# Concentrating on Customers *

Alex Blumenfeld — UC Berkeley
Christina Patterson — Chicago Booth & NBER
Joseph Vavra — Chicago Booth & NBER

June 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 credit- and debit-card transactions covering the near universe of firms in a broad set of U.S. consumer-facing retail and service categories, we measure firms' "effective" market shares, defined as their shares of category spending among the particular customers they serve. Effective shares rise only weakly with conventional measures of firm size, meaning that large differences in aggregate market shares correspond to much smaller differences in effective shares. This distinction is economically meaningful: effective shares better predict persistent customer relationships and switching responses to new entrants. This weak mapping reflects two opposing forces. As firms grow, they deepen relationships with existing customers, raising effective shares, but they also reach more marginal customers, whose lower effective shares offset this increase. Prominent sources of recent firm growth, including geographic expansion and e-commerce, tilt toward reach rather than depth. Thus, aggregate sales concentration has risen substantially over the last decade without a commensurate increase in average effective shares. Interpreted through the lens of a variable-markup model with persistent customer heterogeneity, this empirical pattern suggests that rising concentration has not led to a comparable 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, the Empirical Macro Workshop-Phoenix and the Federal Reserve Bank of Philadelphia 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 credit- and debit-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 spending among the customers it serves. By measuring these customer-level spending shares, effective shares capture information about a firm's importance to its customers. Comparing effective shares with conventional measures of firm size 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 in one year predict customer retention in future years better than conventional market shares do. Third, we show that the weak relationship between effective and conventional market shares reflects a tension between customer reach and depth as firms grow. Increasing sales among incumbent customers deepens existing customer relationships and raises effective shares, whereas acquiring new customers expands reach but lowers effective shares because new customers typically have shallower relationships with the firm. Fourth, we show that this distinction reshapes our understanding of recent trends. Prominent sources of firm growth, including geographic expansion and e-commerce, disproportionately expand firms' reach rather than deepen their customer relationships. More broadly, we find that over the last decade, superstar firms have substantially increased their conventional market shares but have not seen similar increases in 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), the size-dependent component of a firm's demand elasticity declines with its aggregate market share. We show that when this framework is generalized to incorporate persistent customer heterogeneity, the size-dependent elasticity instead declines with the firm's effective 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] Our finding that effective shares predict the persistence of customer relationships provides one source of empirical support for this interpretation. We also provide more direct evidence from competitor entry, showing that customers are less likely to switch away from firms that previously accounted for a larger share of their spending. Taken together, our results suggest that conventional market shares overstate both differences in size-based market power across firms and its growth 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. We use comprehensive 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. These data allow us to measure spending shares at particular firms or establishments within a category, but not spending on specific products or UPCs sold by those merchants.[3] This means that our results are informative about customer dominance and size-based market power at the merchant level rather than about the demand for specific products or manufacturers within a store. 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 to aggregate spending across multiple cards per individual.[4]

Our empirical analysis begins by comparing effective shares to conventional market shares. Effective shares are substantially compressed relative to conventional market shares. The implied firm "size" distribution is therefore much less skewed when measured using effective shares: large differences in conventional market shares correspond to much smaller differences in effective shares. We begin nationally, since national market shares are most easily measured in other data and have received substantial attention in past empirical work. However, similar compression also holds when we compare effective shares to local market shares. This comparison is important, because conventional measures of market shares and their trends depend

[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]TransUnion data also contain demographic information like customer ZIP-11, age and income.

meaningfully on the level of geographic aggregation (Autor et al. (2020), Rossi-Hansberg, Sarte and Trachter (2021), Autor, Patterson and Van Reenen (2023)).[5] One benefit of our approach is that effective shares sidestep the need to define market geography since they impose no restrictions on where customers shop, allowing geographic scope to be revealed by customers' actual shopping behavior rather than imposed ex ante.[6]

Although local market shares are closer in level to effective shares, their relationship is strongly concave—smaller firms tend to have effective shares above their local market shares, while larger firms have effective shares below them. This concavity 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., purchasing online). Together, these forces imply that conventional market shares overstate the dominance of large firms and understate that of small firms. This matters for work studying markup distortions and misallocation (e.g., Baqaee and Farhi (2020), Edmond, Midrigan and Xu (2023), Afrouzi, Drenik and Kim (2025)), where the distribution of firm size is a key input. Relative to a customer-based notion of size, conventional market shares imply a more skewed size distribution with greater concentration.

Our second finding is that *current* effective shares predict *future* customer behavior. High effective shares mechanically reflect realized spending concentration within a measurement year, but we provide two pieces of evidence that effective shares predict behavior beyond the measurement year. First, customers with higher effective shares at a firm this year are substantially more likely to continue shopping there in future years.[7] 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 simply firm-level scale or transitory variation in realized purchases.[8]

Beyond their direct implications for customer behavior, these patterns are also consistent with the market-power interpretation discussed above, in which higher effective shares cor-

[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]Importantly, this does not simply reflect total firm scale: a firm's conventional market share has no additional predictive content after controlling for effective shares.

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

respond to lower demand elasticities.[9] 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 builds on this evidence by showing that changes in customer relationships as firms grow can help explain the weak relationship between effective and conventional shares. The key distinction is between customer depth and reach. 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 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 40% lower than those of incumbent customers. The positive but small overall relationship between effective and conventional shares reflects the net effect of these opposing forces.[10]

Our fourth set of results shows that this distinction reshapes our understanding of recent trends in firm size and the rise of "superstar" firms. We first document that two prominent growth channels in recent years tilt toward reach rather than depth and therefore generate particularly muted changes in effective shares. First, 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. Second, growth in online sales expands firm reach but also exposes customers to a broader choice set. We find that customers who purchase only online have lower effective shares than customers who purchase in person at the same firm, and firms that gain market share primarily through online growth see much smaller increases in effective shares than firms whose growth is primarily in person.

Extending beyond these specific cross-sectional growth patterns, we then show that the distinction between effective and conventional market shares matters more broadly for interpret-

[9] 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).

[10] 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 these higher-intensity incumbent relationships receive more weight in the effective-share calculation, yielding a positive but small overall relationship between effective shares and market shares.

ing longer-run trends.[11] Over the last decade, superstar firms gained substantial aggregate market share but their share of their customers' spending grew much less. Interpreted through the lens of the variable-markup framework discussed above, this empirical pattern suggests that rising concentration has not led to a comparable rise in size-based product-market power.[12]

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 shopping patterns across firms, which can arise from a variety of factors such as geography, tastes and shopping technology. Income heterogeneity and customer dominance may be related in principle since income can 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 these differences, we share the conclusion that customer heterogeneity, rather than firm size alone, is central to understanding variation in markups 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 importantly by customer acquisition. Our results also complement Afrouzi, Drenik and Kim (2025), who show 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 proxies for demand elasticity 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, as intensified competition for customers raises demand elasticities.

Argente, M'endez and Van Patten (2026) constructs retailer-specific catchment areas and

[11]Most of our analysis uses the linked TransUnion sample, which gives us the highest quality cross-sectional coverage but is available beginning only in 2021. To study longer trends, we complement this sample with a historical, less granular version of the transaction data available back to 2014.

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

uses 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. Our focus differs in two respects. First, we measure customer dominance directly using individual customers' spending baskets rather than geographically defined catchment areas. Second, rather than studying markup responses in a particular policy setting, we use effective shares to reassess broad cross-sectional and time-series patterns in firm size and concentration.

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 retains, making them useful for studying customer dominance and size-based market power across many sectors, locations, and periods. More generally, 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 generally infeasible, effective shares provide a scalable, customer-based alternative to relying on 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 market-share-based variable-markup mechanism familiar from 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 through the more elastic within-category margin. With persistent customer heterogeneity, this same mechanism instead 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 size-dependent market power is determined by a firm's 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 $\eta$ and across categories with elasticity $\theta$, 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 $\omega_{ij}$ 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 $\eta$ and $\theta$, 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 there is no persistent heterogeneity in customers' relative demand for firms (e.g., $a_{ij} = a_j$ for all $i$), then[13]

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

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. 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:

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

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

Thus, aggregate market share is sufficient in this framework only if customers do not differ in their relative demand for firms. Our empirical analysis focuses on measuring the wedge between $\omega_j^{\text{eff}}$ and $\Omega_j$.

Although nested CES provides a particularly simple benchmark, the distinction between aggregate market share and customer-level dominance is not unique to this setting. In Appendix A2.2, we derive the same effective-share formula in a parsimonious nested-logit discrete-choice model building on Mongey and Waugh (2025), and discuss how related logic extends to broader demand environments. This theoretical mapping gives effective shares 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.[14] In Section 5, we provide empirical evidence consistent with this interpretation. However, our empirical contribution is to measure whether firms become more important within the spending baskets of individual customers as they grow in aggregate, rather than to structurally estimate elasticities from observed shares.

# 3 Data Description and Measurement Details

## 3.1 Data Description

Our primary data sample begins from the universe of all U.S. credit and debit card transactions of a major card-payment processor from 2018-2025. The scope of the data is large: it covers more than 600 million cards and more than 4 trillion dollars of spending each year. This corresponds to roughly one-fifth of aggregate annual personal consumption expenditures (PCE).

