP3) or firm level instead of estimating specifications for customer.[40] In those specifications, establishments and firms with higher effective shares retain a larger share of their prior-year sales from customers who continue shopping there, and effective shares outperform conventional market shares in predicting this outcome whenever they are constructed from individual customers' spending.[41]

One concern is that retention may partly reflect passive inertia in stable shopping environments rather than meaningful differences in attachment. For example, customers may continue shopping at a firm because there is currently little reason to change, even if they would switch in response to a small change in prices, quality, or available alternatives. Column (5) of Ta-

[40] Appendix Table A7 also compares effective shares to two related establishment-level statistics– sales per customer and the number of customers. Sales per customer positively predicts retention, while the number of customers has little predictive power once sales per customer is included, echoing the finding in Afrouzi, Drenik and Kim (2025) that markups are more strongly related to sales per customer than to customer counts. However, effective shares remain the stronger predictor: specifications using effective shares have higher explanatory power, and sales per customer no longer predicts retention once effective shares are included. This pattern suggests that sales per customer is informative because it proxies for customer depth, which effective shares measure more directly.

[41] The exception is the synthetic-customer specification (Appendix Table A7, column 5), where pooling households removes customer-level heterogeneity and makes effective shares mechanically closer to market shares. Even in this specification, effective shares remain significantly related to retention.

ble 1 therefore restricts attention to establishments experiencing large declines in market share, where many customer relationships are being disrupted. This specification continues to include establishment-by-year fixed effects, so the coefficient is identified by comparing customers of the same declining establishment in the same year. Effective shares remain strongly predictive of retention in this subsample. This suggests that effective shares capture customer relationships that matter even when establishments are experiencing disruption or declining relative appeal.

Taken together, these results show that effective shares are not merely noisy spending realizations or proxies for overall firm scale. They predict which relationships persist, both across establishments and among customers of the same establishment, indicating that effective shares are a meaningful measure of customer attachment. The next subsection asks whether this attachment is also informative about substitution behavior when customers face a new alternative.

5.2 Responses to Establishment Entry

The evidence above shows that effective shares predict persistent customer–firm relationships beyond what is captured by conventional market shares. We next show that these same relationships predict substitution when customers face a new alternative. The theoretical framework in Section 2 provides one interpretation of this result: demand from a customer is less elastic when a firm accounts for a larger share of that customer's category spending. The framework therefore predicts that customers with higher effective shares at an incumbent should be less likely to reallocate spending to a new entrant.

We examine this prediction using local entry events, defined as the first appearance of a firm in a 5-digit ZIP code. We discuss this sample as well as various additional details in Appendix A4. Local entry is endogenous, so all specifications include entry-event fixed effects and thus compare customers exposed to the same entry event, rather than comparing across markets with different entry decisions. For each entry event, we define the exposed customers as those living in the 5-digit ZIP code that accounts for the largest number of customers visiting the entrant in the entry year.[42]

Let e index an entry event, defined by an entrant and entry year, and index time relative to the entry event, with 0 denoting the entry year and -1 denoting the year before entry. For each exposed customer i who shopped at the incumbent establishment k in year -1 we estimate

[42]The ZIP code with the largest number of customers visiting the entrant need not be the ZIP code in which the entrant is located.

$$Switch_{i,k,e,0} = \beta_1 \log \omega_{i,j(k),e,-1} + \beta_2 \log \Omega_{k,e,-1} + \gamma_e + \epsilon_{i,k,e,0}.$$

Here $Switch_{i,k,e,0}$ is an indicator for whether customer i both stops purchasing at k and purchases from the entrant in the entry year.[43] The variable $\omega_{i,j(k),e,-1}$ is customer i's effective share at incumbent firm $j(k)$ in the year before entry, measured using the customer's firm-wide spending. $\Omega_{k,e,-1}$ is incumbent k's conventional market ZIP3 market share in that same year, and γ_e is an entry-event fixed effect. We again standardize regressors so that β_1 and β_2 correspond to one standard-deviation effects. Together they show whether switching to a new entrant is more closely related to the customer's effective share at the incumbent or to the incumbent's overall share of local sales.

Table 2: Effective Shares Predict Less Switching to Entering Firms

	(1)	(2)	(3)	(4)	(5)
$\log \omega_{i,j(k),-1}$	-0.028***		-0.027***	-0.029***	-0.031***
	(0.001)		(0.001)	(0.001)	(0.001)
$\log \Omega_{k,-1}$		-0.010***	-0.003***		
		(0.001)	(0.000)		
Entry Event FE	X	X	X	X	X
Incumbent*Entry Event FE				X	X
Drop in $\log \Omega$					X
Observations	35,164,534	35,164,534	35,164,534	34,495,972	6,347,760
R^2	0.062	0.052	0.062	0.124	0.138
Within R^2	0.012	0.001	0.012	0.011	0.011

Notes: This table reports regressions of switching on the customer's log effective share at the incumbent's firm ($\log \omega_{i,j(k),-1}$) and the incumbent's log conventional ZIP3 market share ($\log \Omega_{k,-1}$), both measured in $t-1$ and standardized within category. Switching is an indicator for whether a customer stops shopping at the incumbent establishment in the entry year t and shops at the entrant. The sample is customer-incumbent-entry-event observations in which the customer shopped at the incumbent establishment in $t-1$ and remains active in the category in t. The analysis focuses on a 10 percent sample of identified entry events (about 1,000 events) for computational efficiency. All columns include entry-event fixed effects. Column (4) adds incumbent-by-entry-event fixed effects. Column (5) additionally restricts the sample to incumbent establishments whose decline in log market share from $t-1$ to t exceeds the median decline among establishments with falling market share. Within-R^2 is relative to all included fixed effects. Standard errors are clustered by entry event.

Table 2 shows that customers are less likely to switch to a newly entering firm if the incumbent previously accounted for a larger share of their category spending. A one standard deviation in a customer's log effective share is associated with a 2.8 percentage point decline in

[43] To reduce concerns about customer attrition, we restrict the sample to customers who make at least one purchase at any firm in the incumbent's category during the entry year.

their switching probability, which is about 40 percent of the sample mean switching probability of 7 percent. This relationship is accounted for by customer effective shares rather than by the incumbent's conventional market share: an increase in $\omega_{i,j(k),e,-1}$ strongly predicts a decline in switching, while $\Omega_{j,e,-1}$ has little relationship once effective shares are included. Column (4) adds incumbent-by-entry event fixed effects, which absorb the incumbent's market share and any incumbent-level exposure to the entrant, so identification comes from comparing customers of the same incumbent who devote different shares of their spending to it. This also addresses the concern that some incumbents are closer substitutes for the entrant than others. Effective shares continue to predict a decline in switching in this specification. Column (5) shows that the same pattern also holds when we further restrict this specification to only include incumbents that experience large declines in log market share in the entry year, a proxy for incumbents who are particularly exposed to the entrant's arrival.

Appendix A4 reports several robustness exercises: replacing the incumbent's market share with a customer-specific analog (the spending-weighted market share of the establishments where the customer shops), restricting to incumbents located in the entrant's 3-digit ZIP code, excluding restaurants (which dominate the baseline sample), and estimating separate relationships for 200 separate entry events rather than imposing a common relationship across all entry events. In all cases, the patterns are consistent with the interpretation that effective shares capture meaningful variation in customer attachment and substitution behavior.

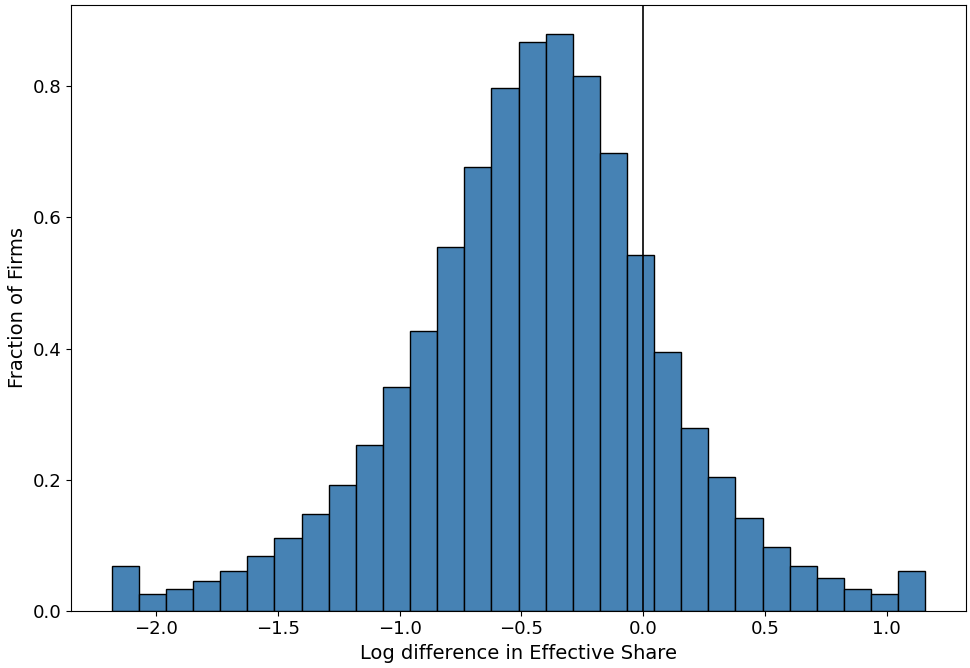
6 Effects of Customer Acquisition on Firm Growth and Effective Shares

In this section, we turn to understanding the muted relationship between effective and conventional market shares. We show that growing firms expand their reach by acquiring new customers, but these customers are initially less attached and therefore attenuate effective-share growth. At the same time, growing firms also deepen relationships with existing customers, which raises effective shares directly. Together, these two opposing forces help explain why effective shares rise only weakly with conventional market shares as firms grow.

Prior work emphasizes the role of customer acquisition for firm growth (Klenow et al., 2022; Argente et al., 2025; Afrouzi, Drenik and Kim, 2025), and we find the same pattern in our data: changes in the number of customers account for roughly 80 percent of the cross-firm variation in sales growth between 2021 and 2025.[44] The attachment of newly acquired customers is therefore

[44]Specifically, the coefficient from a regression of $\Delta \log(\text{sales})$ on $\Delta \log(\text{no. customers})$ with category fixed effects is 0.8.

Figure 7: Effective Shares for New versus Continuing Customers



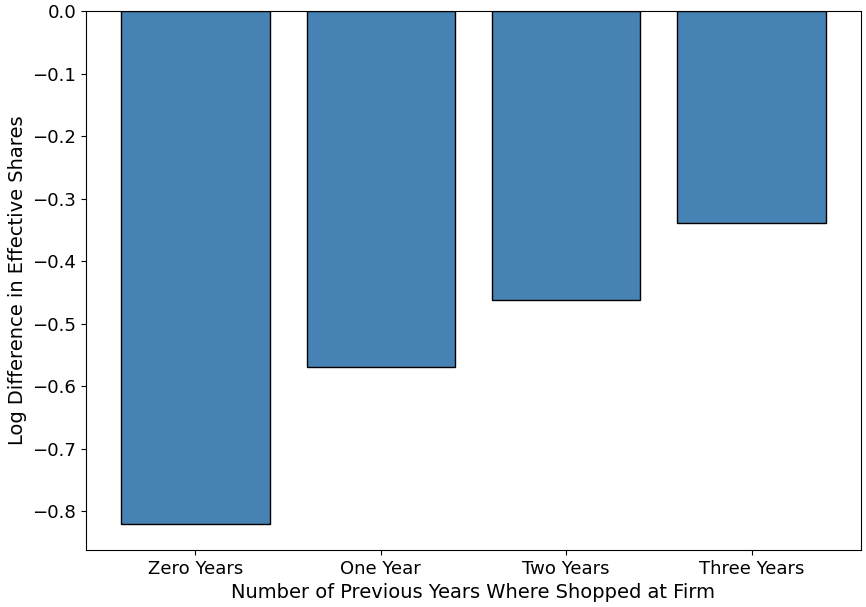
Notes: This figure shows distribution across firms (each firm weighted equally) of $d_j \equiv \log \omega^{\text{eff}}_{j,\text{new}} - \log \omega^{\text{eff}}_{j,\text{cont}}$ in 2025, where $\omega^{\text{eff}}_{j,\text{new}}$ is the effective share of customers who purchase from firm j in 2025 but not in 2021-2024, and $\omega^{\text{eff}}_{j,\text{cont}}$ is the effective share of customers who purchase from j in 2025 and in at least one year from 2021 to 2024. Negative values indicate that newly acquired customers have lower effective shares than continuing customers. Only firms with at least 20 customers in both groups are included, and the distribution is winsorized at the top and bottom 0.5 percent. The vertical line indicates $d_j = 0$, where the two group-specific effective shares are equal.

central to understanding how firm growth translates into effective-share growth.

We begin by comparing effective shares of continuing customer to newly acquired customers. For each firm, we partition customers who we observe spending at the firm in 2025 into two groups: new customers, who purchase from the firm in 2025 but not in any earlier sample year, and continuing customers, who purchased from the firm in 2025 and at least one earlier sample year. We then compute the firm's effective share separately within each group, weighting each customer by its spending at the firm relative to the group's total spending at the firm so that weights sum to one within each group. Figure 7 plots the distribution across firms of the log difference between the effective share of new customers and the effective share of continuing customers.

The vast majority of firms have negative gaps: newly acquired customers devote a smaller share of category spending to the firm than continuing customers do. At the median firm, the

Figure 8: Effective Shares Rise with Prior Years at the Firm



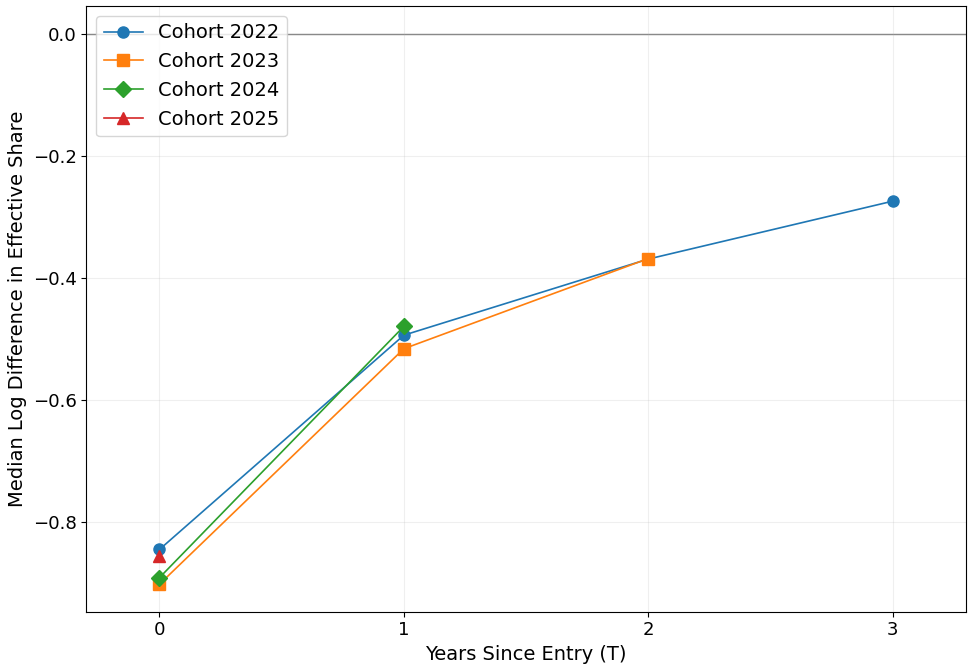
Notes: This figure groups 2025 customers by the number of prior sample years (2021-2024) in which they shopped at the firm. For each firm and each $k \in \{0, 1, 2, 3\}$, we compute the log difference between the effective share of customers with k prior years at the firm and the effective share of customers with four prior years at the firm. Only firms with at least 20 customers in each of the five groups are included, and effective shares are computed within each group using customer sales weights that sum to one within the group. The figure plots the median of the log difference across firms.

effective share of new customers is 35 percent lower than that of continuing customers.[45] The median gap is negative in every category, ranging from 8 to 51 percent (Appendix Table A12).

The new-versus-continuing comparison summarizes prior exposure with a single indicator: whether a 2025 customer purchased from the firm in any earlier sample year. We next explore whether effective shares vary systematically with relationship tenure measured over the 2021 to 2025 sample. For each firm, we group 2025 customers by the number of prior years they shopped at the firm and compare each group's effective share to that of customers who shopped at the firm in all four prior years. Figure 8 plots the median of these firm-level relative differences for each group. Attachment rises steadily with relationship length: customers new in 2025 have effective shares about 56 percent below the most-tenured group, and this deficit shrinks to roughly 43, 37, and 29 percent for customers with one, two, and three prior years at

[45] $0.35 = (1 - e^{-.426})$. If we require that continuing customers must purchase in every year from 2021 through 2025, rather than in at least one year, then the median gap rises to 49 percent.