The unit of observation is an individual transaction between an (anonymized) credit card and a merchant. For each transaction, we observe the exact dollar amount, time and date, the unique identifier of the card that made the transaction, and the identity of 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 would be separate from Amazon). For in-person transactions, in addition to the identity of the firm, we observe the 5-digit zip code (and typically the exact address) of the store where the transaction was made, allowing us to identify purchases in different establishments of the same firm. To account for the fact that nearby establishments of the same firm are close sub-

[14]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.

stitutes, we aggregate all establishments with a firm at the ZIP3 and henceforth use this as our measure of "establishment".[15] For remote purchases (e.g. online, phone, recurring purchases), we observe the firm identifier but no physical location for the transaction. 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.[16]

Each transaction in the dataset is assigned a merchant category code (MCC). This 4-digit code captures the type of good or service that the firm provides. This code is assigned at the time that a merchant registers to accept card payments, and it determines the merchants' fee structure and corresponding credit card rewards, so is a fundamental characteristic of each transaction. We aggregate very similar MCC codes into a single MCC category and subset to the MCC codes that are customer-facing with good coverage of individual firms, leaving us with 33 MCC codes in our analysis sample.[17] Some merchants have multiple MCC codes corresponding to establishments providing different products (e.g. Costco wholesale vs. Costco gas stations). We assign all transactions within a firm to its MCC code with the most transactions. The largest categories in our sample are General Merchandise Stores (31% of our sample), Grocery Stores (18% of our sample), and Restaurants (9% of our sample) (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.[18] We therefore interpret consumer choice as occurring over retailers within a category. While this reflects a data limitation, it is also an economically relevant margin in many retail and service markets. To address potential concerns, we show that our results are robust in categories where firms and products are more closely aligned, such as Auto Repair, Health and Beauty Spas, and Restaurants (see Appendix Table A9).

The raw transactions dataset is a card-level panel, and has no links that aggregate multiple cards owned by the same customer. However, beginning in 2021, this card-level panel can be linked to TransUnion credit records, allowing us to group multiple cards belonging to the same person. The TransUnion linkage also provides a set of key consumer observables, including home location, age, and estimated income. The linked sample covers around 100 million

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

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

[17]We drop some categories like 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 A12 for a detailed list of categories.

[18]Neiman and Vavra (2023) document related but distinct trends at the UPC level for grocery-store products in Nielsen data. They also find growing divergence between household and aggregate concentration, although aggregate product concentration declines in their setting as new UPCs are introduced.

individuals and $3 trillion of spending and appears broadly representative of the adult U.S. population on a number of demographic characteristics. 41% of the spending in the full transaction dataset is included in the merged sample. Credit records do not capture debit-cards, so this is the main source of dropped spending.[19] Our conclusions are robust to using card-level instead of person-level spend to construct effective shares. Performing analysis at the card-level also lets us verify that patterns are very similar if we expand the sample to include the debit cards and credit cards that are not linked to TransUnion credit records.[20] In the linked sample, the average person has 1.6 credit cards per year and 2.2 cards over the sample from 2021-2025, with 13.76% of people having at least 3 cards within a year.[21]

Starting from the linked sample, we impose additional 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. This restriction excludes idiosyncratic travel spending (e.g., vacation purchases) and focuses on the set of stores plausibly relevant for regular shopping behavior.[22] 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.[23] Fourth, we drop all customers whose home address in TransUnion changes 3-digit zip codes within the year. Fifth, we focus our attention on category-zip3 pairs where we observe at least $450K in spending between 2021-2025. Sixth, to limit potential sampling error in measured effective shares, we focus only on firms with at least 200 customers observed in our linked subsample in a year. We henceforth refer to this filtered sample as our "Primary" sample.

Figure 1 shows that the credit card spending we observe aligns well with firm spending data from the Census, suggesting that our sample does a good job of capturing the true distribution of spending across firms. In particular, this figure shows that the national sales share of the top 4 firms (CR4) in each category in our card data aligns well with the same measure for similar

[19] 76% of cards that are not linked are debit cards; the remaining unlinked cards arise because not all credit-issuers report to TransUnion.

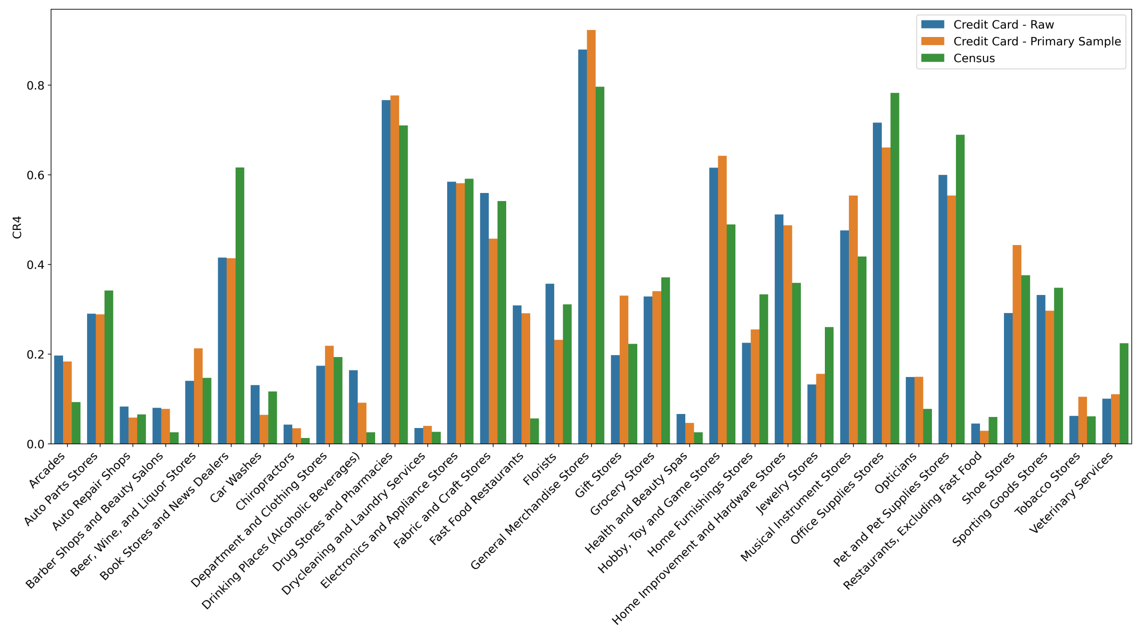
[20] Specifically, we find that the relationship between firm-level market shares and effective shares calculated using all cards and that using just credit cards are very similar. See Appendix Table A9 for the details.

[21] 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.

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

[23] We are able to include transactions made through Apple and Google Pay, as that affects the mode of transaction but not the firm identifier. In our final subsample of 33 categories, 5.4% of all transactions are at payment facilitators, a number that rose from around 3% in 2018 to 8% in 2022. Payment facilitators are most prevalent in barber shops and restaurants. Excluding these transactions will bias our results only to the extent that these transactions are unevenly distributed across firms or locations.

Figure 1: Benchmarking Credit Card Data to Census Data




Notes: This figure plots the 2022 four-firm concentration ratio (CR4) by category in three data sources: the raw credit card transactions sample, the primary credit card sample merged with TransUnion, and U.S. Census data. For each category, CR4 is defined as the sum of the market shares of the four largest firms. The primary credit card sample is restricted to categories with at least 20 distinct firms in each year. To harmonize category definitions across sources, category labels are standardized prior to merging. Our sample is further restricted to categories observed in all three datasets with non-missing CR4 values. Used Merchandise Stores, Gyms and Fitness Clubs, and Gas Stations are excluded. Census NAICS code in 2022 are manually mapped and verified to our credit card category definitions.

categories in the Census.[24] This is true both in the more comprehensive raw card data as well as in our more restricted primary sample. The category with the most notable deviation is "Fast-Food Restaurants", and this is because the Census treats franchises of the same chain as separate firms while our data combines all franchises in the same firm. We think our measure is likely more relevant for understanding customer demand and firm product market power, since consumer shopping choices as well as pricing and product assortment are often made at the firm rather than 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.[25] The main benefit of this histor-

[24]The crosswalk we use to map NAICS codes used by Census to MCC categories in credit card data are shown in Appendix Table A12. It is worth noting that NAICS and MCC-based categorizations of firms will not perfectly align, so we should not expect a perfect match even if spending for every individual firm exactly matches.

[25]A version of this data goes back to 2008 but its coverage is extremely sparse before 2014, so we begin analysis in 2014.

ical panel is the longer time coverage, since our primary sample starts in 2021. However, this historical panel has two important limitations. First, it cannot be linked to TransUnion, and therefore for this analysis, our unit of observation is the card, not the person. Second, this historical panel measures total spending by individual cards in each zip code, but it only further dis-aggregates this spending within zip codes for a subset of (relatively large) "named firms". Just like in our primary data set, the spending that cards do at these named firms is separately identified, but the remainder of spending is grouped together as spending at "Other-un-named firms". Although we can see a card's *total* spending at these unnamed firms we are unable to measure its spending at any specific unnamed firm within this group. This means that the data is well-suited for studying spending trends for larger, named firms but not for studying spending trends at the smallest firms in the economy.[26] We restrict our analysis to only categories with at least 10 named firms in every year 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 named firms while we only need to reliably capture the top 4 firms for our analysis. Despite these limitations, we confirm that the key relationships between effective shares and market shares are similar in our primary and historical panel in the years where they overlap, which gives us confidence that the historical panel captures the same underlying patterns as our primary data. We discuss this benchmarking more in Section 6.

## 3.2 Measuring Effective Shares

Using this analysis sample, we now construct the empirical analogues of the spending shares defined in Section 2. That static definition has no notion of time. In our baseline measure we take it to the data by constructing each customer's spending shares 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

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

[26] Importantly, 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* spending at all other firms but not how that other spending is split across specific firms.

Firm $j$'s effective share is then

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

Thus $\omega_{ij}$ measures the firm's importance within a customer's spending basket, while $\alpha_{ji}$ weights customers by their importance to the firm's sales. Because we build effective shares from customers' own spending baskets within an MCC category, they have the advantage of imposing no geographic market boundaries on firms. We need not take a stand on whether the relevant market is a ZIP code, an MSA, or the nation as a whole. The geographic scope of each firm is instead revealed by where its customers actually shop. Incorporating e-commerce 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 handle online spending.

Although transaction data let us observe spending directly, we see only a finite number of purchases. This means that measured spending shares might reflect random noise rather than persistent customer–firm relationships or the full set of options a customer considers. Concretely, at the level of a single purchase, realized shares are necessarily zero or one regardless of 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 further attenuate these idiosyncratic realizations, we restrict attention in each category to customers who make at least 5 transactions in that category during the year.[27] The main concern with this filter is the potential for selection, since it drops the least frequent shoppers. However, it is important to note that customers with few transactions tend to also have low $\alpha_{ji}$, and thus tend to have a small influence on effective shares even if included. All of our conclusions are also similar if we focus on only categories where this restriction rarely binds.[28]

We also construct two alternative measures which pool transactions more extensively. This simultaneously reduces concerns about measurement error and about selection induced by the minimum-transaction filter. However, each alternative approach comes with its own trade-offs. The first approach pools each individual's transactions over the full five-year period from 2021 to 2025. This means that we observe many more transactions for every person, but it assumes that customer-firm relationships are stable over a long period and precludes annual

[27]Simulations suggest that with a 5+ transaction filter, we can reliably distinguish true variation in $\omega_j^{\text{eff}}$ from variation generated by random discrete realizations. We also discuss a specific small-sample bias correction.

[28]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 0.4 percent of spending in General Merchandise Stores but 80 percent of spending in Jewelry Stores.

panel analysis. The second approach constructs ”synthetic customers” by pooling transactions across demographically similar individuals living in the same location—specifically, individuals in the same census block with similar age, income, and (inferred) car-ownership status. These synthetic customers have many more transactions, but the cost of this approach is that it averages over some genuine heterogeneity in customer-firm relationships and it precludes individual level analysis. In practice, these measures are all very similar: the firm-level effective shares calculated using 5-year pooled samples or using synthetic customers have a correlation with our baseline measure of 0.96 and 0.94, respectively. Table A1 provides summary statistics for these three pooled samples as well as for an even larger sample that measures spending at the card level without requiring TransUnion links. We show below that our main results are robust across these alternative constructions.

# 4 The Difference between Effective Shares and Market Shares

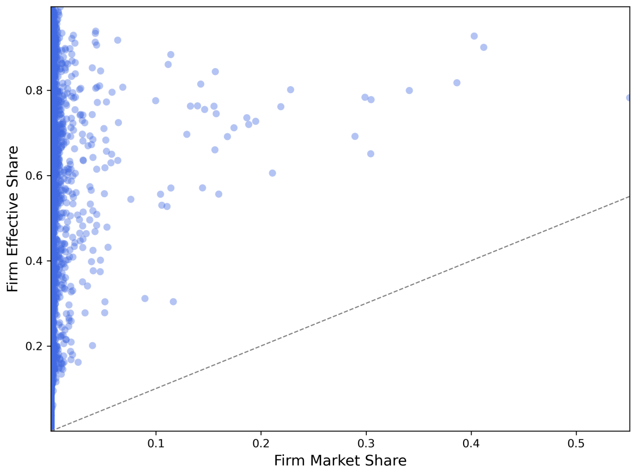
Our first empirical finding is that differences in conventional market shares substantially overstate how much firms differ in their importance to customers: large differences in conventional market shares correspond to much smaller differences in effective shares. This compression is apparent whether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

The left panel of Figure 2 shows the relationship between firms' conventional national market shares and effective shares in 2025. The x-axis shows each firm's conventional national market share, while the y-axis shows its effective share. All points lie mechanically above the dashed 45-degree line, since a firm must account for at least as large a share of category spending among its own customers as it does among all customers nationally. More notably, the gap is large for most firms, and the relationship between conventional and effective shares is weak. There is also substantial heterogeneity in effective shares among firms with similar national market shares.

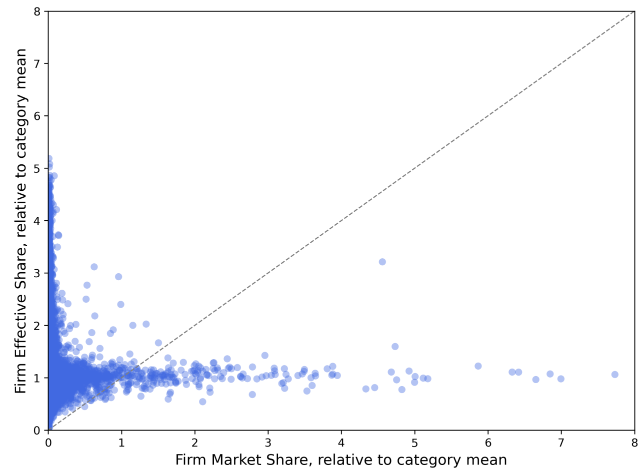
The left panel of Figure 2 pools firms across all categories. However, categories differ substantially in their market-share distributions, and this cross-category variation may obscure within-category patterns that are of greater interest.[29] In the right panel, we therefore divide each variable by its respective sales-weighted category average. The x-axis thus shows each firm's conventional market share relative to the category average, and the y-axis shows its effective share relative to the category average. After this normalization, the relationship is even

[29]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 2: Effective shares differ from national market shares




(a) Overall




(b) Within Category

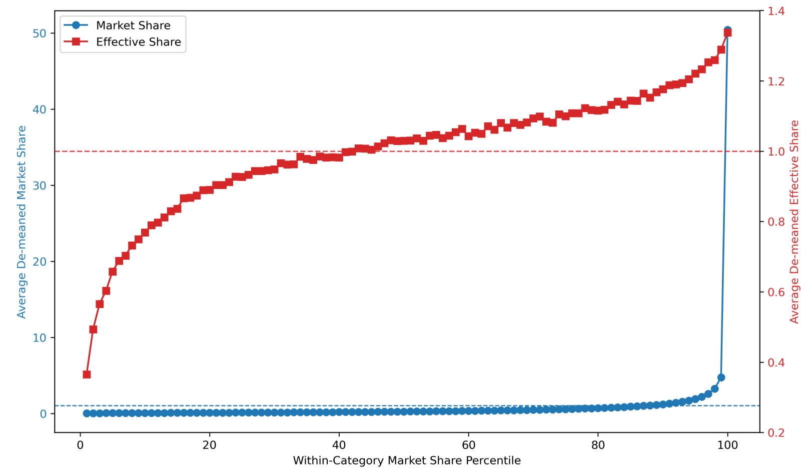
Notes: In each panel, each dot reflects 1 firm in 2025. In panel (b), each value is divided by the sales-weighted average value in the category. Effective shares are defined using annual customer observations. In the right panel, both the x and y axis are truncated at 8.

weaker: firms that are large outliers in conventional market share relative to others in their category are generally not large outliers in effective share.

Since the distribution of firm size is highly skewed, most firms have small national market shares so there is a large mass clustered near zero in both panels of Figure 2. This makes it difficult to see potential patterns among most firms in the raw scatter plots. Figure 3 provides an alternative visualization of how conventional and effective shares vary across the entire distribution. Specifically, we rank firms by their market shares within their categories and plot the average market share relative to the category average (left axis) and the average effective share relative to the category average (right axis) for firms in each percentile. The market-share series is weakly increasing by construction, since the percentiles are defined using market shares. The relevant feature is the shape of the distributions. 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 across the distribution. The largest firms, as defined by conventional market share, capture a larger share of their customers' spending, but their effective shares are only about 30 percent above the category average.