Figure 9: Effective Shares of Entrant Customers by Years Since Customer Acquisition



Notes: This figure follows cohorts of customers whose first observed purchase from a firm occurs in year c (cohorts 2022 to 2025; 2021 is excluded because first purchases in that year cannot be distinguished from earlier ones). For each cohort c and year $t \geq c$, we keep customers who purchase from the firm in every year from c through t and compute the log difference between this cohort's effective share in year t and the year-t effective share of the firm's established customers, defined as customers who shop at the firm in all five sample years. The figure plots the median of this firm-level log difference across firms, by cohort and years since entry. Only firms with at least twenty customers in each relevant cohort and comparison group are included. Unlike Figure A5, this figure conditions only on survival through each horizon t.

the firm.[46] Customers with a longer history at the firm thus tend to direct more of their category spending there.

The tenure gradient reflects both deepening within customer-firm relationships as they age and the survival of more attached relationships, consistent with the retention results in Section 5. To examine these forces, Figure 9 follows cohorts of newly acquired customers by years since their first observed purchase at the firm, conditioning at each horizon on continued purchasing through that horizon. Customers who continue purchasing through longer horizons direct more of their category spending to that firm, indicating that persistence on the extensive margin is closely linked to depth on the intensive margin. This is not an accounting identity, since a customer could purchase from a firm every year while still doing little of their category spending there. Notably, the gap does not close within our sample window: three years after first

[46] The corresponding median log differences are -0.82, -0.57, -0.46, and -0.34.

purchase, the 2022 cohort's effective share remains around 24 percent below that of customers who purchase from the firm in every sample year.[47]

Taken together, these customer-level patterns help explain why customer acquisition weakens the link between conventional market-share growth and effective-share growth. Acquiring new customers expands reach, but newly acquired customers have lower effective shares than continuing customers. However, this does not imply that growing firms have declining effective shares. The same forces that attract new customers may also make the firm more attractive to its incumbent customers and raise those customers' spending shares. Figure 10 evaluates the balance of these forces by grouping firms into terciles based on the fraction of their 2021 to 2025 sales growth accounted for by net customer acquisition and plotting the relationship between effective-share growth and market-share growth separately for each group.[48]

Consistent with the customer-level evidence, the relationship between effective-share growth and market-share growth is much weaker for firms whose growth is more customer-acquisition-driven. The slope is 0.43 in the lowest customer-acquisition tercile, 0.23 in the middle tercile, and 0.06 in the highest. At the same time, the slope remains positive even among firms whose growth is most concentrated in customer acquisition. This suggests that firms gaining customers are not merely adding shallow relationships but are, at the same time, deepening relationships with their incumbent customers.[49]

7 Growth Channels and Concentration Trends

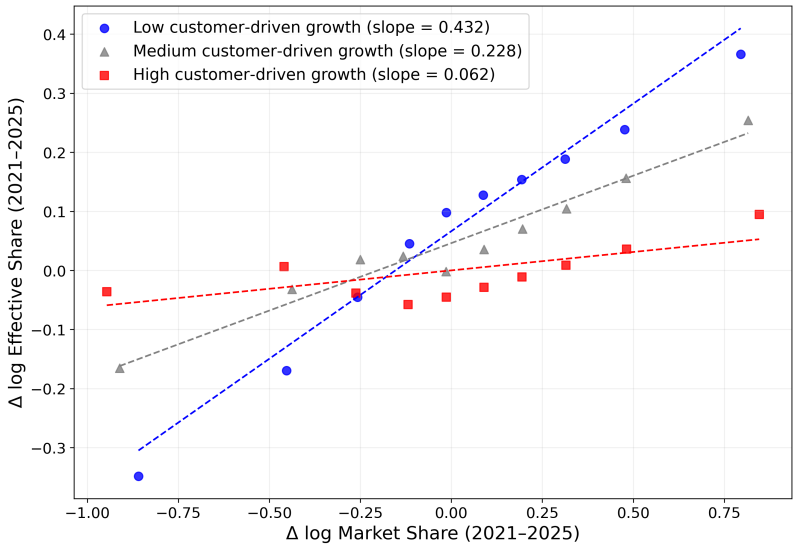
The previous section showed that when firms grow by reaching new customers, conventional market shares rise more than effective shares. We now show that two prominent sources of recent firm growth, geographic expansion and online sales, operate importantly through this reach margin. We then revisit concentration trends and show that the rise of superstar firms in aggregate sales has not been accompanied by a comparable rise in effective shares.

[47] Appendix Figure A5 instead conditions on survival through 2025, holding the set of customers fixed within each first-purchase cohort as years since first purchase increase. There the gap narrows sharply after the first-purchase year and changes little afterward, suggesting that within-relationship changes are concentrated in the first year, while the continued rise in later years in Figure 9 reflects the survival of more attached relationships.

[48] Specifically, letting $g_{N,j} = \Delta \log(\text{No. Customers}_j)$ and $g_{S,j} = \Delta \log(\text{Sales per Customer}_j)$, we group firms into terciles of $|g_{N,j}|/(|g_{N,j}| + |g_{S,j}|)$. This measure is bounded between zero and one and is well defined even when a firm has little overall growth or the two terms offset each other.

[49] A simple example illustrates how customer acquisition can dominate sales growth while incumbent deepening still raises effective shares. Suppose there are 200 customers, each spending \$100 in the category. A firm initially sells \$20 to each of 100 customers. It then adds 100 customers who each spend \$10 at the firm, while incumbent customers increase spending at the firm from \$20 to \$25. Sales rise from \$2000 to \$3500, with \$1000 of the \$1500 increase coming from new customers. Customer acquisition accounts for most sales growth, and the effective share rises, from 0.2 to $.25(2500/3500) + .1(1000/3500) \approx .207$.

Figure 10: Effective-Share Growth and the Customer-Acquisition Margin



Notes: This figure plots the relationship between firms' change in log effective share and change in log market share between 2021 and 2025. The sample is constructed from the national panel and restricted to firms observed in both years. The figure trims the bottom and top 2.5 percent of the distribution of the change in log market share. Firms are divided into terciles based on the share of overall growth attributable to customer growth, defined as $\frac{|g_N|}{|g_N|+|g_S|}$, where g_N is log customer growth and g_S is growth in sales per customer. Within each tercile, points denote bin means from a binscatter of $\Delta \log$ effective share against $\Delta \log$ market share, using common bins constructed from the full sample. Dashed lines show tercile-specific weighted linear fits.

Geographic Expansion. One important channel for recent firm growth is geographic expansion. Prior work argues that rising national concentration partly reflects firms expanding into more local markets rather than becoming larger within each market (Hsieh and Rossi-Hansberg, 2023). Related evidence shows that entrants build market share by entering more geographic markets and opening more stores within those markets (Argente et al., 2025).

Geographic expansion tends to raise conventional market shares more than effective shares because it expands customer reach rather than deepening existing customer relationships. If customers in new markets looked identical to the firm's existing customers, geographic expansion would increase the firm's aggregate sales and conventional market share without changing its effective share. If these new customers are more marginal, geographic expansion instead pushes effective shares down. Consistent with this mechanism, Appendix Figure A6 shows that when a firm enters a new ZIP3 market, the effective share of its new establishment is substantially below that of its existing establishments.

Panel (a) of Figure 12 shows the implications of changing geographic footprints for the relationship between effective-share growth and market-share growth. The purple squares show

firms that changed the number of 3-digit ZIP codes in which they operated between 2021 and 2025, while the green circles show firms whose geographic footprint remained fixed. For the same change in market share, effective shares move substantially more for firms with a fixed geographic footprint than for firms whose footprint changes. Changing the footprint adds or removes customers in other locations, which changes reach without a corresponding change in the depth of customer relationships.

Online Sales. A second prominent source of recent firm growth is the expansion of online sales. Appendix Figure A7 shows that within our categories, online sales rose from around 20 percent of sales in 2014 to over 37 percent in 2025.[50] Online sales jumped in 2020 during the pandemic, but the online share remained substantially above its pre-pandemic level and returned to a positive trend thereafter. Like geographic expansion, online growth can weaken the relationship between conventional market shares and effective shares by expanding reach to less attached customers with access to a broader set of firms.

We begin by showing that customers who buy from a firm online allocate a smaller share of their category spending to it than customers who buy from the same firm in person. Figure 11 reports estimates of β from regressions of the form

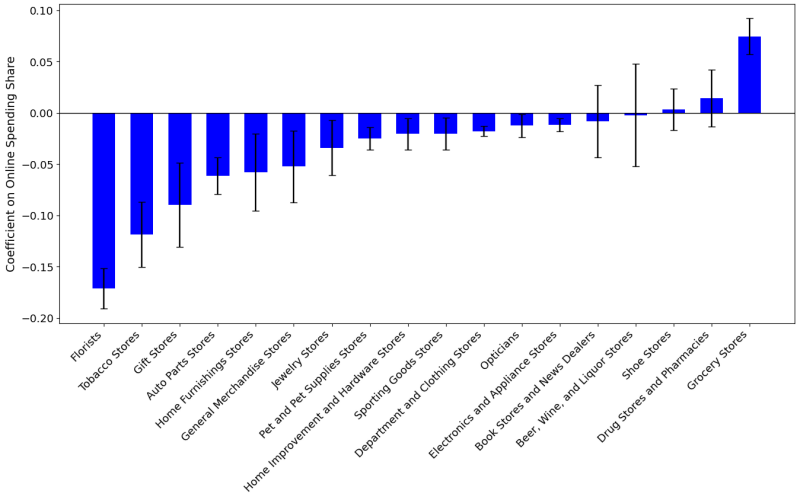
$$\omega_{i,j,t} = \beta \times \text{online share}_{i,j,t} + \alpha \log(x_{i,j,t}) + \gamma_j + \varepsilon_{i,j,t}, \tag{14}$$

where $\omega_{i,j,t}$ is the share of customer i's category spending that goes to firm j, $x_{i,j,t}$ is the customer's total spending at firm j in year t, γ_j are firm fixed effects, and $\text{online share}_{i,j,t}$ is the fraction of customer i's spending at firm j that occurs online. For computational reasons, we restrict the sample to 2025. The analysis focuses on retail categories since services require in-person consumption. We control for total spending at the firm because online customers tend to have higher incomes (Dolfen et al., 2023) and thus spend more at any given firm. With firm fixed effects the coefficient β measures whether customers who spend similar amounts at the same firm allocate a smaller share of their category spending to that firm when more of their spending there occurs online.

The estimated coefficient is negative in most categories. Conditional on total spending at a firm, customers who do more of their spending there online allocate a smaller share of their category spending to it. In some categories, such as florists, the difference is large. In others, such as department and clothing stores, it is smaller but still negative. Groceries are a notable exception. Unlike many other forms of online retail, online grocery purchases often involve buying

[50] We use the more aggregated historical panel available back to 2014, since this calculation does not require disaggregated firm data.

Figure 11: The Relative Effective Share of Online Customers



Notes: This figure reports category-specific estimates of β from Equation 14, estimated separately for each category using 2025 data. For each category, the regression specification includes firm fixed effects and controls for log of the customer's total spending at that firm. Vertical lines indicate 95 percent confidence intervals constructed using standard errors clustered at the firm level.

from the same local stores through a different transaction mode, so online grocery shopping may save time without substantially expanding the customer's relevant choice set.

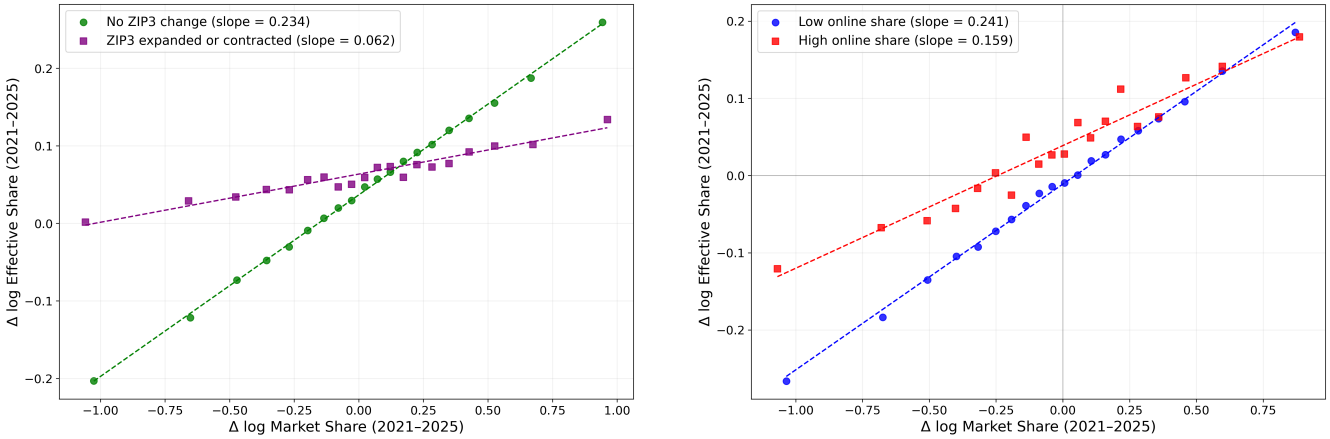
If online access allows firms to expand by reaching less attached customers, then growth among predominantly online firms should translate less strongly into customer depth than growth among predominantly in-person firms. To test this prediction, we split firms by their initial online spending share and compare the relationship between effective-share growth and market-share growth across the two groups.[51] Panel (b) of Figure 12 shows that firms with greater online sales exhibit a weaker relationship between market-share growth and effective-share growth.[52]

Superstar Firm Growth. Since superstar firms have grown large in part by expanding customer reach through channels such as online sales and geographic expansion, we next ask

[51]This specification asks whether the relationship differs across firms with different initial modes of selling. A complementary exercise classifies firms by the channels through which their sales change between 2021 and 2025. Specifically, we divide the absolute change in online sales by the sum of the absolute changes in online and in-person sales, which captures the relative importance of online sales in gross sales changes. Splitting firms by this measure also yields a weaker relationship for the high online-contribution group (.19 vs .23).

[52]Online spending shares differ both within and across categories, but the difference in slopes persists when we remove cross-category variation: allowing for category-specific means and slopes to absorb industry heterogeneity yields an average difference in slopes of .058 compared to a difference of 0.082 when including cross-industry effects.

Figure 12: The Nature of Firm Growth and Effective Shares



(a) Geographic Expansion

(b) Online Spending

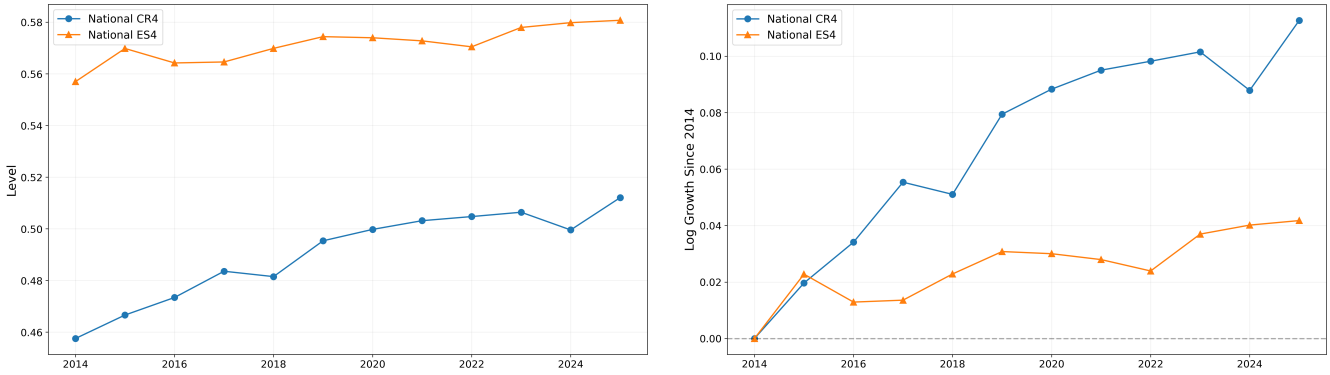
Notes: The figure plots binscatters of $\Delta \log$ effective share against $\Delta \log$ market share between 2021 and 2025 for firms observed in both years. Panel (a) splits firms by whether the number of ZIP3 markets in which they operate changes between 2021 and 2025. Panel (b) restricts the sample to retail categories and compares firms whose online share of sales in 2021 was below 20 percent with firms whose online share was above 80 percent. In both panels, the bottom and top 2.5 percent of the distribution of $\Delta \log$ market share are trimmed. Points denote sales weighted bin means and the same bins are used for both groups within each panel. Dashed lines show group-specific linear fits.

whether their rising aggregate sales shares have been accompanied by a comparable rise in effective shares.

For this analysis we turn to the historical panel data. As discussed in Section 3, this covers a sample from 2014–2025. and measures total spending by individual cards in each ZIP code, but only further disaggregates this spending for a subset of relatively large "named firms." Because the panel records each card's total spending at unnamed firms, we can compute effective shares for named firms, but not for all firms, so we implement a top-firm analog of effective concentration. We compute the average effective share for customers of the top four firms in each category, which we term ES4, and compare it to the conventional sales share of those same four firms, CR4.