A natural explanation for the gap between national market shares and effective shares is

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



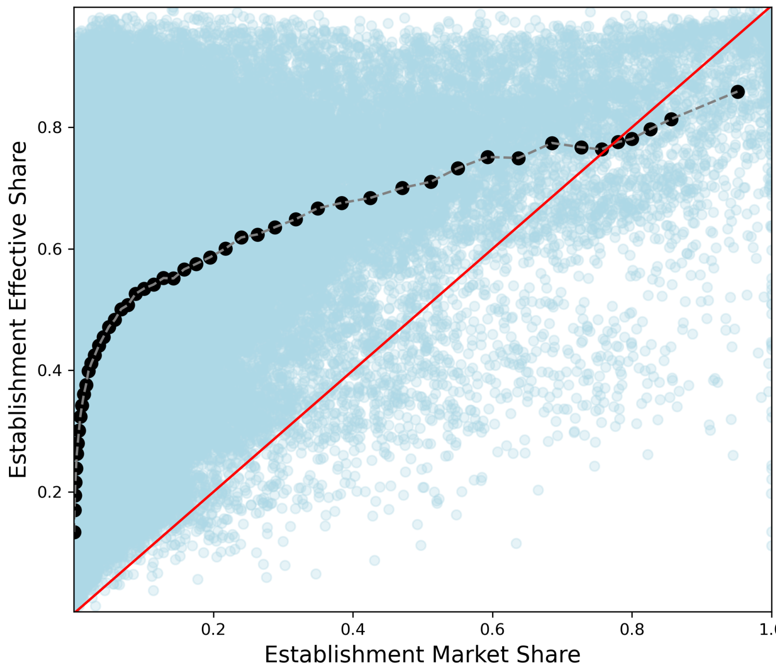

Notes: The x-axis reports within-category market share percentiles, constructed by ranking firms within each category by their market share and assigning them to up to 100 percentile groups. Percentile 1 corresponds to the lowest market share within a category, and percentile 100 to the highest. The left vertical axis (blue) plots the average ratio of market share within each percentile group to the category average market share, meaning that a value of 1 indicates being at the category average. The right vertical axis (red) plots the average effective share within each percentile group, normalized by the category mean effective share, so that values greater than 1 indicate above-average effective share within a category. Horizontal dashed lines at 1 show the category mean. The figure aggregates observations across categories after within-category normalization, and all firms receive equal weight within percentile groups.

geographic aggregation: many consumer markets are local rather than national. In sectors such as restaurants and grocery stores, national market shares may therefore be the wrong benchmark, because they average over many distinct local markets. We test this by moving to the local level and asking whether effective shares align more closely with local market shares. If the divergence is mainly a market-definition issue, then effective shares should line up much more closely with conventional shares once they are computed in the local markets where firms actually compete.

To test this, we move to the local level, where an observation is an establishment (a firm–ZIP3 pair, as defined in Section 3.1). For each establishment $k$ of firm $j$ in period $t$, we define the local market share as the sales at that establishment divided by the total category sales in the ZIP3. Constructing establishment effective shares requires no new customer-level object. This is because we treat all of a firm's establishments and its online channel as substitutes for a given customer, which means that there is no customer-establishment spending share $\omega_{ijk}$, only the customer-firm share $\omega_{ij}$ defined in Section 3.2.[30] We therefore construct establishment

[30]Results are robust to instead treating each establishment as a distinct entity in customers' baskets, i.e., defining

Figure 4: Effective shares differ from local market shares




Notes: Each blue dot is an establishment (firm-zip3 combination). The black dots show binscatter relationships with 50 bins, where the bins are defined weighting establishments by their total sales. For each bin, we plot the weighted mean effective share against the weighted mean market share, where weights are given by establishment sales.

$k$'s effective share by re-weighting these customer-firm shares by each customer's share of the establishment's sales.[31] 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 resulting estimates. The y-axis shows the effective share of each establishment, while the x-axis shows the establishment's share of total category sales in its ZIP3. Each blue dot is an establishment, while the black series traces binned averages, with each bin representing the same amount of total sales. The red 45-degree line shows the benchmark along which the two measures would coincide.

There are several notable patterns. First, the overall levels of effective shares and market

$\omega_{ijk}$ as establishment $k$'s own share of customer $i$'s category spending and aggregating with establishment-level sales weights.

[31] Spending at online-only firms is included in customers' total category spending and therefore reduces $\omega_{ij}$ at brick-and-mortar firms. However, online-only firms have no physical location over which to compute conventional local market shares and so are not included as observations in the establishment-level analysis.

shares are much more comparable than in the national comparison, primarily because local market shares are much higher than national market shares. Nevertheless, the individual relationship remains weak: establishments with similar market shares often differ substantially in their effective shares. More importantly, the compression documented at the national level remains clearly visible at the local level. This compression appears as a strongly concave relationship between the two measures. Establishments with small local market shares have effective shares far above their market shares – those establishments are capturing a much larger share of their customers' spending than of the local market as a whole. At the other extreme, establishments with very large local market shares have effective shares that are somewhat *below* their market shares. A store can appear dominant according to local market shares while still accounting for a more moderate share of its customers' total category spending once one accounts for purchases made outside the local market (either in-person or online).[32]

One possibility is that the divergence at the ZIP3 level reflects this particular choice of local geography, but Appendix Figure A1 shows that this tension remains under alternative geographic definitions. When we define markets using 5-digit ZIP codes, which are substantially smaller than ZIP3s, conventional market shares exceed effective shares for most establishments, suggesting that ZIP5 markets are typically too narrow. We also consider larger markets defined at the state level. Finally, we implement a more flexible, 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 specification that lets geographic scale vary by category, substantial divergence remains between the two measures.

More fundamentally, the fact that establishments draw different sets of customers limits the usefulness of any common geographic market definition. Establishments differ in the geographic scope of their customers' shopping patterns, either because they draw customers with different travel patterns (e.g., some stores have parking and others don't) or because they have different online footprints. Even for establishments in the same location and category, a common geographic partition may therefore be too broad for some and too narrow for others. This does not rule out selecting a geographic definition that closely aligns the two measures for a particular establishment or subset of establishments. But it suggests that no common partition is likely to work well for all establishments.

To better understand the wedge between effective shares and market shares, we decompose it into two forces. Figure 5 illustrates this decomposition. 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 this within-region effective share.

[32] In Appendix Figure A2, we show that this same pattern holds under the alternative pooling choices discussed in Section 3.2.

It necessarily lies above the red 45-degree line, which represents equality with the local market share, 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 it does 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 important throughout the distribution, but is especially pronounced for small and mid-sized establishments. Even when these establishments account for only a small share of local sales, they capture a substantial share of the within-region spending of the customers they serve.

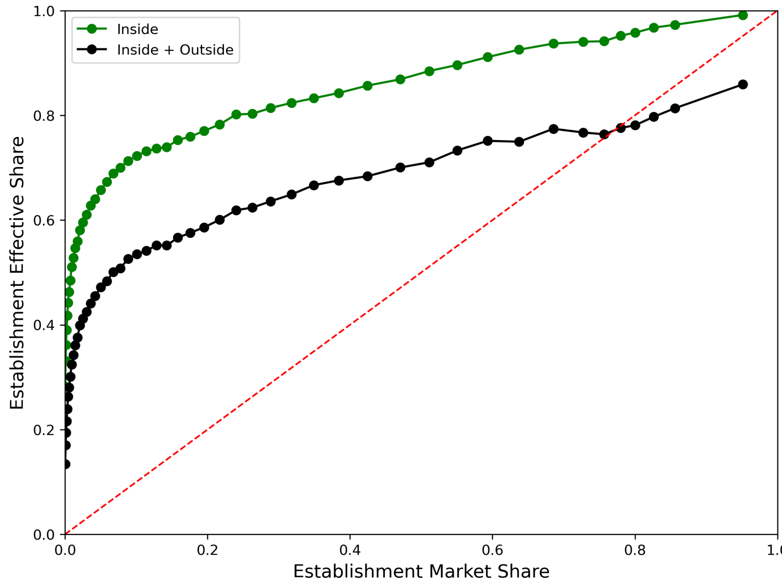
Within-market segmentation is partly offset by spending that customers do outside the ZIP3, either online or in person. The black series in Figure 5 reincorporates this out-of-region spending and therefore reproduces the black series from Figure 4. Relative to the green series, the curve shifts downward almost uniformly. The direction of the average shift is unsurprising: when customers shop outside their ZIP3, they often do so at a different set of firms than when they shop within the ZIP3. More notable is that the shift is similar across the local-market-share distribution.[33] This means that out-of-region spending reduces effective shares broadly, while the pronounced concavity in the relationship between effective and local market shares arises primarily from within-market segmentation.

The concave relationship between effective shares and market shares in levels suggests that the relationship may be closer to linear in logs. The left panel of Figure 6 shows that this is the case. It plots binscatters of log of effective shares against log market shares in 2025. The blue circles show the relationship between effective shares and national market shares for firms (as in Figure 2 above) and the orange squares show the relationship between effective shares and local market shares for establishments (as in Figure 4). In both cases, the relationship is close to linear but is substantially attenuated relative to the dashed 45-degree line. The slope is 0.16 nationally and 0.17 locally. Thus, a 10 percent higher market share is associated with only about a 1.6 percent higher effective share nationally and a 1.7 percent higher effective share locally. The right panel shows that this linear relationship is also present when looking at changes within firms and establishments between 2021-2025. The corresponding slopes remain far below one: a 10 percent increase in market share is associated with only a 1.8 percent increase in effective share when looking at national relationships and a 2.1 percent increase when looking at local relationships. Appendix Table A9 shows that this estimate is very similar when weighting firms by their sales and is slightly smaller in services and larger in retail.[34]

[33] For single-market firms with no online presence, reincorporating out-of-region spending can only weakly reduce effective shares, since the firm captures none of its customers' spending outside the region. For multi-region firms or firms that sell online, the effect is ambiguous in principle, because customers may continue shopping at the same firm in other regions or online.

[34] In Appendix Figure A3, we show the relationship between log establishment effective shares and the log of

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.

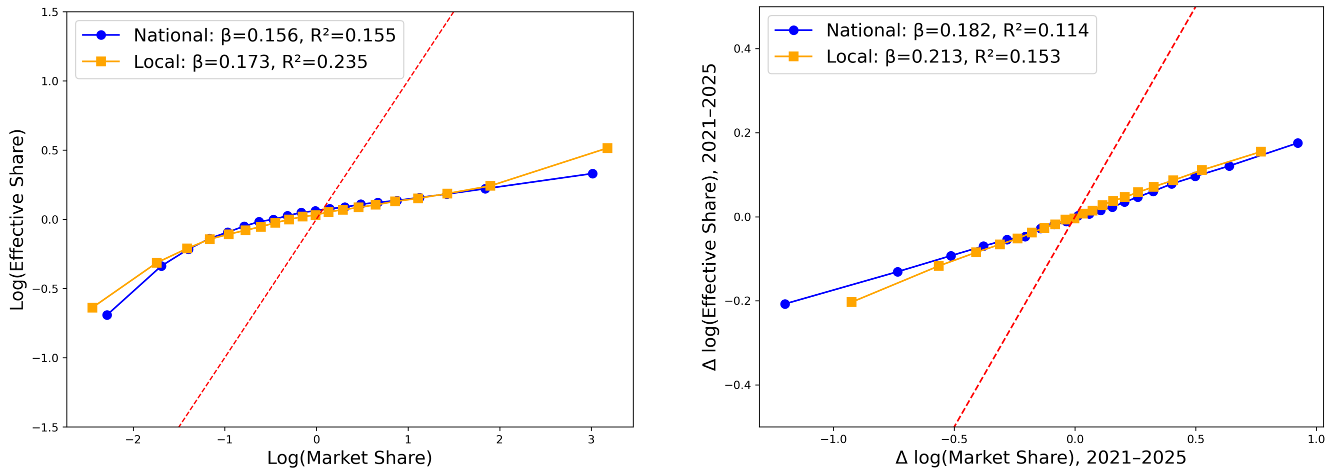
The results above show that conventional market shares are a poor summary of firms' importance to their customers. This raises two related questions. First, are high effective shares primarily explained by observable differences in the types of customers firms serve? Appendix Table A9 shows that controlling 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.[35] 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 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). This measure requires only total firm sales and customer counts, rather than information on the distribution of spending across customer-firm pairs. It is naturally related to effective shares because a customer shifting category spending towards a

establishment market shares in each of our categories. The attenuated relationship is pervasive, ranging from 0.25 in General Merchandise Stores to around 0.05 in Arcades or Office Supply Stores.

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

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




(a) 2025 Levels

(b) Changes from 2021-2025

Notes: This figure presents binscatter relationships between market share and effective share estimated using national and establishment data. The left panel plots the relationship in levels in 2025, while the right panel plots the relationship in changes between 2021 and 2025. In both panels, the horizontal axis reports log market share and the vertical axis reports log effective share. Circles correspond to the national sample, while squares correspond to the establishment sample. In the left panel (levels), log market share and log effective share are demeaned using fixed effects: category fixed effects in the national sample and category × ZIP3 fixed effects in the local sample. The plotted points represent averages within 20 bins of the residualized log market share. In the right panel, extreme changes in log market share are trimmed symmetrically at the top and bottom 2.5% of the distribution. As in the left panel, changes are residualized with respect to category fixed effects in the national sample and category × ZIP3 fixed effects in the local sample. The binscatter plots averages within 20 bins of residualized changes in log market share. In both panels, the coefficients and $R^2$ of the regressions on the underlying data are reported in the legend. The dashed 45-degree line is shown for reference.

firm will raise both its effective share and its sales per customer. At the same time, the two measures are distinct: sales per customer can increase without a corresponding increase in effective share if the increase in sales per customer comes from serving customers with larger category spending rather than from capturing a larger share of those customers' spending. In addition, because effective share is an average of customer-level spending shares weighted by their share of the firm's total sales, it depends not only on average sales per customer but also on how the firm's sales are distributed across specific customers.

Appendix Table A5 shows regression estimates that summarize the relationship between the log of effective share and the log of sales per customer or log number of customers. Unsurprisingly, we find that effective shares indeed co-vary strongly with sales per customer within category: a 1 percent increase in sales per customer is associated with a 0.55 percent increase in effective shares. However, despite this strong co-movement, there is still substantial variation in a firm's effective share that is unrelated to the firm's average sales per customer, with partial $R^2$ from these regressions hovering around 0.5 depending on the specification. Therefore, we conclude that sales per customer is a useful although 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 continues to refer to a firm-by-ZIP3 pair rather than a single physical address. We focus on this local analysis using establishments rather than national firm-level analysis because Section 4 showed that local market shares are more closely related to effective shares than national market shares are. However, similar results hold when we instead aggregate to the firm level. For expositional simplicity, we use the term firm throughout this section, but the baseline regression is estimated on customer-establishment observations.[36] Among customers who shopped at firm $j$ in year $t-1$, we define retention as an indicator for whether customer $i$ shops at firm $j$ again in year $t$:

$$\text{Retention}_{ijt} = \mathbb{1}\left\{i \text{ shops at } j \text{ in } t\right\}.$$

We then estimate the relationship between retention and the customer-level spending share that enters firm effective-share calculations:

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

Here $\omega_{i,j,t-1}$ is firm $j$'s share of customer $i$'s category spending in year $t-1$, and $\Omega_{j,t-1}$

[36] In the baseline retention analysis, customer effective shares are defined at the firm level, while conventional market shares are defined at the establishment level. Thus, a customer who shops at multiple establishments of the same firm enters the sample separately for each establishment, with the same firm-level customer effective share but with establishment-specific conventional market shares. Online transactions are excluded before constructing these measures, so online-only firms are not included in this analysis.

is firm $j$'s conventional market share in year $t-1$. To simplify exposition, we refer to the customer-level $\omega_{i,j,t-1}$ in these specifications as customer effective shares, since firm-level effective shares are constructed by aggregating these customer shares across the firm's customers. We use the binary retention indicator as our outcome because it directly measures whether a customer-firm relationship continues.[37] We standardize both $\log \omega_{i,j,t-1}$ and $\log \Omega_{j,t-1}$ within category so that coefficients $\beta_1$ and $\beta_2$ are directly comparable and correspond to the effect of a one-standard-deviation increase in the respective variable.[38] We estimate this specification using data from 2021–2025. 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.

This exercise 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, then they should have little predictive power for future behavior. Second, even if customers shopped completely at random, retention might be higher at firms with larger conventional market shares simply because customers are more likely to encounter and purchase from those firms again, not because those firms are especially important to the specific customers they serve. The key question is therefore whether customer effective shares predict retention after controlling for the firm's aggregate market share. If they do, this indicates that customer effective shares contain information about persistent customer-firm relationships, rather than simply reflecting firm-level scale.

Table 1 reports the results. Column (1) relates customer retention to effective shares alone. Customers with higher effective shares are substantially more likely to continue shopping at the firm in the following year. The magnitude is economically large: a one-standard-deviation increase in a customer's effective share is associated with an approximately 18 percentage point higher probability of retention. Column (2) shows that conventional market shares are also positively related to retention, consistent with the fact that customers are more likely to return to firms with larger overall market presence. However, Column (3) includes both variables simultaneously and shows that the predictive content of conventional market shares disappears once customer effective shares are included. The coefficient on the conventional market share

[37] A natural alternative is to use $\omega_{i,j,t}$ 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 A7 shows that results are similar if we instead study the persistence of $\omega_{i,j,t}$, or if we measure retention over longer horizons.

[38] 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) |
|---|---|---|---|---|---|
| $\log \omega_{i,j,-1}$ | 0.179*** | | 0.180*** | 0.178*** | 0.184*** |
| | (0.002) | | (0.002) | (0.002) | (0.004) |
| $\log \Omega_{j,-1}$ | | 0.013** | -0.014*** | | |
| | | (0.005) | (0.005) | | |
| Market FE | X | X | X | X | X |
| Establishment-by-Year FE | | | | X | X |
| Drop in $\log \Omega$ | | | | | X |
| Observations | 8,291,837 | 8,291,837 | 8,291,837 | 8,291,649 | 1,249,007 |
| $R^2$ | 0.244 | 0.172 | 0.244 | 0.301 | 0.326 |
| Within $R^2$ | 0.087 | 0.000 | 0.087 | 0.084 | 0.086 |

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. 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. All standard errors are clustered at the category-by-zip3-by-year level.

falls sharply, while the fit of the regression is essentially unchanged: the within-$R^2$ is 0.087 with effective shares alone and remains 0.087 after adding market shares.

Column (4) includes establishment-by-year fixed effects, identifying the relationship by comparing customers who shopped at the same firm in the same year. This specification asks whether, among customers of the same firm, those who devote a larger share of spending to the firm are more likely to continue shopping there in the future. The relationship remains strong and positive, indicating that variation in customer effective shares even within a firm captures economically meaningful variation in customer attachment.

While this analysis focuses on customer-level retention, Appendix Table A6 shows that similar patterns emerge when we aggregate to the establishment (firm-by-ZIP3) or firm level.[39] Establishments and firms with higher effective shares retain a larger fraction of their customers

[39] Appendix Table A6 also compares effective shares to two related firm-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, while effective shares measure customer dominance more directly.

in future periods, and effective shares consistently outperform conventional market shares in predicting retention. We also obtain similar results when effective shares are constructed using synthetic customers as described in Section 3. This robustness suggests that the retention results are not driven solely by the particular way we aggregate purchases at the individual level.