Appendix Figure A8 supports this focus on top firms. In overlapping years, CR4 and ES4 line up closely between the historical and primary panels. In the primary panel, ES4 is also strongly related to $\overline{\omega}$, the dollar-weighted average effective share of all firms, suggesting that ES4 is a reasonable proxy for effective shares more broadly. We focus on national trends because large national firms are unlikely to be omitted from the named-firm panel.

Figure 13: Concentration Levels and Growth Trends in the Balanced Sample



(a) Concentration levels

(b) Concentration growth

Notes: The figure reports concentration measures from 2014 to 2025. CR4 denotes the four-firm concentration ratio, and ES4 denotes the average effective share of the top four firms. Effective shares are constructed from the historical panel, with customers defined at the card level. To limit composition effects, we construct the ES4 for each category using a chained index. After constructing category-level concentration measures, we average across categories using fixed 2025 category total sales weights. The sample is restricted to a balanced panel of categories observed in every year from 2014 through 2025 with at least 10 named firms in each year and where named firms make up at least 25% of category sales. Panel (a) plots levels. Panel (b) plots log growth relative to 2014.

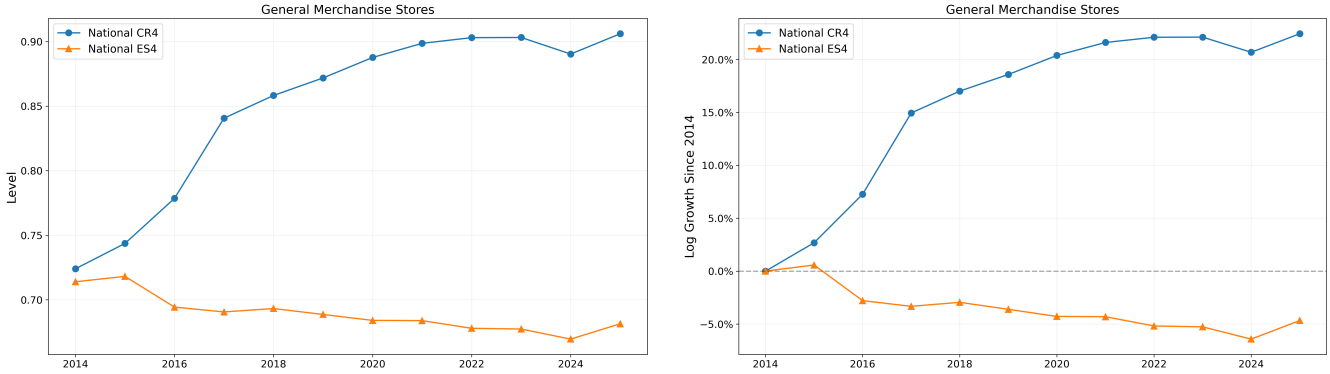
Figure 13 shows concentration trends in CR4 and ES4 in the historical panel.[53] Panel (a) shows levels, and panel (b) normalizes each series to its 2014 value to highlight differences in growth.[54] National CR4 rises by around 12 percent, from 0.458 to 0.512, while ES4 rises by only around 4 percent, from 0.557 to 0.581, roughly a third as much. The rise in national sales concentration therefore substantially overstates the rise in the largest firms' effective shares.

While we restrict our analysis to categories in which named firms account for at least 25 percent of spending, it remains possible that some top firms are included in the unnamed residual category. To mitigate this concern, Figure 14 repeats the analysis for General Merchandise Stores, where approximately 95 percent of sales occur at named firms and the top four firms are therefore measured especially accurately. The divergence is even larger in this category. The conventional sales share of the top firms rises sharply, while their average effective instead *declines* mildly .

[53]To limit composition effects from firms with different levels of effective shares entering and leaving the top four, we measure ES4 as a chained index. The index starts from the sales-weighted average effective share of the top four sales firms in 2014 and then each year adds the average change in effective share from the previous year of that year's top four firms, with sales weight averaged over the two years for each firm. Results are similar if we instead compute the average ES4 directly each year in the raw data.

[54]All of these series average across categories using fixed 2025 category sales weights, since we are interested in within-category changes rather than changes in category composition. Figure A9 shows similar patterns when

Figure 14: Concentration Levels and Growth Trends for General Merchandise Stores



(a) Concentration levels

(b) Concentration growth

Notes: This figure repeats the analysis from Figure 13 but restricts attention to General Merchandise Stores, where coverage of named firms is highest.

More generally, the average trends in Figure 13 mask substantial heterogeneity across categories. Appendix Table A11 reports national CR4 and ES4 in 2014 and 2025 for each of the 17 retail categories in the historical panel. In 11 of the 17 categories, ES4 changes by less than CR4, so switching from CR4 to ES4 moves the measured trend closer to zero, but in some categories where CR4 falls, such as drug stores and office supplies stores, ES4 rises. These patterns reinforce the conclusion that changes in sales concentration do not translate into proportionate changes in effective shares.

These trends also bear on the debate over national versus local concentration trends discussed in the introduction. Measured concentration trends hinge on a choice of geography, which effective shares sidestep by taking a customer-focused perspective. The muted rise in effective shares that we document suggests that the rise of "superstar" firms has not been accompanied by a comparable increase in their importance to customers.

8 Conclusion

Conventional market shares measure where dollars are spent, but not who spends them or how important a firm is to its customers. In this paper, we show that this distinction matters by constructing effective shares, which measure a firm's share of spending within its own customer base. Effective shares predict customer behavior over time, and their evolution depends critically on how firms grow. Measuring effective shares also changes the interpretation of traditional measures of firm size and dominance. Firms that are extremely large by conventional

categories are averaged using time-varying sales shares.

market shares look much less exceptional using effective shares. Similarly, effective shares have risen much less than aggregate concentration over the last decade. Interpreted through the lens of the variable-markup framework that maps market power to firm "size", these patterns suggest that rising concentration has not led to a comparable rise in size-based product-market power.

The broader lesson is that aggregate concentration statistics can be misleading when economic power is relationship-specific. In product markets, the relevant question is not only how much a firm sells, but how central it is within the spending baskets of the customers it sells to. Similar measurement issues arise in other settings where agents have heterogeneous relationships, such as credit markets or labor markets. Measures of within-relationship dominance may therefore provide a useful complement to aggregate concentration statistics in studying economic power more broadly.

References

Afrouzi, Hassan, Andres Drenik, and Ryan Kim. 2025. "Concentration, Market Power, and Misallocation: The Role of Endogenous Customer Acquisition." *NBER Working Paper 31415*.

Argente, David, Doireann Fitzgerald, Sara Moreira, and Anthony Priolo. 2025. "How Do Entrants Build Market Share? The Role of Demand Frictions." *American Economic Review: Insights*, 7(2): 160–176.

Argente, David, Esteban M'endez, and Diana Van Patten. 2026. "De Minimis as a Constraint on Retail Market Power: Local Catchment Areas and Markup Elasticities." Working paper.

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

Autor, David, Christina Patterson, and John Van Reenen. 2023. "Local and national concentration trends in jobs and sales: The role of structural transformation." National Bureau of Economic Research.

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

Baqaee, David Rezza, and Emmanuel Farhi. 2020. "Productivity and Misallocation in General Equilibrium." *The Quarterly Journal of Economics*, 135(1): 105–163.

Batch, Andrea, Benjamin Bridgman, Abe Dunn, and Mahsa Gholizadeh. 2024. "Consumption Zones." *Journal of Economic Geography*, 24(2).

Becker, Jonathan, Chris Edmond, Virgiliu Midrigan, and Daniel Yi Xu. 2026. "The Spatial Distribution of Markups." Working paper.

Benkard, C. Lanier, Nathan H. Miller, and Ali Yurukoglu. 2025. "The Rise of Market Power and the Macroeconomic Implications: Comment." *NBER Working Paper 3145*.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841–890.

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

Dolfen, Paul, Liran Einav, Peter J. Klenow, Benjamin Klopack, Jonathan D. Levin, Larry Levin, and Wayne Best. 2023. "Assessing the Gains from E-Commerce." *American Economic Journal: Macroeconomics*, 15(1): 342–370.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2023. "How Costly Are Markups?" *Journal of Political Economy*, 131(7): 1619–1675.

Einav, Liran, Mariana Guido, and Pete Klenow. 2026. "Customer Overlap and Diversion Ratios." *NBER Working Paper 34693*.

Foster, Lucia, John Haltiwanger, and Matthew Tuttle. 2022. "Rising Markups or Changing Technology?" U.S. Census Bureau, Center for Economic Studies CES Working Paper CES-WP-22-38.

Hsieh, Chang-Tai, and Esteban Rossi-Hansberg. 2023. "The Industrial Revolution in Services." *Journal of Political Economy*.

Hubmer, Joachim, and Jakob Nord. 2026. "Demand Investment, Market Power, and the Rise in Concentration." Working Paper. Working paper.

Klenow, Peter J, Jonathan D Levin, Raviv Murciano-Goroff, and Liran Einav. 2022. "Customers and retail growth." Working Paper.

Mongey, Simon, and Michael E. Waugh. 2025. "Pricing Inequality." National Bureau of Economic Research NBER Working Paper 33399.

Neiman, Brent, and Joseph Vavra. 2023. "The rise of niche consumption." *American Economic Journal: Macroeconomics*, 15(3): 224–264.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter. 2021. "Diverging trends in national and local concentration." *NBER Macroeconomics Annual*, 35(1): 115–150.

Smith, Dominic A., and Sergio Ocampo. 2025. "The Evolution of US Retail Concentration." *American Economic Journal: Macroeconomics*, 17(1): 71–101.

Steve Bond, Arshia Hashemi, Greg Kaplan Piotr Zoch. 2021. "Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data." *Journal of Monetary Economics*.

Syverson, Chad. 2019. "Macroeconomics and Market Power: Context, Implications, and Open Questions." *Journal of Economic Perspectives*, 33(3): 23–43.

Traina, James. 2018. "Is Aggregate Market Power Increasing? Production Trends Using Financial Statements." The University of Chicago Booth School of Business, George J. Stigler Center for the Study of the Economy and the State Working Paper 272.

A1 TransUnion Data Appendix

We first de-duplicate the TransUnion data so that each card is linked to only one person in each month. We follow the procedure developed by the data provider, which the provider had already applied to some years, and we apply it to the remaining years. TransUnion records the account type for each person-card pair, such as an individual account or an authorized-user account. When a card is linked to more than one person in a month, we keep the person whose account type ranks highest in the following order: individual account; account maker (the primary account holder on an account with a guarantor); joint contractual liability; authorized user, guarantor, or participant; undesignated account type; and terminated or deceased account.

We then merge the TransUnion data with the transaction data. Because TransUnion records are quarterly in 2021 and 2022 and monthly from 2023 onward, we merge at the card-quarter level for 2021 to 2022 and at the card-month level thereafter.

Because some cards are linked to different people over time, we next assign each card a single cardholder for the full sample. The cardholder is the person linked to the card for the longest period. Ties are broken first by the number of transactions and then by total spending on the card.

Finally, some cards appear in the transaction data in months in which they are not linked in TransUnion. For these cards, we assign the unmatched months to the card's cardholder and fill in demographic variables from the cardholder's nearest month in the TransUnion data.

A2 Appendix: Derivations for Effective Shares

A2.1 Nested-CES benchmark

This appendix derives the customer-level and firm-level demand elasticities used in Section 2, culminating in the result that a firm's demand elasticity depends on its effective share rather than its aggregate market share.

A2.1.1 Setup

Consider differentiated firms grouped into categories. We focus on one category and suppress its index. Customers are indexed by i, firms within the category by j, and the set of firms in the category by $\mathcal{F}$. Customer i has a CES consumption aggregator across firms in the category:

$$C_i = \left[\sum_{j \in \mathcal{F}} (a_{ij})^{1/\eta} \, q_{ij}^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}}, \qquad \eta > 1, \tag{15}$$

where q_{ij} denotes customer i's demand for firm j and $a_{ij} \geq 0$ is a persistent customer–firm-specific demand shifter. Heterogeneity in a_{ij} allows firms to differ persistently in their importance across customers because of geography, tastes, product availability, shopping technology, or other customer–firm-specific factors.

The corresponding category price index for customer i is

$$P_i = \left[\sum_{j \in \mathcal{F}} a_{ij} p_j^{1-\eta} \right]^{\frac{1}{1-\eta}}, \tag{16}$$

where firm j charges a common price p_j across customers. Conditional demand for firm j is

$$q_{ij} = a_{ij} \left(\frac{p_j}{P_i} \right)^{-\eta} C_i. \tag{17}$$

To capture substitution away from the category, assume that customer i's category demand has elasticity θ with respect to the category price index:

$$-\frac{\partial \log C_i}{\partial \log P_i} = \theta, \qquad \eta > \theta > 1. \tag{18}$$

The assumption $\eta > \theta$ means that customers substitute more readily across firms within a category than away from the category as a whole.

The rest of this section builds up to two main results: a closed-form expression for customer-level demand elasticities in terms of customer-level spending shares, and a closed-form expression for firm-level demand elasticities in terms of effective shares. We then specialize to the case without persistent customer heterogeneity and recover the standard Atkeson and Burstein (2008) formula.

A2.1.2 Customer-level spending shares

We begin by establishing two properties of customer-level spending shares that will be used in subsequent derivations: a closed-form expression in terms of demand shifters and prices, and the identity that links the spending share to the elasticity of the customer's price index with respect to firm prices.

Define firm j's share of customer i's spending in the category as

$$\omega_{ij} \equiv \frac{p_j q_{ij}}{\sum_{k \in \mathcal{F}} p_k q_{ik}}. \tag{19}$$

Substituting equation (17) into equation (19) gives

$$\omega_{ij} = \frac{a_{ij} p_j^{1-\eta}}{\sum_{k \in \mathcal{F}} a_{ik} p_k^{1-\eta}}. \tag{20}$$

This expresses customer-level spending shares directly in terms of the demand shifters a_{ij} and the price vector. We use this expression below to characterize the representative-household special case.

A standard property of CES demand is that this expenditure share equals the elasticity of customer i's category price index with respect to firm j's price. To see this, take logs of equation

(16) and differentiate with respect to $\log p_j$:

$$\frac{\partial \log P_i}{\partial \log p_j} = \frac{a_{ij} p_j^{1-\eta}}{\sum_{k \in \mathcal{F}} a_{ik} p_k^{1-\eta}} = \omega_{ij}. \tag{21}$$

This identity is the key input for the customer-level elasticity derivation that follows.

A2.1.3 Customer-level demand elasticities

We now derive the central customer-level result: customer i's demand elasticity for firm j is

$$\varepsilon_{ij} = \eta - (\eta - \theta)\omega_{ij}. \tag{22}$$

Start by taking logs of equation (17):

$$\log q_{ij} = \log a_{ij} - \eta \log p_j + \eta \log P_i + \log C_i. \tag{23}$$

Then differentiating with respect to $\log p_j$ yields

$$\begin{aligned} \varepsilon_{ij} \equiv -\frac{\partial \log q_{ij}}{\partial \log p_j} &= \eta - \eta \frac{\partial \log P_i}{\partial \log p_j} - \frac{\partial \log C_i}{\partial \log P_i} \frac{\partial \log P_i}{\partial \log p_j} \\ &= \eta - (\eta - \theta) \frac{\partial \log P_i}{\partial \log p_j}, \end{aligned} \tag{24}$$

where the second equality uses equation (18). Substituting equation (21) gives

$$\varepsilon_{ij} = \eta - (\eta - \theta)\omega_{ij}. \tag{25}$$

A2.1.4 Firm-level demand elasticities

We now derive the central firm-level result: firm j's demand elasticity is

$$\varepsilon_j = \eta - (\eta - \theta)\omega_j^{\text{eff}}, \tag{26}$$

where ω_j^{eff} is firm j's effective share, defined below.

Total demand for firm j is $Q_j \equiv \sum_i q_{ij}$. Because firm j charges a common price across customers, differentiating $\log Q_j$ with respect to $\log p_j$ yields a sales-weighted average of customer-level elasticities:

$$\varepsilon_j = \sum_i \alpha_{ji} \varepsilon_{ij}, \qquad \alpha_{ji} \equiv \frac{p_j q_{ij}}{\sum_{i'} p_j q_{i'j}}. \tag{27}$$

Substituting the customer-level elasticity expression (22) and defining firm j's effective share as

$$\omega_j^{\text{eff}} \equiv \sum_i \alpha_{ji} \omega_{ij} \tag{28}$$

yields equation (26). Under Bertrand pricing with constant marginal cost, markups equal $\frac{\varepsilon_j}{\varepsilon_j - 1} =$

$\frac{\eta-(\eta-\theta)\omega_j^{\text{eff}}}{\eta-(\eta-\theta)\omega_j^{\text{eff}}-1}$.

When customers do not differ in their relative demand for firms ($a_{ij} = a_j$ for all i), equation (20) implies $\omega_{ij} = \Omega_j$ for all i, where $\Omega_j \equiv \sum_i p_j q_{ij} / \sum_i \sum_{k\in\mathcal{F}} p_k q_{ik}$ is firm j's aggregate share of category spending. Effective share therefore collapses to aggregate market share, $\omega_j^{\text{eff}} = \Omega_j$, and equation (26) reduces to the familiar Atkeson and Burstein (2008) mapping:

$$\varepsilon_j = \eta - (\eta - \theta)\Omega_j.$$

The two formulas differ only when customers exhibit persistent heterogeneity in their relative demand for firms.

A2.2 Nested-logit discrete-choice model

This appendix shows that the effective-share result derived in Appendix A2.1 for the nested-CES benchmark also holds in a nested-logit discrete-choice setting building on Mongey and Waugh (2025). The customer-level analogue of the spending share ω_{ij} is the within-category choice probability ρ_{ij}, and the firm-level demand elasticity again takes the form $\varepsilon_j = \eta - (\eta - \theta)\omega_j^{\text{eff}}$, where effective share is now the sales-weighted average of customer-level choice probabilities.

A2.2.1 Setup

We focus on one category and suppress its index. Conditional on a purchase opportunity within the category, household i chooses one firm j. The utility from choosing firm j in period t is

$$U_{ijt} = -\log(p_j) + \phi_j + \mu_{ij} + \zeta_{ijt}, \tag{29}$$

where p_j is the price, ϕ_j is a common quality shifter, μ_{ij} is a permanent customer–firm match component, and ζ_{ijt} is a transitory taste shock.[55] We assume that ζ_{ijt} follows a generalized extreme-value structure, with parameter η governing substitution across firms within the category.[56] The term μ_{ij} captures persistent differences across households in the attractiveness of firms. Because μ_{ij} is fixed over time while ζ_{ijt} is transitory, customer-level choice probabilities differ persistently across households.

Let

$$v_{ij} \equiv \phi_j + \mu_{ij} - \log(p_j) \tag{30}$$

denote the deterministic component of utility. The within-category choice probability is

$$\rho_{ij} = \frac{\exp\{\eta v_{ij}\}}{\sum_{k\in\mathcal{F}} \exp\{\eta v_{ik}\}}, \tag{31}$$

[55] This functional form ensures that markup variation across firms arises only through demand elasticity. The simplification is not essential for the conclusion that the relevant notion of firm size depends on customer-level shares.

[56] Larger η corresponds to closer within-category substitutes, which is the source of the $1/\eta$ scaling in the inclusive value I_i below.

and the category inclusive value is

$$I_i = \frac{1}{\eta} \log \left(\sum_{k \in \mathcal{F}} \exp\{\eta v_{ik}\} \right). \tag{32}$$

To capture substitution away from the category without specifying the full outer demand system, we assume that customer i's expected category demand, Q_i, depends on the category's attractiveness as summarized by the inclusive value I_i:

$$\frac{\partial \log Q_i}{\partial I_i} = \theta, \qquad \eta > \theta > 1. \tag{33}$$

The inclusive value summarizes the expected utility customer i derives from within-category choices: when prices fall or quality rises within the category, I_i rises and the customer expects to make more purchases. Under a common proportional increase in category prices, this implies a category-demand elasticity of θ. As in the nested-CES benchmark, $\eta > \theta > 1$ means that customers substitute more readily across firms within the category than across categories.

Expected demand from customer i for firm j is therefore

$$q_{ij} = \rho_{ij} Q_i. \tag{34}$$

The choice probability ρ_{ij} is the discrete-choice analogue of the customer-level spending share ω_{ij} of the earlier nested-CES benchmark; it is the latent object that pooled transaction data can recover under repeated purchase opportunities.[57]

As in the nested-CES benchmark, we derive closed-form customer- and firm-level demand elasticities, with the customer-level choice probability ρ_{ij} now playing the role of the spending share ω_{ij}, and arrive at the same firm-level formula $\varepsilon_j = \eta - (\eta - \theta)\omega_j^{\text{eff}}$.

A2.2.2 Customer-level demand elasticities

We now derive the central customer-level result: customer i's demand elasticity for firm j is

$$\varepsilon_{ij} = \eta - (\eta - \theta)\rho_{ij}. \tag{35}$$

This is the discrete-choice analogue of equation (22) in the nested-CES benchmark, with the customer-level choice probability ρ_{ij} playing the role of the customer-level spending share ω_{ij}.

Because $\partial v_{ij} / \partial \log p_j = -1$ and $\partial v_{ik} / \partial \log p_j = 0$ for $k \neq j$, the within-category elasticity is

$$-\frac{\partial \log \rho_{ij}}{\partial \log p_j} = \eta(1 - \rho_{ij}). \tag{36}$$

[57] Fix prices and quality shifters (p_j, ϕ_j) and consider repeated purchase opportunities for household i within the category. If the customer–firm match components μ_{ij} are fixed over time and the transitory taste shocks ζ_{ijt} are i.i.d. across t, then the share of customer i's category purchases made at firm j converges to the latent conditional choice probability ρ_{ij}. If expenditure conditional on a purchase is constant across firms, purchase shares also coincide with expenditure shares. Pooling transactions within a customer over time therefore recovers the persistent customer-level expenditure shares relevant for the share-dependent component of demand elasticity.

The between-category elasticity follows from $\partial I_i / \partial \log p_j = -\rho_{ij}$, which gives

$$-\frac{\partial \log Q_i}{\partial \log p_j} = \theta \rho_{ij}. \tag{37}$$

Combining the two yields

$$\varepsilon_{ij} \equiv -\frac{\partial \log q_{ij}}{\partial \log p_j} = \eta(1 - \rho_{ij}) + \theta \rho_{ij} = \eta - (\eta - \theta)\rho_{ij}, \tag{38}$$

which delivers equation (35).

A2.2.3 Firm-level demand elasticities and effective shares

We now derive the central firm-level result: firm j's demand elasticity is

$$\varepsilon_j = \eta - (\eta - \theta)\omega_j^{\text{eff}}, \tag{39}$$

where effective share ω_j^{eff} is the expected-sales-weighted average of customer-level choice probabilities, defined below.

Because firm j charges a common price across customers, its demand elasticity is the expected-sales-weighted average of customer-level elasticities:

$$\varepsilon_j = \sum_i \alpha_{ji} \varepsilon_{ij}, \tag{40}$$

where

$$\alpha_{ji} \equiv \frac{q_{ij}}{\sum_{i'} q_{i'j}} = \frac{\rho_{ij} Q_i}{\sum_{i'} \rho_{i'j} Q_{i'}} \tag{41}$$

is customer i's share of firm j's expected sales.[58]

Substituting equation (35) into equation (40) gives

$$\begin{aligned} \varepsilon_j &= \sum_i \alpha_{ji} \left[\eta - (\eta - \theta)\rho_{ij}\right] \\ &= \eta - (\eta - \theta) \sum_i \alpha_{ji} \rho_{ij}. \end{aligned} \tag{42}$$

Define firm j's effective share in the nested-logit setting as

$$\omega_j^{\text{eff}} \equiv \sum_i \alpha_{ji} \rho_{ij}. \tag{43}$$

Equation (39) follows immediately.

When customers do not differ in their relative demand for firms ($\mu_{ij} = \mu_j$ for all i), the

[58] If customers have identical category-demand levels, $Q_i = Q$ for all i, then equation (41) simplifies to $\alpha_{ji} = \rho_{ij} / \sum_{i'} \rho_{i'j}$. More generally, customers with greater category demand receive greater weight in firm sales.

conditional choice probability is common across customers, $\rho_{ij} = \rho_j$ for all i. Firm j's aggregate within-category market share is then

$$\Omega_j \equiv \frac{\sum_i q_{ij}}{\sum_i \sum_k q_{ik}} = \frac{\sum_i \rho_j Q_i}{\sum_i Q_i} = \rho_j. \tag{44}$$

Effective share therefore collapses to aggregate market share, $\omega_j^{\text{eff}} = \Omega_j$, and we again get the familiar Atkeson and Burstein (2008) mapping

$$\varepsilon_j = \eta - (\eta - \theta)\Omega_j.$$

As in the nested-CES benchmark, effective and conventional shares differ only when customers exhibit persistent heterogeneity in their relative demand for firms.

The model can also be extended to allow heterogeneous price sensitivity across customers, as in Berry, Levinsohn and Pakes (1995) and Mongey and Waugh (2025). In this case, aggregating customer-level elasticities introduces both a firm-specific average price-sensitivity term and a price-sensitivity-weighted effective share. Recovering elasticity levels then requires either price variation or additional structure, but the distribution of customer spending shares remains key for quantifying the size-dependent component of market power. In Appendix Table A10, we show that our main findings are similar when we account for observable customer characteristics and when we construct the corresponding price-sensitivity-weighted measure.

We can also generalize instead to an environment with additional heterogeneity. In particular, we can allow heterogeneous price sensitivity across customers in addition to persistent heterogeneity in relative demand, building on Berry, Levinsohn and Pakes (1995) and Mongey and Waugh (2025). For example, in Mongey and Waugh (2025), the customer-level extensive-margin elasticity becomes $\varepsilon_{ij} = [\eta(1 - \rho_{ij}) + \theta\rho_{ij}] \cdot MU_i$, where MU_i captures the customer's marginal value of the expenditure change induced by a price change. Under the common CRRA utility specification used in their quantitative analysis, this term is pinned down by the customer's conditional consumption choice and the common curvature parameter.[59] Aggregating these customer-level elasticities then introduces both a firm-specific average price-sensitivity term and a price-sensitivity-weighted effective share.[60] In Appendix Table A10, we show that our main findings are very similar if we control for income and other demographic characteristics, or if we directly apply the MU_i adjustment from their model.

[59] Berry, Levinsohn and Pakes (1995) instead treats this heterogeneity as exogenous variation that is recovered under demand estimation.

[60] Recovering the level of each firm's elasticity in this extension requires either information on price variation as in Berry, Levinsohn and Pakes (1995) or additional structure as in Mongey and Waugh (2025). Although this introduces additional heterogeneity in elasticities across firms, measuring the distribution of customer spending effective shares remains the central empirical task for quantifying size-based market power in particular.

A2.3 Broader intuition beyond the benchmark models

The nested-CES benchmark and the parsimonious nested-logit model above deliver an exact affine relationship between effective shares and demand elasticities:[61]

$$\varepsilon_j = \eta - (\eta - \theta)\omega_j^{\text{eff}}. \tag{45}$$

Effective share is therefore sufficient for the share-dependent component of firm-level elasticity in these benchmarks when customers have common values of η and θ within a category.

The broader intuition does not require this exact sufficient-statistic result. Marshall's second law of demand states that demand becomes more elastic as a firm's relative price rises, or equivalently less elastic as the firm moves down its demand curve. Applied customer by customer, this suggests that demand will tend to be less elastic among customers for whom a firm accounts for a larger share of category spending. In many economically natural environments, such as models with persistent differences in match quality, geography, product availability, or shopping technology, the customer-level spending share ω_{ij} is therefore a natural measure of how important firm j is to customer i and how difficult it is for that customer to substitute away from the firm. Importantly, this mapping is a ceteris paribus relationship: it concerns how elasticity varies with a customer's spending share holding fixed the other features of that customer's demand. It is not automatic in arbitrary demand systems, because cross-customer differences in spending shares may instead reflect shifts in customer-specific demand curves that leave elasticities unchanged. The benchmark models are special in this respect, since in those models the elasticity depends on ω_{ij} regardless of why the share is high.

A reduced-form representation of the ceteris paribus relationship between shares and elasticities is

$$\varepsilon_{ij} = g_i(\omega_{ij}), \qquad g_i'(\omega) < 0. \tag{46}$$

The condition $g_i'(\omega) < 0$ captures the relationship between customer-level spending shares and elasticities holding fixed other determinants of customer-level demand elasticity. Aggregating across customers gives

$$\varepsilon_j = \sum_i \alpha_{ji} g_i(\omega_{ij}). \tag{47}$$

Customer elasticities may depend nonlinearly on spending shares, and customers may differ in price sensitivity or other demand characteristics. This means that two firms with the same effective share may face different elasticities if the distributions of their customer-level spending shares or customer characteristics differ. Nevertheless, effective share remains informative about the customer-dominance margin: whether the customers generating a firm's sales allocate a large or small share of their category spending to that firm.

A Hotelling-style framework provides one concrete example of this broader logic. Suppose customers and firms differ in their locations, and customers make repeated purchases subject to transitory taste shocks. Customers located closer to firm j allocate a larger share of their category purchases to that firm, so persistent differences in customer locations generate variation

[61] For simplicity, in this section we use ω_{ij} to refer to the customer-level share emerging from the underlying demand system whether this reflects the spending share in nested CES or the conditional choice probability in nested logit.

in customer-level spending shares. Customers for whom a firm is relatively more attractive are also less willing to substitute away from it, generating a negative relationship between customer-level spending shares and elasticities consistent with $g_i'(\omega) < 0$. The relationship need not be affine, however, so effective share need not be sufficient for firm-level elasticity outside the benchmark models. Throughout the paper, we therefore interpret effective shares as a measure of customer-firm importance, but note that any exact mapping to demand elasticities depends on the underlying demand structure.

A2.4 Market-level concentration

The benchmark mapping from effective shares to elasticities also delivers a useful market-level result: the sales-weighted average firm elasticity in a market depends on the concentration of customers' spending baskets, not the concentration of firms' aggregate sales. As in the previous subsection, ω_{ij} denotes the customer-level share emerging from the underlying demand system.

The sales-weighted average elasticity is

$$\varepsilon \equiv \sum_j \Omega_j \varepsilon_j = \eta - (\eta - \theta) \sum_j \Omega_j \omega_j^{\text{eff}}. \tag{48}$$

In the representative-household case, $\omega_j^{\text{eff}} = \Omega_j$ for all firms, and the relevant concentration object is the sales Herfindahl index:

$$\sum_j \Omega_j \omega_j^{\text{eff}} = \sum_j \Omega_j^2 = HHI. \tag{49}$$

This linearly declining elasticity in the sales Herfindahl is one reason the prior literature has focused on measuring the concentration of market shares.

With heterogeneous customers, however, the relevant concentration object is different. Letting y_{ij} denote spending of customer i at firm j, $Y_j \equiv \sum_i y_{ij}$, $Y_i \equiv \sum_j y_{ij}$, and $Y \equiv \sum_i Y_i$, the sales-weighted sum of effective shares can be rewritten as

$$\sum_j \Omega_j \omega_j^{\text{eff}} = \sum_j \left(\frac{Y_j}{Y}\right) \left(\sum_i \frac{y_{ij}}{Y_j} \omega_{ij}\right) = \sum_i \left(\frac{Y_i}{Y}\right) \left(\sum_j \omega_{ij}^2\right) \equiv HHI^{\text{cust}}. \tag{50}$$

With heterogeneous customers, the relevant concentration object for size-based market power is therefore the spending-weighted average Herfindahl of individual customers' baskets, not the Herfindahl of firms' aggregate sales.

A3 Appendix: Finite Transactions and Effective Share Measurement

Equation 13 defines effective shares directly from realized spending as the sales-weighted share of each customer's category spending that goes to a firm. While this is the object that enters all of our empirical results, the theoretical framework in Section 2 interprets effective shares as the persistent spending shares that would obtain over many purchases, holding preferences fixed. With a finite number of purchases, the empirical object that we measure can potentially deviate from these persistent shares. This appendix explores the importance of finite transactions for the interpretation of our results. Specifically, it shows that finite samples do not drive our concavity conclusions, provides a justification for the focus on customers with at least 5 transactions in a category, and it provides a correction that removes finite sample bias in our empirical measure relative to the persistent spending shares implied by Section 2.