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 little reason to reconsider, even if they would switch in response to a small change in prices, quality, or available alternatives. Column (5) therefore restricts attention to firms experiencing large declines in market share, where many customer-relationships are being disrupted. This specification continues to include firm-by-year fixed effects, so the coefficient is identified by comparing customers of the same declining firm 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 firms 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 customer–firm relationships persist, and they do so both across firms and among customers of the same firm. This establishes that effective shares capture meaningful 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 Firm 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 behavior when customers face a new alternative. The theoretical framework in Section 2 provides one interpretation of this empirical result: demand from a customer is less elastic when a firm accounts for a larger share of that customer's category spending. This means that customers with higher effective shares at an incumbent firm should be less likely to reallocate spending to a new entrant.

We examine this prediction using local firm entry events.[40] Importantly, local entry is endogenous. For this reason, all specifications include entry-event fixed effects and compare customer-incumbent observations exposed to the same local entry event, rather than compar-

[40]As in the retention analysis, customer effective shares are defined using a customer's firm-wide spending, while conventional market shares and entry events are measured locally. Thus, throughout this subsection we use firm language for exposition, but the entry event and conventional market share variables are local measures.

ing across markets with different endogenous entry decisions. For each entry event, we identify customers who previously shopped at incumbent firms and ask whether those with higher effective shares at an incumbent are less likely to switch to the entrant after it appears.[41]

Let $e$ index an entry event, defined by an entering firm, location, and entry year. We index time relative to the entry event, with $0$ denoting the entry year and $-1$ denoting the year before entry. We then estimate

$$Switch_{ije,0} = \beta_1 \log \omega_{ij,e,-1} + \beta_2 \log \Omega_{j,e,-1} + \gamma_e + \epsilon_{ij,e,0}.$$

Here $Switch_{ij,e,0}$ is an indicator for whether customer $i$, who purchased from incumbent firm $j$ in the year before entry event $e$, both stops purchasing from $j$ and purchases from the entrant in the entry year.[42] The variable $\omega_{ij,e,-1}$ is customer $i$'s effective share at incumbent firm $j$ in the year before entry, measured using the customer's firm-wide spending. The variable $\Omega_{j,e,-1}$ is incumbent $j$'s local conventional market share in that same pre-entry year. The entry-event fixed effect $\gamma_e$ implies that the regression compares customer-incumbent observations exposed to the same entry event, rather than comparing across markets with different endogenous entry decisions.

Our estimates of $\beta_1$ and $\beta_2$ therefore reveal whether switching to a newly entering firm is more closely related to the customer's effective share at the incumbent or to the incumbent's overall share of local sales. We again standardize regressors so that coefficients can be directly compared and interpreted as one standard-deviation effects. For each entry event, we define the exposed customer set as customers living in the 5-digit ZIP code that accounts for the largest number of customers visiting the entrant in the entry year.[43]

Table 2 shows that customers are less likely to switch to a newly entering firm when the incumbent previously accounted for a larger share of their category spending. This relationship is driven by customer effective shares rather than by the incumbent's conventional market share: and increase in $\omega_{ij,e,-1}$ strongly predicts a decline in switching, while $\Omega_{j,e,-1}$ has little additional predictive power once effective shares are included. Column (4) adds incumbent-by-entry event fixed effects and shows that effective shares continue to strongly predict a decline in switching. With these fixed effects, identification comes only from comparing customers

[41]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). Further details on exactly how we define entry can be found in Appendix A1.2.

[42]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.

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

Table 2: Effective Shares Predict Less Switching to Entering Firms

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $\log \omega_{j,-1}$ | -0.028*** | | -0.027*** | -0.029*** | -0.031*** |
| | (0.001) | | (0.001) | (0.001) | (0.001) |
| $\log \Omega_{j,-1}$ | | -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 | 34,345,857 | 34,345,857 | 34,345,857 | 33,681,551 | 6,310,933 |
| $R^2$ | 0.061 | 0.051 | 0.061 | 0.123 | 0.135 |
| Within $R^2$ | 0.012 | 0.001 | 0.012 | 0.011 | 0.011 |

Notes: This table shows the relationship between effective shares and switching to an entering establishment, where switching is an indicator for leaving the incumbent establishment in $t$ and shopping at the entering establishment. We use a 10% sample of all identified firm-entry events. This sample corresponds to around 1,000 entry events. The sample is customer-incumbent-entry-event observations, where the customer shopped at the incumbent establishment in $t-1$ and remains active in the category in $t$. Column (5) further restricts the sample to incumbent establishments with declines in conventional market share from $t-1$ to $t$ where the decline exceeds the median decline in log market shares among establishments with falling market shares. All standard errors are clustered by entry event.

who shopped at the same incumbent prior to the entry event but did so with different intensity. Since the fixed effects absorb any incumbent-level exposure to the entrant, this also addresses a concern that some incumbents may be closer substitutes for the entrant than others. 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 that are especially exposed to the entrant's arrival.

Appendix A1.2 discusses various robustness exercises. In all cases, the patterns are consistent with the interpretation that effective shares capture economically 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 attenuated 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 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 firms' sales growth between 2021 and 2025.[44] The attachment of newly acquired customers is therefore central to understanding how firm growth translates into effective-share growth.

We begin by comparing effective shares among continuing and 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 in at least one earlier sample year. We then compute the firm's effective share separately within each group, using customer sales weights that sum to one within the group.[45] 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, implying that newly acquired customers allocate a smaller share of category spending to the firm than continuing customers do. The magnitude is economically meaningful: at the median firm, the effective share of new customers is 39 percent lower than that of continuing customers.[46] This pattern is also pervasive across categories (see Appendix Table A11).

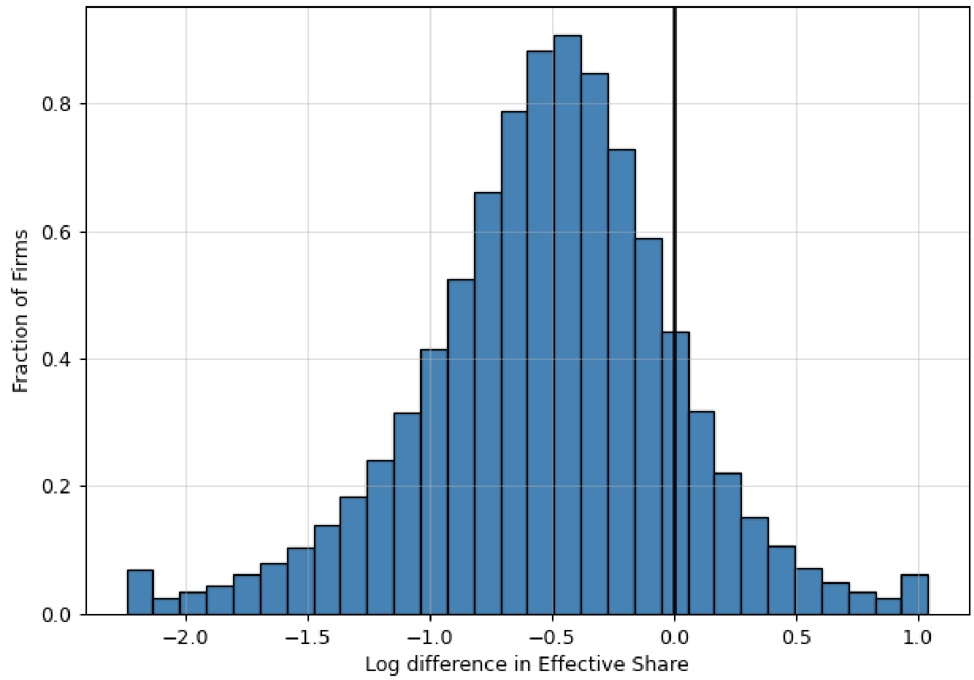
The new-versus-continuing comparison summarizes prior customer-firm exposure with a single indicator: whether a 2025 customer purchased from the firm in any earlier sample year. We next unpack the role of relationship tenure by asking whether effective shares vary systematically with the number of prior years in which a customer shopped at the firm. For each firm, we group 2025 customers according to the number of prior years, from 2021 through 2024, in which they shopped at the firm. We then compare each group's effective share to the effective share 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 smoothly with relationship length: customers new in 2025 have effective shares about 62 percent below the most-tenured group, and this deficit shrinks to roughly 50, 42, and 31 percent for customers

[44]Specifically, the coefficient on $\Delta\log(\text{sales})$ from a regression on $\Delta\log(\text{sales})$ on $\Delta\log(\text{no. customers})$ and category fixed effects is 0.78.

[45]That is, the effective share for each group is computed using the relative sales weights of customers in that group.

[46]$0.39 = (1 - e^{-.494})$. If we impose a stricter requirement that a continuing customer must purchase in every year from 2021 through 2025, rather than in at least one year, then the gap rises to 64% at the median firm.