The potential difference between a firm's empirical effective share and the theoretical share in Section 2 is easiest to see in a benchmark with no persistent customer heterogeneity. Suppose every customer makes n_i independent purchases of equal size and chooses firm j for each transaction with probability Ω_j. By definition, the individual's persistent probability of shopping at firm j is its aggregate market share Ω_j. However, it is straightforward to show that ω_j^{eff} will exceed Ω_j when n_i is small. Let k_{ij} denote the number of customer i's n_i purchases made at firm j. k_{ij} is binomial with parameters (n_i, Ω_j), so customer i's realized share at j is $\omega_{ij} = k_{ij}/n_i$ and i's weight in j's sales is $\alpha_{ji} = k_{ij}/\sum_{i'} k_{i'j}$. Substituting into equation 13 implies that

$$\omega_j^{\text{eff}} \;=\; \sum_i \alpha_{ji}\,\omega_{ij} \;=\; \frac{\sum_i k_{ij}^2/n_i}{\sum_i k_{ij}}.$$

Since $\mathbb{E}[k_{ij}] = n_i\Omega_j$ and $\mathbb{E}[k_{ij}^2] = n_i\Omega_j(1-\Omega_j) + n_i^2\Omega_j^2$, as the number of customers N grows the numerator and denominator converge to $N\Omega_j(1-\Omega_j) + \Omega_j^2\sum_i n_i$ and $\Omega_j\sum_i n_i$, so

$$\omega_j^{\text{eff}} \;\to\; \Omega_j + \frac{1-\Omega_j}{\bar{n}} \;=\; \Omega_j\Big(1-\frac{1}{\bar{n}}\Big) + \frac{1}{\bar{n}}, \tag{51}$$

where $\bar{n}$ is the mean number of transactions per customer $\bar{n} \equiv \sum_i n_i/N$. This means that the customer's persistent shopping probability at firm j is Ω_j but the measured effective share exceeds Ω_j. This is because a customer who randomly purchases from firm j more frequently than implied by Ω_j has both a higher realized share at j and a larger weight in $j's$ sales.

However, this finite transaction wedge does not explain the weak relationship between conventional and effective shares that we document in the paper. We show this in several ways. First, if customers at all firms have the same average number of transactions, then this wedge is simply an affine transformation of Ω_j and so cannot explain the strong concavity in Figure 4. Explaining a concave relationship would require a systematic correlation between firm size and the total number of transactions made by their customers across all firms. While possible, even such a correlation could not explain the quantitative patterns in Figure 4, since the largest firms have effective shares which are below their local market shares.

Importantly, we can go beyond this hypothetical discussion. In particular, the same argument that introduces this finite sample wedge also delivers a correction. The finite-transaction

component arises because each transaction enters both a customer's share and that customer's weight in the firm's sales. Replacing each customer's share with the share computed over the customer's other transactions removes this mechanical bias. Indeed, it is straightforward to recompute firms' "bias-adjusted" effective shares using this leave-one-out procedure. The scatter plot in the left panel of Figure A10 shows that the bias corrected effective share is highly correlated with the raw effective share at the individual establishment level, with a slope of 1.025 and an R^2 of .998. While bias adjusted effective shares are weakly below raw effective shares, there are no systematic patterns by firm size. Given the similarity between raw and bias-adjusted effective shares, for simplicity we focus on the simpler raw effective shares calculations in the paper. However, we have verified that all of our patterns are similar when applying these bias-corrections. The right panel of Figure A10 illustrates this for the key Figure 4 and shows that the concave relationship between effective and conventional market shares is not driven by finite transactions.

We next ask a related but distinct question: how many transactions must an individual make in a category to produce useful information about shopping probabilities. Clearly, if an individual only makes a single transaction, then they will appear with an effective share of one at whatever firm they shop at, regardless of the true distribution of underlying shopping probabilities and so they will not contribute useful information about shopping propensities. Conversely, an individual with a very large number of transactions will have well-observed shopping probabilities, but restricting to individuals with a large number of transactions discards a large share of spending and customers in many categories and thus raises concerns about whether such individuals in the data are representative of typical shopping patterns.

Figure A11 uses a simulation with persistent heterogeneity in shopping probabilities to show how the effective share measured from equation 13 compares to the persistent share under different sample restrictions. We simulate 150 markets, each with between 5 and 60 firms and 10,000 customers calibrated to match the distribution of shopping patterns underlying the local analysis from Figure 4. Within a market, firm sizes are drawn from a log-normal distribution with a standard deviation of 1.5 in logs to approximately match the within-market dispersion of log market shares in the data. Customer i makes n_i purchases of equal size, with $n_i = \lceil \exp(1 + 1.2 z_i) \rceil$ and z_i standard normal. This mimics the distribution of transaction counts in the lower-frequency categories in Table A3. Persistent heterogeneity is introduced through firm-specific loyalty. Each firm j has a loyalty parameter λ_j drawn uniformly on $[0.3, 0.8]$. Each customer is assigned a "preferred" firm $p(i)$, with probability proportional to market shares, and buys from firm j with probability

$$P_{ij} \;=\; (1 - \lambda_{p(i)})\, \Omega_j \;+\; \lambda_{p(i)}\, \mathbf{1}\{j = p(i)\},$$

so that a customer directs a fraction $\lambda_{p(i)}$ of purchases to their preferred firm and spreads the remainder across the market in proportion to market shares. Because λ_j varies across firms, firms with the same market share have different persistent effective shares. The "true" persistent effective share of firm j is equation 13 evaluated at purchase probabilities rather than realized shares, $\omega_j^{\text{true}} = \sum_i n_i P_{ij}^2 / \sum_i n_i P_{ij}$, computed over all customers in the market. Realized purchases are multinomial draws with probabilities P_{ij}, and the measured effective share is computed using equation 13 applied to these realized purchases.

The left panel of Figure A11 shows the relationship between measured effective shares and true underlying effective shares when there are no restrictions on transaction frequency (i.e. every customer and firm is included). If measured shares equaled true effective shares then all points would lie on the 45-degree line. Instead, measured effective shares lie above the 45-degree line, with the largest gap for firms with small shopping probabilities. This pattern reflects the finite transaction bias discussed above. In addition, there is non-trivial variance across firms with the same true effective shares, reflecting the fact that some firms have effective shares computed over a small number of random customers. Thus, with no restrictions on transaction voume, measurement error is indeed a concern. However, it is important to note that even in this extreme case, there is still substantial signal: regressing measured on true effective shares gives a slope of 0.82. The right panel shows that once we restrict to customers with at least five transactions and to firms with at least 200 such customers as in our primary sample, measurement error essentially disappears. Imposing this restriction substantially tightens the relationship, with the correlation rising to 0.995, and is our basis for the use of this restriction throughout our analysis.

A4 Appendix: Details for Firm Entry Analysis

In this section, we provide several additional details of the analysis showing that effective shares predict customer switching in response to firm entry. The key question is whether customers with higher effective shares at an incumbent prior to entry are less likely to switch to a new entrant. If effective shares capture information relevant for customer switching in the face of shocks, then customers with higher effective shares should be less responsive to the new option.

We identify a new establishment in a particular 5-digit ZIP code based on the first quarter in which a firm records a transaction in that ZIP code in our data (meaning that the firm has an establishment there, not just that it sells to a customer who lives there).[62] We require the following conditions in order to consider an establishment as an "entrant": 12 consecutive months of nonzero transaction activity, at least 25 transactions in every month after entry, and at least $50,000 of sales in its first year after entry. Since all analysis in our paper focuses on customers' annual spending, we restrict the sample to establishments that enter in the first quarter of the year.[63]

In order to increase statistical power as well as for computational feasibility, we focus this analysis on a specific set of customers who could feasibly shop at the new store if they wanted to.[64] For each entry event, we identify the 5-digit ZIP that had the most customers visit the

[62]We focus on the first entry of a firm in a 5-digit ZIP rather than entry of additional establishments in ZIP5s where a firm already has a presence. Furthermore, we also drop from our analysis any incumbent establishments that are part of the same firm as the entering establishment (e.g. we do not analyze within-firm switching behavior across establishments).

[63]We independently verified 220 of these entry events using independent audits from Yelp customer review dates and google searches for information on openings. Results are similar if we re-run analysis using this sample.

[64]Computation requires narrowing the set of customers because the sample size in the regression is number of entry events $\times$ the number of exposed customers, which becomes enormous as the number of customers counted as exposed grows. We focus on a location-based measure of exposure since distant stores are unlikely to be rele-

entering firm in the year of entry, and the exposed sample for that event is then the set of *all* consumers who lived in that 5-digit ZIP in both $t-1$ and t (whether or not they shopped at the new entrant).[65] The underlying rationale is that if an individual's neighbors shop at a new entrant, then it was likely a potential shopping option for that individual as well.

In our baseline specification in main text Table 2, we use a 10% sample of all identified firm-entry events. This sample corresponds to around 1,000 entry events and more than 35 million exposed customer $\times$ incumbent observations. However, we find nearly identical results when we restrict to a subset of entry events that we have manually verified. In particular, we manually confirmed a subset of the entry events identified from our card data using Yelp data as well as manual searches for the firm's entry date using various other web sources. This is important because there are various sources of potential spurious entry in the data like firm name changes (e.g. through mergers or acquisitions), re-configuration of payment terminals, etc. Details on how we implemented this can be found in the next subsection.

Appendix Table A9 reports various robustness versions of the entry results in main text Table 2. In column (1), we replace the incumbent's market share $\Omega_{k,e,-1}$ with a customer-specific analog, $\Omega^{market}_{i,e,-1}$, defined as the average ZIP3 market share of the incumbent establishments where customer i shopped in the category in $t-1$, weighted by i's spending at each establishment. For customers who shop at a single incumbent, this equals $\Omega_{k,e,-1}$. Because $\omega_{i,j(k),e,-1}$ varies across customers of the same incumbent while $\Omega_{k,e,-1}$ does not, this provides a conventional-share control that also varies at the customer level.

The vast majority of entry events in our sample are in the restaurant category, since that category has many establishments and a high level of churn. However, Appendix Table A9 shows that results are very similar when focusing on non-restaurant entrants.

We also run a version of this exercise that restricts attention to incumbent firms located in the entrant's 3-digit ZIP code, to focus only on incumbent firms that are "most exposed" to a given entry event. Results are very similar, and this is again true whether restaurants are included or not.

Finally, in Appendix Figure A4, we estimate the relationship between switching to the new entrant and the customer's effective share separately for 200 entry events.[66] We plot the distribution of the coefficients, finding that almost all point estimates are negative. This suggests that this is a general pattern that is not driven by a specific large entry event that may happen to appeal to specific types of customers.

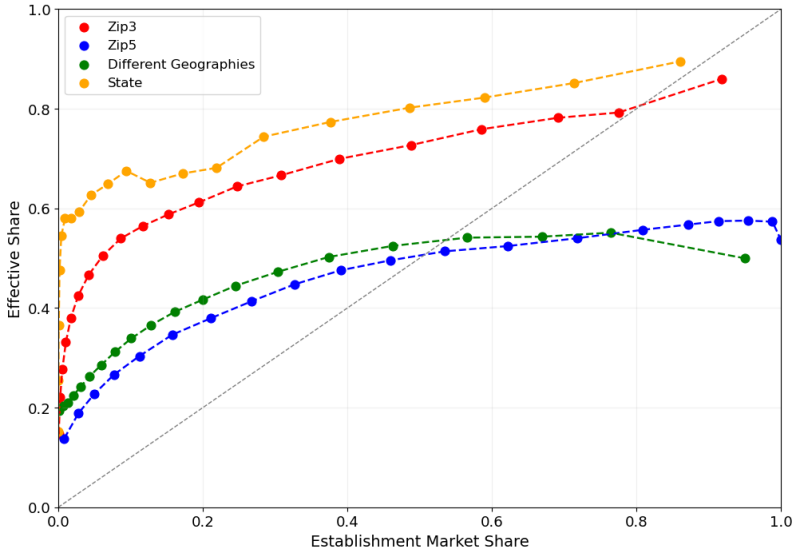
A5 Appendix Tables and Figures

vant options for most customers. Including a broader set of exposed customers in our analysis reduces average switching probabilities but has little effect on the size of regression coefficients *relative* to this average and so does not change our conclusions.

[65] Note that the ZIP with the largest number of customers may or may not be the same as the ZIP of the entrant.

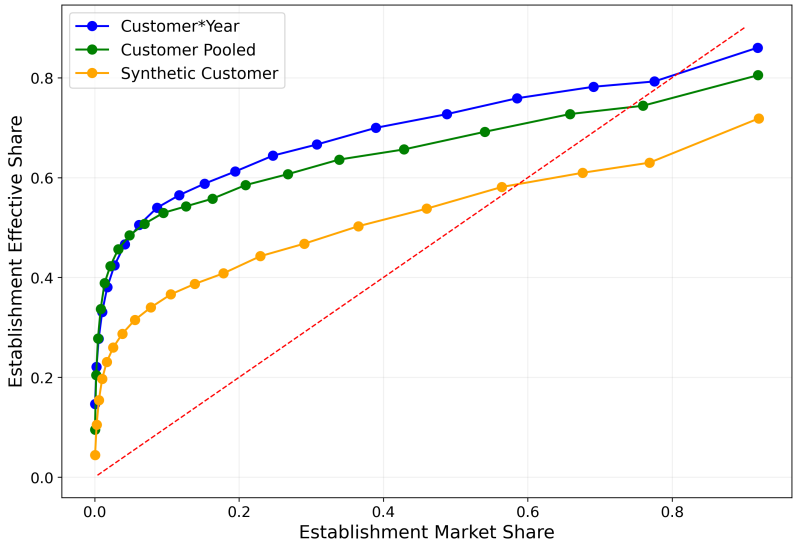
[66] Since this figure focuses on a smaller number of entry events, we use the verified sample discussed above.

Figure A1: Robustness: Alternative Geographic Market Definitions



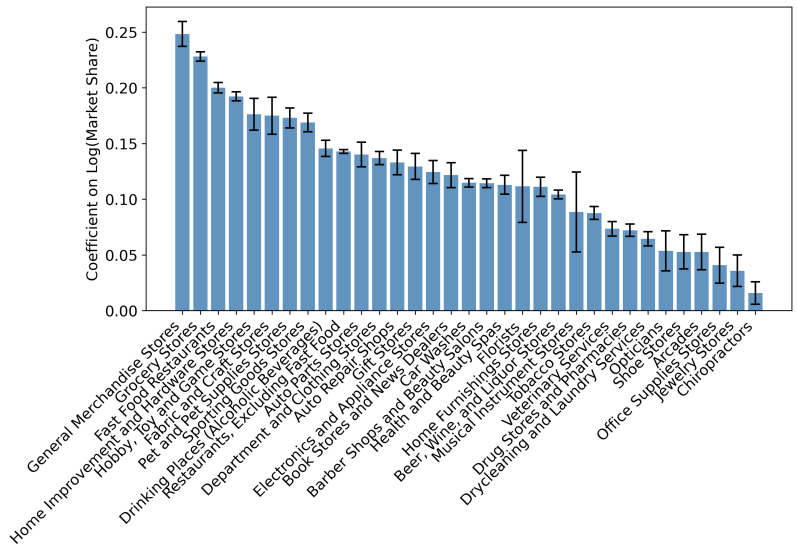
Notes: The dots are binscatter relationships with 20 bins, where the bins are defined by weighting firms by total establishment sales. Each bin plots the mean market effective share against the mean market share. Data is from 2025 and only includes in-person establishments with at least 200 customers. Zip5 Level Stores are establishments and markets defined at the zipcode level. Zip3 Level Stores are establishments and markets defined at the zip3 level. State Level Stores are defined establishments and markets defined at the state level. In the mixed geography version, the follow categories are defined at the state level: Book Stores and News Dealers, Chiropractors, Electronics and Appliance Stores, Fabric and Craft Stores, Gift Stores, Hobby, Toy, and Game Stores, Home Improvement and Hardware Stores, Musical Instrument Stores, and Office Supplies Stores. The following categories are defined at the zip3 level: Arcades, Beer, Wine, and Liquor Stores, Car Washes, Department and Clothing Stores, Drug Stores and Pharmacies, Fast Food Restaurants, Florists, General Merchandise Stores, Grocery Stores, Health and Beauty Spas, Home Furnishings Stores, Opticians, Pet and Pet Supplies Stores, Shoe Stores, Tobacco Stores, and Veterinary Services. The remaining are defined at the 5-digit zip.

Figure A2: Patterns by the level of customer aggregation



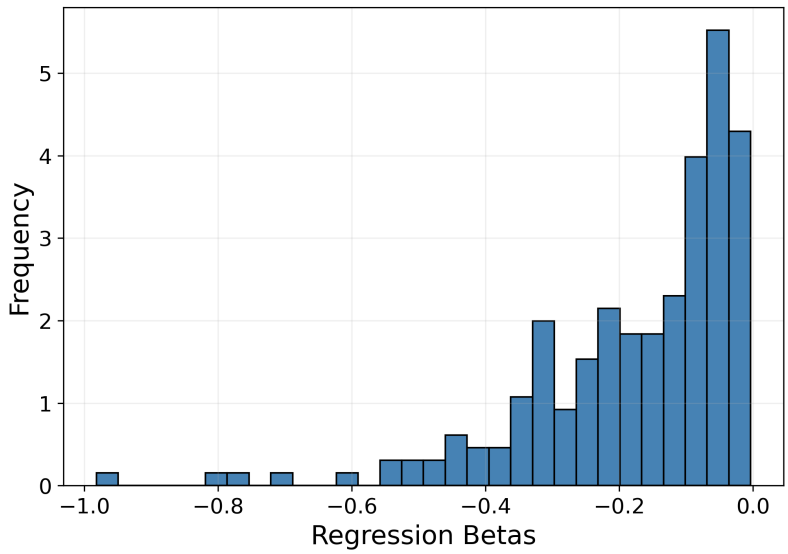
Notes: This figure shows binned scatterplots of effective shares and market shares computed using three distinct aggregation methods: pooling all spending within a year for each customer (our baseline approach), pooling all spending across the years 2021-2025 for each customer, or using our synthetic customer definitions that group together customers with similar demographics in a location. In all three cases, small establishments (as measured by their traditional market shares) tend to have effective shares that are much larger than their market shares, while the largest establishments have effective shares that are below their market shares.