Figure 7: Effective Shares for New versus Continuing Customers



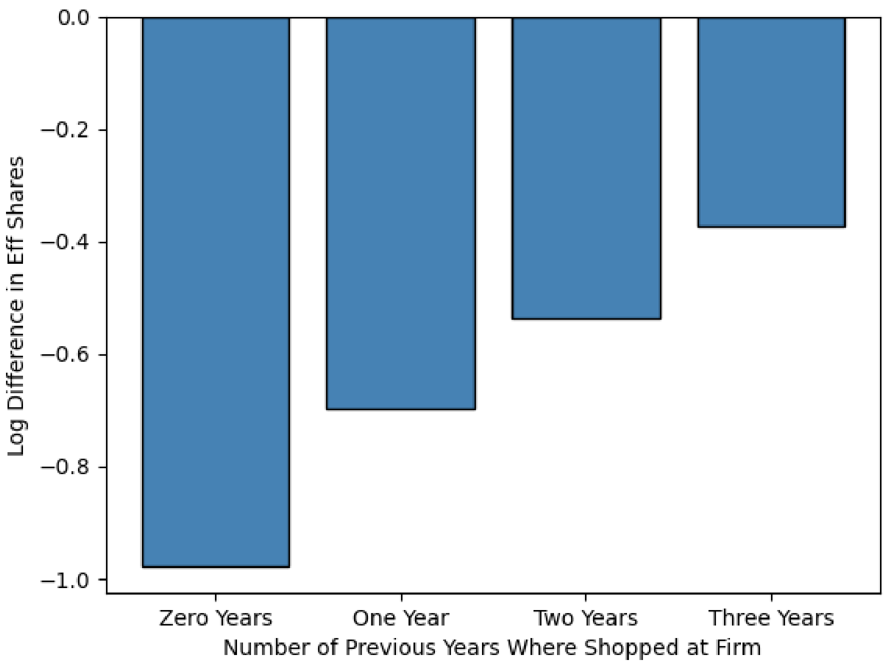

Notes: This figure shows the equally weighted distribution across firms of the log difference between the effective share of new customers and the effective share of continuing customers in 2025. Specifically, the plotted variable is $d_j \equiv \log(\omega^{eff}_{j,\text{new}}) - \log(\omega^{eff}_{j,\text{continuing}})$, so negative values indicate that newly acquired customers have lower effective shares than continuing customers. Only firms with at least 20 new and continuing customers are included, and the distribution is winsorized at the top and bottom 2.5 percent. The vertical line indicates $d_j = 0$, where the two group-specific effective shares are equal.

with one, two, and three prior years at the firm.[47] Thus, the new-customer gap reflects a broader tenure gradient: customers with a longer history at the firm tend to direct more of their category spending there.

The tenure gradient reflects both changes within customer-firm relationships as they age and selection into which relationships persist. 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. As we move to longer horizons since first purchase, the customers who continue purchasing from the firm through that horizon direct more of their category spending to that firm. This pattern is not mechanical: a customer could purchase from a firm every year while still doing little of their category spending there. The positive gradient therefore indicates that persistence on the extensive margin is

[47]The corresponding median log differences are $-0.98$, $-0.70$, $-0.54$, and $-0.37$.

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



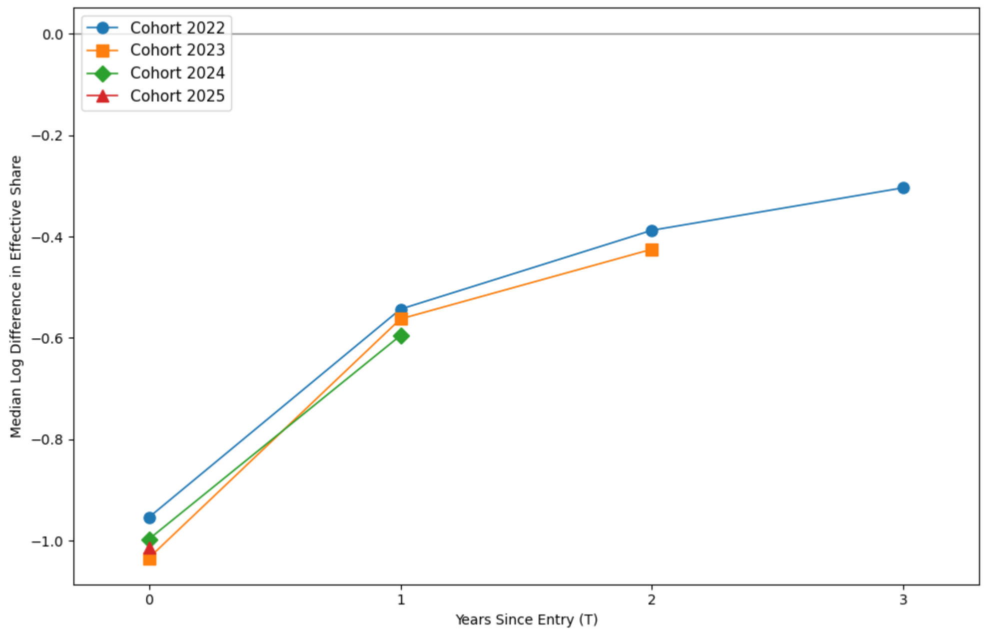

Notes: This figure groups 2025 customers by the number of prior sample years, from 2021 through 2024, in which they shopped at the firm. For each firm and each group $k \in \{0, 1, 2, 3\}$, we compute the log difference between the effective share of customers with $k$ prior years at the firm and the 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. The figure plots the median of this firm-level log difference across firms for each value of $k$. Effective shares are computed separately within each group using customer sales weights that sum to one within the group.

closely linked to depth on the intensive margin. Notably, the gap does not close within our sample window: even three years after first purchase, the earliest cohort we observe remains meaningfully less attached than customers who purchase from the firm in every sample year.[48]

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

[48]Appendix Figure A5 reports a complementary exercise that conditions on survival through 2025, holding the set of customers fixed within each first-purchase cohort as years since first purchase increase. In that figure, the gap narrows sharply after the first-purchase year but shows less additional convergence afterward, suggesting that the monotone gradient in Figure 9 reflects both within-relationship changes and selection into persistent relationships.

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



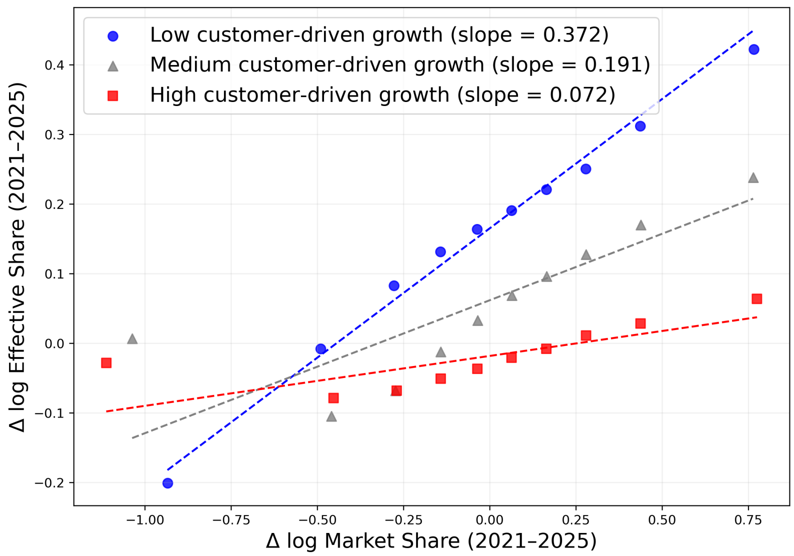

Notes: This figure follows cohorts of customers whose first observed purchase from a firm occurs in year $c$. For each cohort $c$ and each year $t \geq c$, we restrict to customers who continue to purchase from the firm in every year from $c$ through $t$. We then compute the log difference between this 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 each relevant cohort and comparison group are included. Unlike Figure A5, which restricts to relationships that survive through 2025, this figure conditions only on survival through each horizon $t$.

to existing customers, raising spending shares among incumbents. Figure 10 evaluates the balance of these forces by grouping firms into terciles based on the fraction of their 2021–2025 sales growth accounted for by net customer acquisition and then plotting the relationship between effective-share growth and market-share growth separately for each group.[49]

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 elasticity is roughly five times smaller for firms in the highest customer-acquisition tercile than for firms in the lowest tercile. At the same time, the elasticity 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. They also appear

[49]Specifically, we group firms into terciles based on their value of $\frac{|\Delta \log(\text{No. Customers}_i)|}{|\Delta \log(\text{No. Customers}_i)|+|\Delta \log(\text{Sales per Customer}_i)|}$. This measure is bounded between zero and one for all firms and is well-defined even in cases where the firm has very little growth overall or these two terms offset each other.

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.

to become more attractive to existing customers, who direct more of their category spending to the firm. Thus, customer acquisition attenuates effective-share growth, but it often occurs alongside incumbent deepening, which pushes effective shares up.[50]

# 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 primarily through this reach margin. We then revisit concentration trends and show that the rise of superstar firms in

[50]A simple example illustrates how customer-acquisition growth can dominate sales growth while incumbent deepening still raises effective shares. Suppose each customer spends \$100 in the category. A firm initially sells \$20 to each of 100 customers. It then adds 100 new 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. Thus, customer acquisition accounts for most sales growth, but the effective share rises from 0.2 to $(2500/3500) \cdot .25 + (1000/3500) \cdot .10 \approx .207$.

aggregate sales has not been accompanied by a comparable rise in effective shares.

**Geographic Expansion.** One important channel for recent firm growth, particularly in services, 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).