Figure A3: Category-level heterogeneity in relationship between effective and market shares



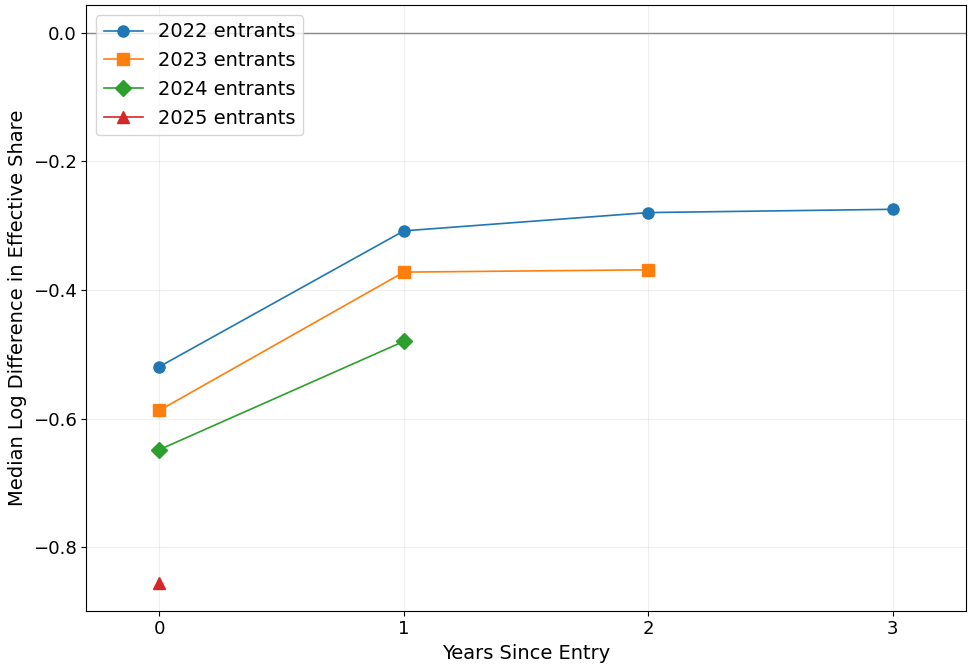
Notes: The figure reports category-specific estimates of the relationship between log firm effective market share and log firm market share. For each category, the plotted bar is the coefficient from a separate regression. Vertical black lines denote 95 percent confidence intervals. Categories are ordered by the estimated coefficient. All observations are unweighted.

Figure A4: The Relationship between Effective Shares and Switching: Distribution



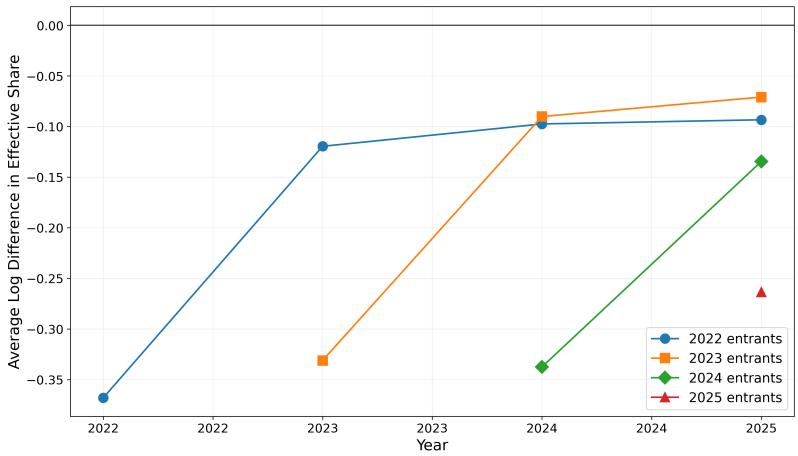
Notes: The figure plots the distribution of the coefficient on $\omega_{ij,e,-1}$ from the specification shown in Column 4 of Table 2 for 200 randomly selected entry events e.

Figure A5: Effective Shares by Cohort in a Balanced Customer Panel



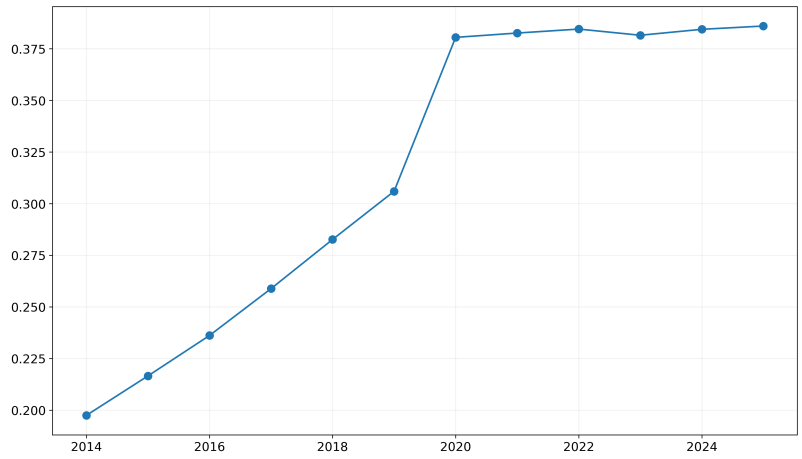
This figure follows cohorts of customers whose first observed purchase from a firm occurs in year c and who continue to purchase from that firm in every year through 2025. For each cohort c and each year $t \geq c$, we compute the log difference between the cohort's effective share in year t and the effective share in year t of the firm's established customers, defined as customers who shop at the firm in all five sample years. The figure plots the median of this firm-level log difference across firms, separately by cohort and years since entry. Only firms with at least twenty customers in all cohort and comparison group are included. Because the sample is restricted to relationships that survive through 2025, the figure holds the set of customers fixed within each cohort as relationship age increases.

Figure A6: Growth through Geographic Expansion: New Establishments



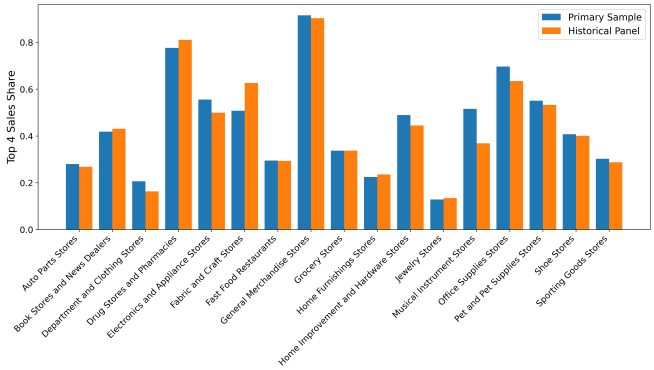
Notes: This figure presents differences in effective market share between newly entered and incumbent ZIP3 markets within the same firm. The sample consists of multi-establishment firms that operated in at least one ZIP3 market in 2021 and subsequently entered at least one additional ZIP3 market between 2022 and 2025. ZIP3 entry cohorts are defined by the first year in which a firm appears in a given ZIP3 market. For each firm and calendar year, effective market shares are averaged separately across (i) incumbent ZIP3 markets in which the firm operated in 2021 and (ii) ZIP3 markets entered in a given cohort year. The plotted outcomes are log differences between these entrant and incumbent averages, computed within firm and then aggregated across firms weighting firms by the share of sales in 2021.

Figure A7: The Growth in Online Spending in Retail Categories

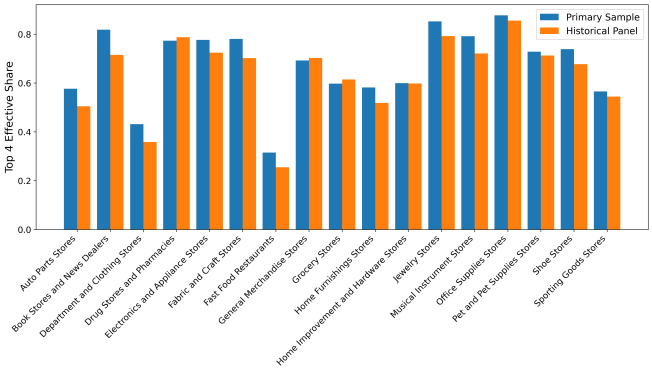


Notes: This figure plots the annual share of spending conducted online in the historical panel across retail categories from 2014 through 2025. Online spending share is measured as the card-not-present ("CNP") share of total sales, aggregated across retail categories in each year. The retail category here includes Shoe Stores; Gas Stations; Sporting Goods Stores; Department and Clothing Stores; Electronics and Appliance Stores; Home Furnishings Stores; Drug Stores and Pharmacies; Auto Parts Stores; Grocery Stores; General Merchandise Stores; Pet and Pet Supplies Stores; Florists; Jewelry Stores; Beer, Wine, and Liquor Stores; Opticians; Home Improvement and Hardware Stores; Book Stores and News Dealers; Gift Stores; Used Merchandise Stores; and Tobacco Stores.

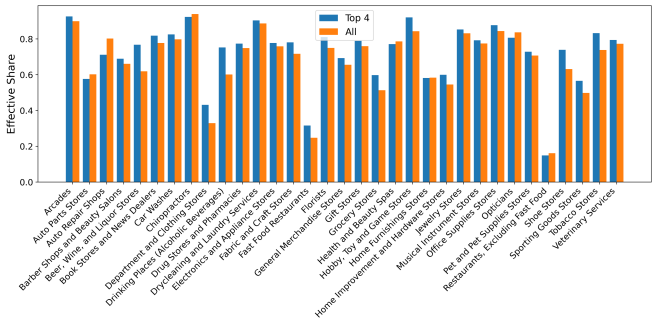
Figure A8: Validation of Historical Panel Data



(a) Top 4 Sales Shares: Primary vs Historical Sample



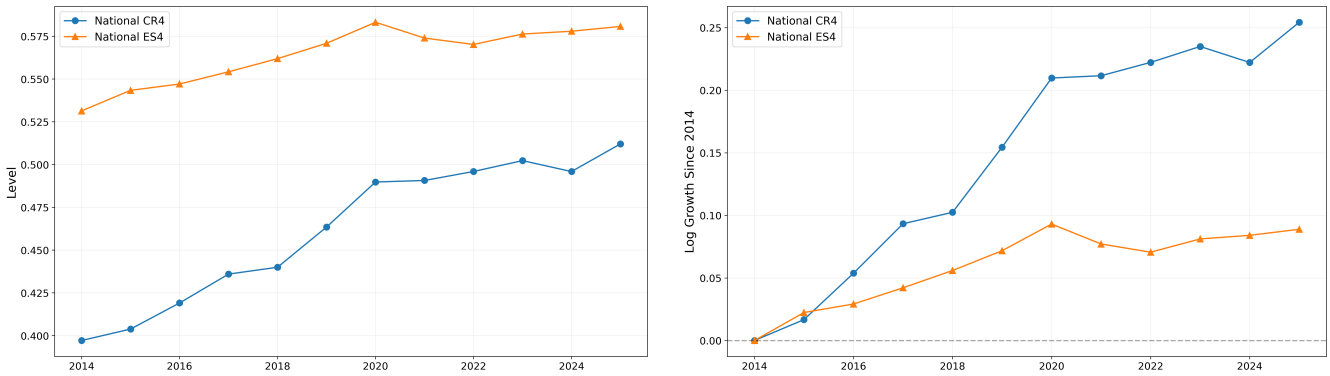
(b) Average of Top 4 Effective Shares: Primary vs Historical Sample



(c) Top 4 vs All Firms Average Effective Share: National

Notes: Panel A compares the four-firm concentration ratio (CR4) across categories in 2022 for two samples: the Primary Sample and the Historical Panel. For each category, CR4 is defined as the combined sales share of the four largest firms. Panel B compares, for the same categories and year, the average effective share of the top four firms (ES4) across the two samples. Panels A and B include only categories observed in both datasets with non-missing values for the relevant measure. Panel C compares, by category in 2022, ES4 with the average effective share computed over all firms in the national sample. "Top 4" denotes the average effective share of the four largest firms in each category, while "All" denotes the corresponding measure computed over the full firm distribution.

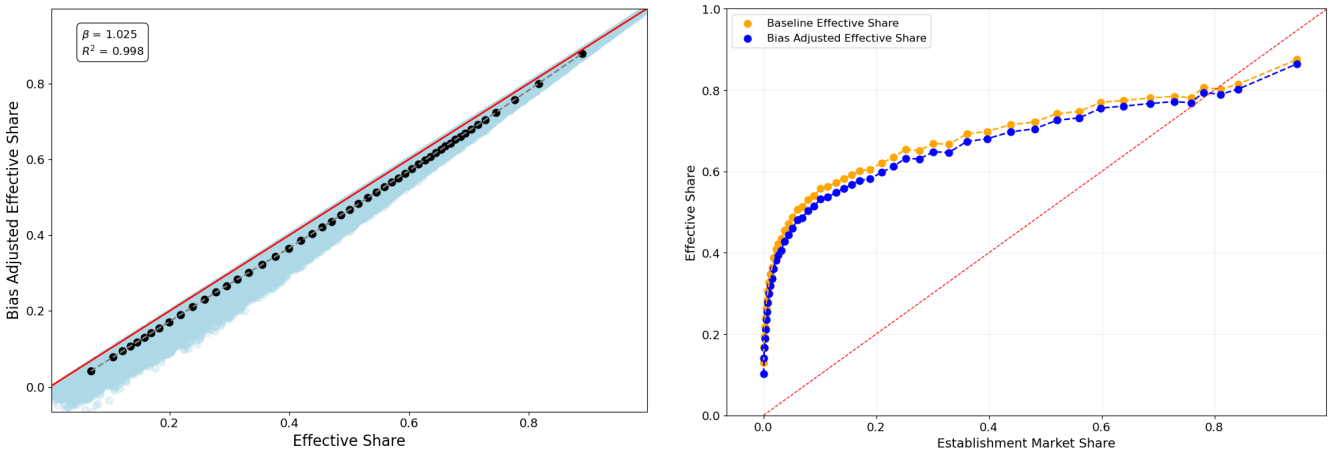
Figure A9: Concentration levels and growth trends w/ time-varying category weights



(a) Concentration levels trend (b) Concentration growth trend

Notes: Panel A reports annual weighted averages of national concentration measures from 2014 to 2025: the four-firm concentration ratio (CR4) and the average effective share of the top four firms (ES4), with ES4 for each category constructed as in footnote 53. National series are constructed by averaging category-level concentration measures using time-varying category total sales weights. Panel B reports cumulative log changes in each series relative to 2014, computed as $\log(X_t) - \log(X_{2014})$.

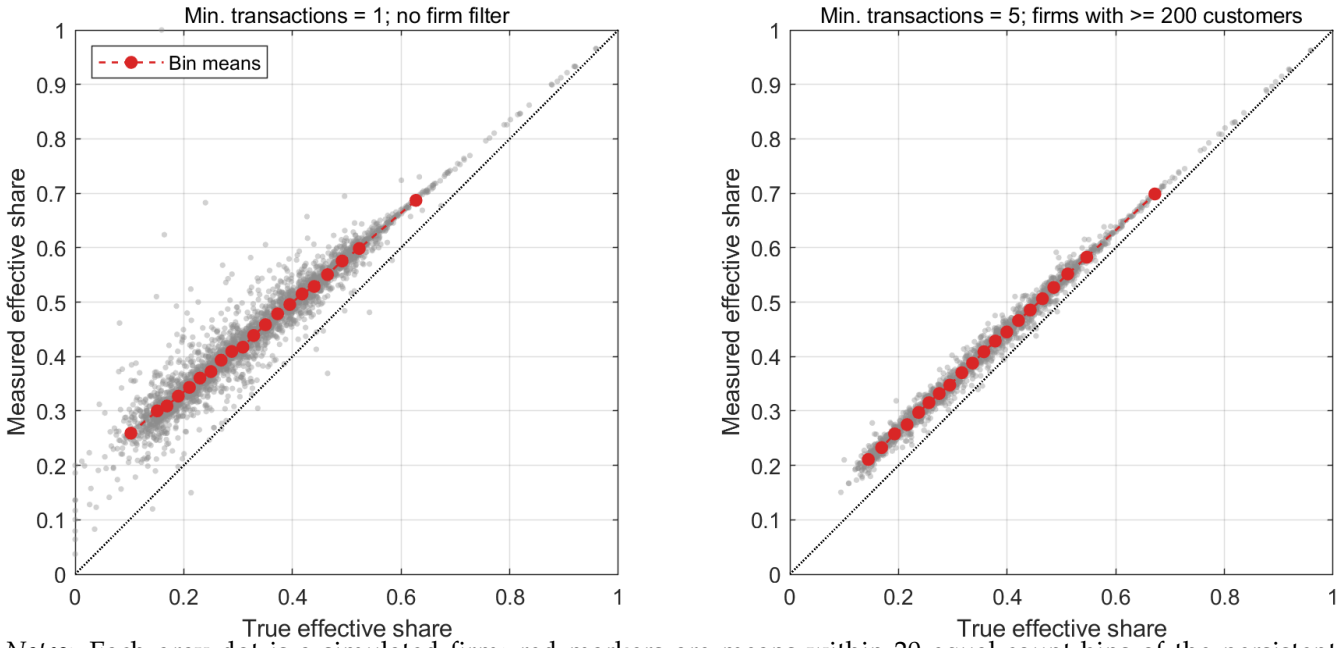
Figure A10: Effects of Finite Transaction Counts



(a) Establishment-specific bias correction (b) Patterns are not driven by finite transactions

Notes: This figure shows the results of correcting for finite sample bias in the data using a leave-one-transaction out effective share construction. The left panel shows a scatter plot of raw and bias-adjusted effective shares at the establishment level, and the right panel recomputes Figure 4 using the raw and bias-adjusted effective shares.

Figure A11: Measured versus persistent effective shares in simulated markets



Notes: Each gray dot is a simulated firm; red markers are means within 20 equal-count bins of the persistent effective share; the dotted line is the 45-degree line. The left panel includes all customers and firms. The right panel restricts to customers with at least five transactions and to firms with at least 200 such customers. The simulation is described in the appendix.

Table A1: Summary Statistics for Various Analysis Samples

	Customer*Year	Customer Pooled	Synthetic Customer	Transaction Data
Customer Count	73.95 million	29.87 million	8.77 million	1.69 billion
Firm Count	2.48 million	2.45 million	2.70 million	4.11 million
Establishment Count	2.66 million	2.67 million	2.77 million	4.68 million
Total Sales	3.06 trillion	2.19 trillion	3.73 trillion	13.31 trillion
Online Share	39.37%	37.85%	39.02%	32.00%
Average Annual Spend	13,803.48	14,689.49	144,864.27	3,141.30

Notes: All columns include data from 2021-2025. The Customer*Year is our baseline sample. Customers are only kept if they meet a minimum activity threshold: median monthly transactions adding across all categories must be at least 5, and customers must not move within the year. Customer pooled expands to the 2021-2025 period. Median monthly transactions over all 5 years must be at least 5 and they must not move over the 5 year period. Synthetic customers pools across similar demographic households within the same census block group and counts these combined observations as a single "customer". Transaction data treats the card as the customer instead of linking cards to people. This means that cards (including debit cards) not linked to TransUnion can be included. Firm count is the number of national firms in the sample. Establishment count refers to the number of in-person "establishments" (our measure combines all establishments of the same firm in the same ZIP3); it does not include online stores. Average annual spend refers to the average annual customer spend — person spend for the first two samples, synthetic customer spend for the synthetic customer, and card spend for the transaction data.

Table A2: Impact of Industry Transaction Filter

	Percent of People Dropped		Percent of Spending Dropped	
Category	5 Txns	10 Txns	5 Txns	10 Txns
Sporting Goods Stores	68.4	86.3	24.7	44.1
Office Supplies Stores	86.2	95.7	48.3	64.0
Arcades	88.4	95.9	53.2	66.2
Shoe Stores	86.4	96.4	53.3	70.4
Veterinary Services	66.8	85.8	31.0	53.3
Jewelry Stores	94.9	98.0	67.3	77.0
Opticians	93.2	98.1	68.9	78.3
Gift Stores	88.2	96.1	52.5	66.4
Drinking Places (Alcoholic Beverages)	75.3	89.0	34.0	51.3
Book Stores and News Dealers	83.0	93.6	42.7	60.3
Tobacco Stores	71.7	84.9	19.6	35.1
Fast Food Restaurants	19.1	32.3	2.0	5.3
Electronics and Appliance Stores	84.3	95.3	53.4	70.9
Barber Shops and Beauty Salons	62.2	83.2	21.3	43.0
Drug Stores and Pharmacies	42.2	62.6	8.5	19.8
Home Improvement and Hardware Stores	45.2	64.8	9.5	20.3
Musical Instrument Stores	85.5	93.7	47.3	63.2
Car Washes	72.4	88.4	35.5	55.3
Chiropractors	65.4	82.0	28.9	45.6
Florists	92.6	97.5	56.7	67.8
General Merchandise Stores	6.5	13.2	0.3	1.1
Restaurants, Excluding Fast Food	19.3	35.0	2.2	6.7
Grocery Stores	10.6	20.8	0.7	2.2
Hobby, Toy and Game Stores	82.0	93.3	39.9	57.6
Beer, Wine, and Liquor Stores	61.2	77.8	14.6	28.4
Auto Parts Stores	78.3	91.5	41.3	58.0
Fabric and Craft Stores	83.5	93.4	36.5	52.0
Auto Repair Shops	88.9	97.3	60.9	77.0
Health and Beauty Spas	77.9	90.7	38.6	58.7
Home Furnishings Stores	74.6	90.9	36.4	56.2
Pet and Pet Supplies Stores	63.3	81.5	19.1	37.8
Department and Clothing Stores	41.8	64.5	8.6	20.3
Drycleaning and Laundry Services	78.1	89.3	27.6	42.4

Notes: Includes anyone who appears in the main sample from 2021-2025. "5 Txns" includes people who have at least five transactions in the industry in a year. "10 Txns" includes people who have at least ten transactions in the industry in a year. This is in comparison to the baseline sample with distance from home, industry, home zip3, and median monthly transaction filters already applied.

Table A3: Average and Median Industry Transactions

	Person*Year		Synthetic Customer		Person Pooled	
Category	Average	Median	Average	Median	Average	Median
Tobacco Stores	16.7	10	31.5	12	32.4	14
Pet and Pet Supplies Stores	15.1	9	47.3	15	31.8	16
Grocery Stores	68.5	36	574.7	137	290.4	147
Opticians	13.6	6	24.2	9	12.1	7
Sporting Goods Stores	14.8	8	55.8	16	28.4	13
Chiropractors	15.7	10	29.7	11	27.5	13
Barber Shops and Beauty Salons	13.4	8	53.8	17	30.7	16
Hobby, Toy and Game Stores	13.5	7	29.9	11	20.6	10
Fabric and Craft Stores	15.3	8	30.9	10	22.6	10
Health and Beauty Spas	12.9	8	28.6	11	22	11
Department and Clothing Stores	20.3	11	116.6	34	54.4	23
Fast Food Restaurants	53	24	400.9	95	196.5	77
General Merchandise Stores	92.9	57	816.8	255	421.4	262
Restaurants, Excluding Fast Food	37.7	20	305.7	71	142.5	68
Car Washes	12.9	8	40.4	14	23.4	12
Drug Stores and Pharmacies	21.3	12	127	34	61.7	28
Veterinary Services	13.2	8	38.3	13	24.7	13
Auto Repair Shops	11.8	6	29.9	12	14.1	9
Shoe Stores	12.1	7	29.5	11	15.3	9
Home Improvement and Hardware Stores	23.1	12	131.8	33	61.2	25
Jewelry Stores	21	8	27.1	8	16.8	7
Home Furnishings Stores	12.7	7	41.4	12	19.1	10
Electronics and Appliance Stores	12.5	7	33	11	15.3	9
Auto Parts Stores	13.8	8	43	17	21.7	11
Musical Instrument Stores	13.3	8	24.6	8	19.5	10
Drycleaning and Laundry Services	17.2	9	38.9	12	29.6	12
Office Supplies Stores	12.5	7	30.1	10	16.3	9
Arcades	15.1	7	22.9	8	18.4	8
Gift Stores	13.8	7	28.7	10	17.1	9
Drinking Places (Alcoholic Beverages)	15.4	8	41.4	14	24.9	11
Beer, Wine, and Liquor Stores	21	10	78.8	22	47.7	18
Florists	14.2	7	26.9	8	14.9	8
Book Stores and News Dealers	14.4	8	33.3	11	20.7	10

Notes: The data is from 2021-2025. For the first four columns, the values are the average and median yearly transaction counts per industry. Columns one and two (three and four) have an activity filter applied that only includes customers (synthetic customers) with at least 5 transactions in that industry in a year. The final two columns are the average and median transaction count in an industry over 5 years. For these two columns, there is an activity filter of at least 5 transactions within the industry over 5 years.

Table A4: Transunion Industry Spending Shares

Category	Studied Categories	Total Spend
General Merchandise Stores	0.2920	0.2586
Grocery Stores	0.1752	0.1552
Restaurants, Excluding Fast Food	0.0958	0.0848
Home Improvement and Hardware Stores	0.0743	0.0658
Department and Clothing Stores	0.0667	0.0591
Fast Food Restaurants	0.0371	0.0329
Home Furnishings Stores	0.0363	0.0322
Sporting Goods Stores	0.0247	0.0219
Electronics and Appliance Stores	0.0235	0.0208
Drug Stores and Pharmacies	0.0202	0.0179
Veterinary Services	0.0193	0.0171
Auto Parts Stores	0.0177	0.0157
Auto Repair Shops	0.0177	0.0157
Beer, Wine, and Liquor Stores	0.0134	0.0119
Jewelry Stores	0.0116	0.0103
Barber Shops and Beauty Salons	0.0110	0.0098
Pet and Pet Supplies Stores	0.0099	0.0088
Opticians	0.0074	0.0065
Shoe Stores	0.0071	0.0063
Drinking Places (Alcoholic Beverages)	0.0052	0.0046
Health and Beauty Spas	0.0047	0.0042
Gift Stores	0.0036	0.0032
Hobby, Toy and Game Stores	0.0034	0.0030
Book Stores and News Dealers	0.0032	0.0028
Car Washes	0.0028	0.0025
Fabric and Craft Stores	0.0024	0.0022
Florists	0.0024	0.0022
Office Supplies Stores	0.0024	0.0021
Tobacco Stores	0.0024	0.0021
Drycleaning and Laundry Services	0.0022	0.0019
Chiropractors	0.0019	0.0017
Musical Instrument Stores	0.0018	0.0016
Arcades	0.0005	0.0004
Total	1.0000	0.8857

Notes: This is for the Transunion sample prior to applying any filters, including activity filters for cards, and distance from home zipcode. The left column is the share of spending of only the industries included in the final analysis sample. The right column represents the industry's share of the total TransUnion sample. Rows are only included for industries in the final sample.

Table A5: The Relationship between Effective Shares and Sales per Customer

	No Fixed Effects			Industry Fixed Effects		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: ZIP3 Panel						
$\log(\text{SPC})$	0.607		0.605	0.512		0.512
	(0.000)		(0.000)	(0.000)		(0.000)
$\log(\text{Customers})$		-0.076	-0.058		0.051	-0.001
		(0.000)	(0.000)		(0.000)	(0.000)
R^2	0.508	0.012	0.515	0.789	0.606	0.789
Partial R^2	0.508	0.012	0.515	0.471	0.012	0.471
Observations	2,596,932	2,596,932	2,596,932	2,596,932	2,596,932	2,596,932
Panel B: National Panel						
$\log(\text{SPC})$	0.581		0.571	0.477		0.481
	(0.000)		(0.000)	(0.001)		(0.001)
$\log(\text{Customers})$		-0.151	-0.098		0.012	-0.022
		(0.001)	(0.000)		(0.000)	(0.000)
R^2	0.512	0.039	0.528	0.788	0.615	0.788
Partial R^2	0.512	0.039	0.528	0.448	0.001	0.450
Observations	1,856,736	1,856,736	1,856,736	1,856,736	1856736	1856736

Notes: The dependent variable in all columns is the log effective share. The regressors of interest are log sales per customer, $\log(\text{SPC})$, and log customers. Panel A reports estimates using the establishment-level panel, and Panel B reports estimates using the national panel. Columns (1)–(3) include no fixed effects; columns (4)–(6) include industry fixed effects. All observations are unweighted. Standard errors are reported in parentheses. R^2 is the regression R^2. Partial R^2 is equal to the regression R^2 in specifications without fixed effects and, in specifications with industry fixed effects, measures the additional explanatory power of the reported regressors relative to a specification containing only industry fixed effects.

Table A6: Effective Shares and Customer Retention with Sales Weights

	(1)	(2)	(3)	(4)	(5)
$\log \omega_{i,j(k),-1}$	0.083***		0.081***	0.084***	0.096***
	(0.002)		(0.002)	(0.001)	(0.003)
$\log \Omega_{k,-1}$		0.036***	0.024***		
		(0.004)	(0.004)		
Market FE	X	X	X	X	X
Establishment-by-Year FE				X	X
Drop in $\log \Omega$					X
Observations	19,005,540	19,005,540	19,005,540	19,004,762	2,468,369
R^2	0.289	0.278	0.290	0.374	0.425
Within R^2	0.017	0.002	0.018	0.018	0.021

Notes: This table shows the relationship between retention and effective shares, where retention is an indicator for shopping at the establishment in the given year. The sample is customer-establishment pairs, restricted to customers who are active in the category in the given year and shopped at the establishment in the previous year. The dataset is further restricted to customer-establishment pairs from a random sample of about 22,000 establishments. Column (5) further restricts the sample to establishments with declines in log conventional market share from $t-1$ to t that exceed the median decline among establishments with falling log market share. Establishments are weighted by sales in $t-1$. All standard errors are clustered at the category-by-zip3-by-year level.

Table A7: Firm-level relationship between effective shares and customer retention

	(1)	(2)	(3)	(4)	(5)	(6)
$\log \omega_{j(k),t-1}$	0.078***		0.104***	0.055***	0.013***	0.087***
	(0.006)		(0.007)	(0.001)	(0.003)	(0.006)
$\log \Omega^k_{t-1}$	0.028**			0.034***	0.058***	-0.007
	(0.013)			(0.001)	(0.013)	(0.006)
$\log$ Sales / Customer		0.049***	-0.022			
		(0.013)	(0.014)			
$\log$ # of Customers		0.012	0.017*			
		(0.016)	(0.009)			
Firm FE				X		
Synthetic Customer Shares					X	
Nationally Defined Firm						X
Observations	1,848,909	1,848,909	1,848,909	1,776,845	1,848,909	1,286,066
R^2	0.470	0.340	0.467	0.849	0.367	0.425
Within R^2	0.261	0.080	0.257	0.144	0.117	0.194

Notes: This table looks at the relationship between effective shares and the share of establishment sales in $t-1$ that belong to customers who continue to shop at the establishment in t. The sample is at the establishment level, aggregated from customer-level data, where the customer shopped at the establishment in $t-1$ and continued to shop within the category in t. Columns (2) and (3) refer to the log sales per customer and the log number of customers in $t-1$. Column (5) uses establishment effective shares aggregated from synthetic customers, which pools similar demographic households across the same block group into a single customer. The synthetic customers are pooled over age, income, and car ownership demographic groups. Column (6) uses data aggregated to the national firm level, instead of the establishment level, continuing to aggregate from customer-level data where the customer shopped at the firm in $t-1$ and remains active in the category in t. All independent variables are standardized within the category. Standard errors are clustered at the category level.

Table A8: Customer Retention over Longer Horizons

	Loyalty				Effective Share		
	t	$t+1$	$t+2$	All years	t	$t+1$	$t+2$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\log \omega_{i,j(k),t-1}$	0.033***	0.031***	0.028***	0.041***			
	(0.001)	(0.001)	(0.001)	(0.001)			
$\log \Omega_{k,t-1}$	0.031***	0.030***	0.026***	0.013***			
	(0.003)	(0.003)	(0.004)	(0.003)			
$\omega_{i,j(k),t-1}$					0.350***	0.304***	0.270***
					(0.006)	(0.007)	(0.007)
$\Omega_{k,t-1}$					0.048**	0.046	0.026
					(0.024)	(0.036)	(0.044)
Market FE	X	X	X	X	X	X	X
Observations	9,183,632	7,954,376	7,125,073	5,727,440	9,183,632	7,954,376	7,125,073
R^2	0.110	0.141	0.164	0.165	0.153	0.138	0.134
Within R^2	0.007	0.007	0.006	0.014	0.084	0.055	0.039

Notes: Columns (1), (2), (3), and (4) show the relationship between log effective shares in $t-1$ and customer retention in different time periods, where retention is an indicator for shopping at the establishment in the given future time period (or all time periods in column (4)). Columns (5), (6), and (7) show the relationship between customer-level effective shares in $t-1$ and in the given future time period. All columns use customer-by-establishment data, where the customer shopped at the establishment in $t-1$ and is active in the category in the given time period (all years for column (4)). The customer-by-establishment data is a random sample of about 22,000 establishments. The market fixed effects is defined as category-ZIP3-year fixed effects. All standard errors are clustered at the category-by-ZIP3-by-year level.

Table A9: Robustness Checks for Consumer Responses to Firm Entry

	(1)	(2)	(3)	(4)
$\log \omega_{i,j(k),t-1}$	-0.028***	-0.030***	-0.029***	-0.028***
	(0.001)	(0.003)	(0.001)	(0.004)
$\log \Omega^{\text{market}}_{i,t-1}$	-0.001	-0.007***		
	(0.001)	(0.002)		
$\log \Omega^{k}_{t-1}$			-0.008***	-0.015***
			(0.001)	(0.002)
Exlude Restaurants?		X		X
Incumbents Included	All	All	Entrant's ZIP3	Entrant's ZIP3
Observations	35,164,534	2,869,751	20,598,751	1,820,632
R^2	0.062	0.086	0.058	0.089
Within R^2	0.012	0.017	0.015	0.024

Notes: Columns (1) and (2) show the relationship between customer effective shares, the customer sales-weighted average conventional market share in the given category, and customers' response to firm entry. Customers are indicated as switching to the entering firm if they do not shop at the incumbent establishment in t and shop at the entering establishment. The sample is customer-by-incumbent-establishment-by-entering-establishment triplets, where the customer shopped at the incumbent establishment in $t-1$ and remained active in the category in t. Columns (3) and (4) show the relationship between customer effective shares, establishment conventional market share, and customer responses to entering firms, where the sample is limited to incumbents in the same ZIP3 as the entering establishment. All columns include entry event fixed effects and have standard errors clustered at the entry event.

Table A10: Robustness Checks for Log Relationship between Effective and Market Shares

	(1) Base	(2) Sales Wtd	(3) Firm Spec. Prod	(4) Service	(5) Retail	(6) Demo Control	(7) MU	(8) Card	(9) Card & Debit
Panel A: Local (ZIP3)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log \Omega_t^j$	0.178***	0.175***	0.150***	0.150***	0.211***	0.185***	0.178***	0.088***	0.070***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.727	0.789	0.655	0.678	0.740	0.746	0.727	0.831	0.842
Observations	523,499	523,499	339,697	348,339	175,160	506,141	523,499	753,452	997,446
Within R^2	0.248	0.340	0.162	0.163	0.385	0.300	0.248	0.107	0.079
Panel B: National									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\log \Omega_t^j$	0.161***	0.084***	0.148***	0.148***	0.180***	0.168***	0.161***	0.082***	0.081***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
R^2	0.669	0.714	0.626	0.648	0.660	0.689	0.669	0.800	0.807
Observations	378,536	378,536	253,341	261,339	117,197	373,570	378,536	548,674	744,969
Within R^2	0.167	0.082	0.139	0.140	0.216	0.216	0.167	0.076	0.083

Notes: Column (1) shows the baseline sample results. Column (2) uses the same sample as (1), weighting the results by firm sales, winsorized at the 97.5th percentile. Column (3), firm-specific products, subsets the sample to include only categories where firms produce one product. These categories are auto repair shops, barber shops and beauty salons, car washes, chiropractors, florists, health and beauty spas, veterinary services, fast food restaurants, restaurants excluding fast food, and dry cleaning and laundry services. Column (4) only includes service categories, and column (5) restricts the sample to retail categories. Column (6) controls for each firm's weighted average customer demographics. The demographics include log income, log age, the log of the customer's ZIP code's population density, the customer's gender, and whether the customer is a car owner. Column (7) weights effective shares by customers' marginal utility, which relies on their income. Customer marginal utility is computed under a constant relative risk aversion specification with a curvature parameter of 1.57. Income and marginal utility are industry-normalized. Customer-level effective shares receive a market power adjustment of $8.9(1 - \omega_{i,j})$. Column (8) aggregates firm-level effective shares from only credit card-level data. Column (9) aggregates to firm-level effective shares from credit card and debit card data.

Table A11: Four-Firm Concentration and Effective Share by Retail Category, 2014 and 2025

	2014		2025	
Industry	CR4 (%)	ES4 (%)	CR4 (%)	ES4 (%)
General Merchandise Stores	72.4	71.4	90.6	72.3
Drug Stores and Pharmacies	81.0	77.7	69.5	79.1
Fabric and Craft Stores	57.1	61.8	63.6	74.0
Office Supplies Stores	75.6	80.7	60.5	84.2
Pet and Pet Supplies Stores	48.7	74.6	53.8	73.9
Home Improvement and Hardware Stores	49.2	58.4	48.5	60.5
Musical Instrument Stores	32.8	71.1	45.6	75.6
Electronics and Appliance Stores	49.2	70.4	45.0	72.5
Book Stores and News Dealers	41.6	72.3	43.2	70.2
Shoe Stores	44.0	67.4	37.3	66.4
Sporting Goods Stores	25.9	51.0	34.9	56.4
Grocery Stores	35.4	56.5	33.9	62.7
Auto Parts Stores	25.5	52.3	31.0	53.6
Fast Food Restaurants	26.2	24.0	30.1	26.1
Home Furnishings Stores	21.1	54.6	25.3	55.0
Department and Clothing Stores	18.4	31.3	19.0	38.0
Jewelry Stores	16.2	90.0	13.8	75.5

Notes: The table reports the levels between 2014 and 2025 in the four-firm concentration ratio (CR4) and the average effective share of the top four firms (ES4) by category at the national level. Effective shares are constructed from the historical panel, with customers defined at the card level. All values are reported in percentage points.

Table A12: Median Log Difference in 2025 Effective Shares Between New and Continuing Customers, By Industry

Industry	Median
Grocery Stores	-0.58
Department and Clothing Stores	-0.22
Drug Stores and Pharmacies	-0.38
Electronics and Appliance Stores	-0.09
Home Furnishings Stores	-0.08
General Merchandise Stores	-0.44
Auto Parts Stores	-0.14
Auto Repair Shops	-0.17
Barber Shops and Beauty Salons	-0.39
Beer, Wine, and Liquor Stores	-0.72
Book Stores and News Dealers	-0.41
Car Washes	-0.35
Chiropractors	-0.12
Drinking Places (Alcoholic Beverages)	-0.55
Florists	-0.24
Gift Stores	-0.21
Health and Beauty Spas	-0.31
Home Improvement and Hardware Stores	-0.09
Jewelry Stores	-0.14
Opticians	-0.08
Pet and Pet Supplies Stores	-0.39
Shoe Stores	-0.19
Sporting Goods Stores	-0.12
Tobacco Stores	-0.45
Veterinary Services	-0.17
Fast Food Restaurants	-0.43
Restaurants, Excluding Fast Food	-0.46
Drycleaning and Laundry Services	-0.23
Fabric and Craft Stores	-0.32
Hobby, Toy and Game Stores	-0.32
Musical Instrument Stores	-0.10
Office Supplies Stores	-0.19
Arcades	-0.54

Notes: For each firm, we compute 2025 effective shares separately for new customers, who purchase from the firm in 2025 but not in 2021–2024, and continuing customers, who purchase in 2025 and in at least one year from 2021–2024. Each group's effective share is calculated using customer sales weights normalized within the group. We then subtract the log effective share of continuing customers from that of new customers. The table reports the unweighted median of this firm-level log difference within each category. Only firms with more than 20 customers in each group are included. Negative values indicate lower effective shares among new customers.

Table A13: Mapping of Merchant Category Codes (MCC) to Industry Groups and Census Industries

MCC	MCC Name	Industry Group	Census Industry
7932	BILLIARD/POOL ESTABLISHMENT	Arcades	Amusement, Gambling, & Rec.
7933	BOWLING ALLEYS	Arcades	Amusement, Gambling, & Rec.
5013	MOTOR VEHICLE SUPPLY/PARTS	Auto Parts Stores	Automotive Parts & Tire Stores
5531	AUTO/HOME SUPPLY STORES	Auto Parts Stores	Automotive Parts & Tire Stores
5532	AUTOMOTIVE TIRE STORES	Auto Parts Stores	Automotive Parts & Tire Stores
5533	AUTOMOTIVE PARTS STORES	Auto Parts Stores	Automotive Parts & Tire Stores
7531	AUTO BODY REPAIR SHOPS	Auto Repair Shops	Auto Repair & Maintenance
7534	TIRE RETREAD/REPAIR SHOPS	Auto Repair Shops	Auto Repair & Maintenance
7535	AUTO PAINT SHOPS	Auto Repair Shops	Auto Repair & Maintenance
7538	AUTO SERVICE SHOPS/NON DEALER	Auto Repair Shops	Auto Repair & Maintenance
7230	BEAUTY/BARBER SHOPS	Barber Shops and Beauty Salons	Beauty salons
5309	DUTY FREE STORES	Beer, Wine, and Liquor Stores	Beer, wine, and liquor stores
5921	PKG STORES/BEER/WINE/LIQUOR	Beer, Wine, and Liquor Stores	Beer, wine, and liquor stores
5192	BOOKS/PERIODICALS/NEWSPAPERS	Book Stores and News Dealers	Book stores and news dealers
5942	BOOK STORES	Book Stores and News Dealers	Book stores and news dealers
5994	NEWS DEALERS/NEWSSTANDS	Book Stores and News Dealers	Book stores and news dealers
7542	CAR WASHES	Car Washes	Car washes
8041	CHIROPRACTORS	Chiropractors	Offices of chiropractors
5311	DEPARTMENT STORES	Department and Clothing Stores	Clothing and Accessories Stores
5611	MEN/BOYS CLOTHING/ACC STORES	Department and Clothing Stores	Clothing and Accessories Stores
5621	WOMENS READY TO WEAR STORES	Department and Clothing Stores	Clothing and Accessories Stores
5631	WOMENS ACCESS/SPECIALTY	Department and Clothing Stores	Clothing and Accessories Stores
5641	CHILDREN/INFANTS WEAR STORES	Department and Clothing Stores	Clothing and Accessories Stores
5651	FAMILY CLOTHING STORES	Department and Clothing Stores	Clothing and Accessories Stores
5681	FURRIERS AND FUR SHOPS	Department and Clothing Stores	Clothing and Accessories Stores
5691	MENS/WOMENS CLOTHING STORES	Department and Clothing Stores	Clothing and Accessories Stores
5699	MISC APPAREL/ACCESS STORES	Department and Clothing Stores	Clothing and Accessories Stores
5948	LUGGAGE/LEATHER STORES	Department and Clothing Stores	Clothing and Accessories Stores
5965	COMBINATION CATALOG & RETAIL	Department and Clothing Stores	Clothing and Accessories Stores
5813	BARS/TAVERNS/LOUNGES/DISCOS	Drinking Places (Alcoholic)	Drinking Places (Alcoholic)
5912	DRUG STORES & PHARMACIES	Drug Stores and Pharmacies	Pharmacies and drug stores
5975	HEARING AID/SALES/SERVICE	Drug Stores and Pharmacies	Pharmacies and drug stores
5976	ORTHOPEDIC GOODS	Drug Stores and Pharmacies	Pharmacies and drug stores
5997	ELEC RAZOR STORES/SALE/SERV	Drug Stores and Pharmacies	Pharmacies and drug stores
7210	LAUNDRY/CLEANING/GARMENT SV	Drycleaning & Laundry	Drycleaning & Laundry
7211	LAUNDRIES-FAMILY/COMMERCIAL	Drycleaning & Laundry	Drycleaning & Laundry
7216	DRY CLEANERS	Drycleaning & Laundry	Drycleaning & Laundry
5044	OFFICE/PHOTO EQUIPMENT	Electronics & Appliance Stores	Electronics & Appliance Stores
5045	COMPUTERS/PERIPHERALS/SOFTWARE	Electronics & Appliance Stores	Electronics & Appliance Stores
5722	HOUSEHOLD APPLIANCE STORES	Electronics & Appliance Stores	Electronics & Appliance Stores
5732	ELECTRONICS STORES	Electronics & Appliance Stores	Electronics & Appliance Stores
5946	CAMERA & PHOTO SUPPLY STORES	Electronics & Appliance Stores	Electronics & Appliance Stores
5131	PIECE GOOD/NOTIONS/DRY GOOD	Fabric and Craft Stores	Sewing & Piece Goods Stores
5949	FABRIC STORES	Fabric and Craft Stores	Sewing & Piece Goods Stores
5970	ARTIST/CRAFT SHOPS	Fabric and Craft Stores	Sewing & Piece Goods Stores
5193	FLORIST SUPPLIES/NURSERY STOCK	Florists	Florists
5992	FLORISTS	Florists	Florists
5947	GIFT, CARD, NOVELTY STORES	Gift Stores	Gift, Novelty, & Souvenir Stores
5300	WHOLESALE CLUBS	General Merchandise Stores	General merchandise stores

5310	DISCOUNT STORES	General Merchandise Stores	General merchandise stores
5331	VARIETY STORES	General Merchandise Stores	General merchandise stores
5399	MISC GENERAL MERCHANDISE	General Merchandise Stores	General merchandise stores
5411	GROCERY STORES/SUPERMARKETS	Grocery Stores	Grocery stores
5422	FREEZER/MEAT LOCKERS	Grocery Stores	Grocery stores
5441	CANDY/NUT/CONFECTION STORES	Grocery Stores	Grocery stores
5451	DAIRY PRODUCT STORES	Grocery Stores	Grocery stores
5462	BAKERIES	Grocery Stores	Grocery stores
7297	MASSAGE PARLORS	Health and Beauty Spas	Beauty salons
7298	HEALTH & BEAUTY SPAS	Health and Beauty Spas	Beauty salons
5945	HOBBY, TOY & GAME STORES	Hobby, Toy and Game Stores	Hobby, toy, and game stores
5712	FURNITURE/EQUIP STORES	Home Furnishings Stores	Furniture & Furnishings Stores
5719	MISC HOME FURNISHING SPECIALTY	Home Furnishings Stores	Furniture & Furnishings Stores
5950	GLASSWARE/CRYSTAL STORES	Home Furnishings Stores	Furniture & Furnishings Stores
5039	CONSTRUCTION MATERIALS - DEF	Home Improvement/Hardware	Hardware stores
5072	HARDWARE EQUIPMENT/SUPPLIES	Home Improvement/Hardware	Hardware stores
5074	PLUMBING/HEATING EQUIPMENT	Home Improvement/Hardware	Hardware stores
5198	PAINT, VARNISHES & SUPPLIES	Home Improvement/Hardware	Hardware stores
5200	HOME SUPPLY WAREHOUSE STORES	Home Improvement/Hardware	Hardware stores
5211	LUMBER/BUILD. SUPPLY STORES	Home Improvement/Hardware	Hardware stores
5231	GLASS/PAINT/WALLPAPER STORES	Home Improvement/Hardware	Hardware stores
5251	HARDWARE STORES	Home Improvement/Hardware	Hardware stores
5261	NURSERIES, LAWN/GARDEN SUPPLY	Home Improvement/Hardware	Hardware stores
5713	FLOOR COVERING STORES	Home Improvement/Hardware	Hardware stores
5714	DRAPERY & UPHOLSTERY STORES	Home Improvement/Hardware	Hardware stores
5718	FIREPLACES & ACCESSORIES	Home Improvement/Hardware	Hardware stores
5998	TENT AND AWNING SHOPS	Home Improvement/Hardware	Hardware stores
5094	PRECIOUS STONES/METALS/JEWELRY	Jewelry Stores	Jewelry stores
5944	JEWELRY STORES	Jewelry Stores	Jewelry stores
5733	MUSIC STORES/PIANOS	Musical Instrument Stores	Music Instrum./Supplies Stores
5111	STATIONERY/OFFICE SUPPLIES	Office Supplies Stores	Office Supply/Stationery Stores
5943	STATIONERY STORES	Office Supplies Stores	Office Supply/Stationery Stores
8042	OPTOMETRISTS/OPHTHALMOLOGISTS	Opticians	Offices of optometrists
8043	OPTICIANS	Opticians	Offices of optometrists
8044	OPTICAL GOODS & GLASSES	Opticians	Offices of optometrists
5995	PET STORES/FOOD & SUPPLY	Pet and Pet Supplies Stores	Pet and pet supplies stores
5811	CATERERS	Restaurants (Not Fast Food)	Full-service restaurants
5812	EATING PLACES AND RESTAURANTS	Restaurants (Not Fast Food)	Full-service restaurants
5814	FAST FOOD RESTAURANTS	Fast Food Restaurants	Limited-service restaurants
5661	SHOE STORES	Shoe Stores	Shoe stores
5655	SPORTS/RIDING APPAREL STORES	Sporting Goods Stores	Sporting goods stores
5940	BICYCLE SHOPS/SALES/SERVICE	Sporting Goods Stores	Sporting goods stores
5941	SPORTING GOODS STORES	Sporting Goods Stores	Sporting goods stores
5993	CIGAR STORES/STANDS	Tobacco Stores	Tobacco stores
742	VETERINARY SERVICES	Veterinary Services	Veterinary services

Notes: Industry Group refers to the industry assigned to an MCC in the analysis sample. Census Industry is the industry that the Census assigns to the MCC code.