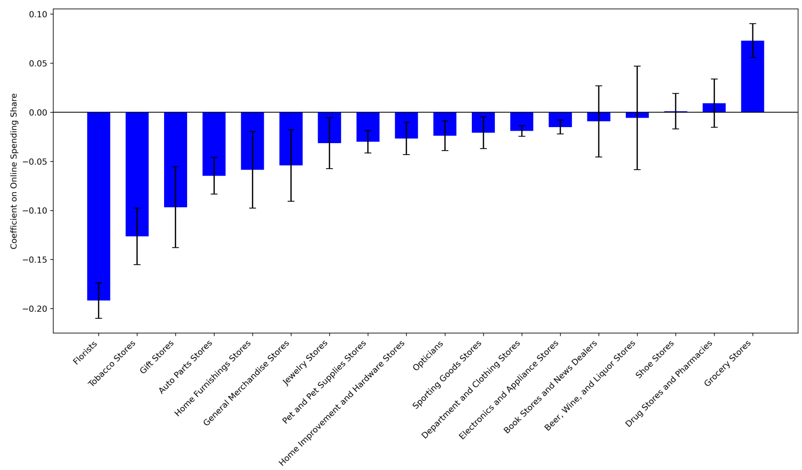
Geographic expansion tends to raise conventional market shares more than effective shares because it expands customer reach rather than deepening existing customer relationships. If customers in new markets looked identical to the firm's existing customers, geographic expansion would increase the firm's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

The left panel of Figure 12 shows the implications of geographic expansion for the relationship between effective-share growth and market-share growth. The purple squares show the relationship for firms that changed the number of 3-digit ZIP codes in which they operated between 2021 and 2025, while the green circles show the same relationship for firms whose geographic footprint remained fixed. For the same increase in market share, effective shares rise substantially more for firms that grow within an existing geographic footprint than for firms that grow by expanding into new markets. Geographic expansion necessarily expands reach by adding customers in new locations, but it does not generate a corresponding increase 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 sample of retail categories, online sales rose from around 10 percent of sales in 2014 to over 25 percent in 2025.[51] There was a pronounced increase 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

[51] 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 $\beta$ from Equation 14, estimated separately for each category. For each category, the regression specification includes firm fixed effects and controls for log of the customer's total spending at that firm, and is estimated using data from 2025. Vertical lines indicate 95 percent confidence intervals, constructed using standard errors clustered at the firm level.

by expanding customer reach. Online shopping makes it easier for customers to buy from a broader set of firms, so online customers may allocate a smaller share of category spending to any one firm than comparable in-person customers. If so, online growth will raise conventional market shares more than effective shares.

We begin by showing that customers who buy from a firm through its online channel allocate a smaller share of their category spending to that firm than customers who buy from the firm in person. Figure 11 reports estimates of $\beta$ from regressions of the form

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

where $\omega_{ijt}$ is the share of customer $i$'s category spending that goes to firm $j$, $x_{ijt}$ is the customer's total spending at firm $j$ in year $t$, $\gamma_j$ are firm fixed effects, and online share$_{i,j,t}$ is the fraction of customer $i$'s spending at firm $j$ that occurs online. For customer-firm pairs in which all spending at firm $j$ occurs online, this share equals one. For computational reasons, we restrict the sample to 2025. We control for total spending at the firm, since online customers tend to have higher incomes (Dolfen et al., 2023) and thus spend more at any given firm. The firm fixed effects compare customers of the same firm. Thus, the coefficient $\beta$ 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 at the firm occurs online.

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

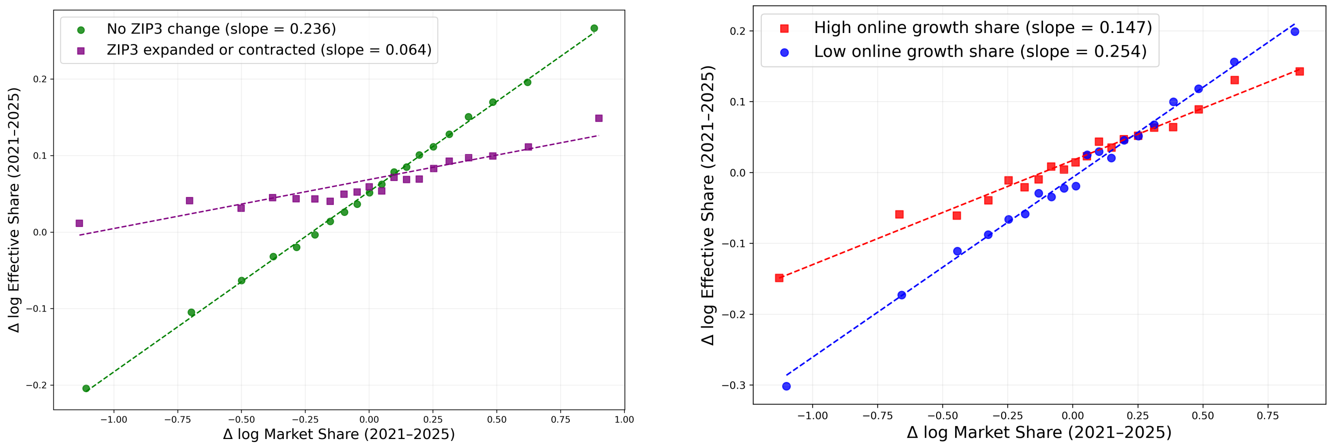
These customer-level patterns imply that, for a given increase in sales, firms should experience larger increases in effective shares when those sales come from in-person customers than when they come from online customers. The right panel of Figure 12 confirms this prediction. We split firms according to whether their sales growth came mostly through online or offline channels and estimate the relationship between effective-share growth and market-share growth separately for each group. Firms whose growth came mostly through online sales have a meaningfully weaker relationship between effective-share growth and market-share growth. Thus, online growth, like geographic expansion, expands customer reach more than customer depth.

**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 whether their rising aggregate sales shares have been accompanied by a comparable rise in customer dominance.

For this analysis we turn to our historical panel data. As discussed in Section 3, this data covers a sample from 2014–2025. This historical panel 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 historical panel does not permit constructing full customer baskets over all firms, 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 top 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 trends in ES4 are informative about trends in effective shares. We focus on national trends because large national firms are unlikely to be omitted from the named-firm panel.

Figure 12: The Nature of Firm Growth and Effective Shares




(a) Geographic Expansion

(b) Online Spending

Notes: Panel A splits firms according to whether the number of ZIP3 markets in which the firm operates changes between 2021 and 2025. Panel B restricts the sample to retail categories and splits firms by the share of total sales growth attributable to online sales growth. Online growth share is defined as $\frac{|g_{\text{Online}}|}{|g_{\text{Online}}|+|g_{\text{Offline}}|}$, where $g_{\text{Online}}$ is growth in log online sales and $g_{\text{Offline}}$ is the residual component of log sales growth not accounted for by online sales growth. The sample in Panel B is further restricted to firms with online growth share below 0.2 or above 0.8, labeled "Low online growth share" and "High online growth share," respectively. In both panels, the bottom and top 2.5 percent of the distribution of $\Delta \log$ market share are trimmed. Points denote weighted bin means from binscatters of $\Delta \log$ effective share against $\Delta \log$ market share, using common bins constructed from the estimation sample in each panel. Dashed lines show group-specific linear fits.

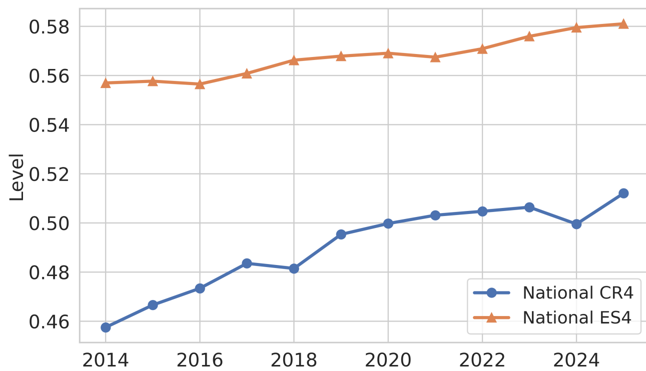
Figure 13 shows concentration trends in CR4 and ES4 in the historical panel.[52] The left panel shows levels, while the right panel normalizes each series to its 2014 value to highlight differences in growth. The blue line shows substantial growth in the national sales share of the top four firms from 2014 to 2025.[53] National CR4 increases by around 12 percent, from 0.457 to 0.512. By contrast, the orange line shows that average effective shares increase by only around 4 percent, from 0.557 to 0.581. Thus, the rise in national sales concentration substantially overstates the rise in firms' effective shares.

While we restrict our analysis to categories in which named firms make up a sizable share 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,

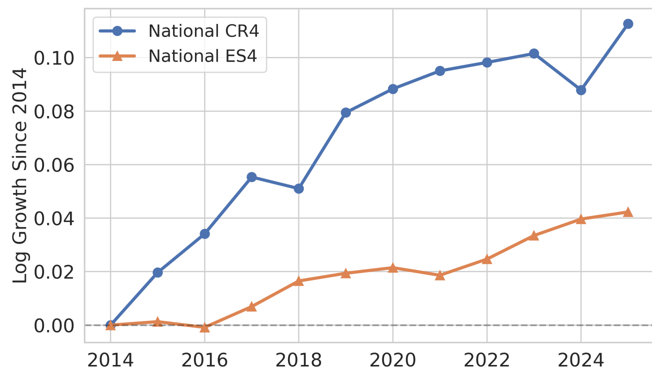
[52]To limit potential spurious composition effects from firms with different levels of effective shares entering and leaving the top four, we measure growth rates for the firms in the top four each year and chain these over time. Results are similar if we instead compute the average ES4 directly each year in the raw data.

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

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




(a) Concentration levels




(b) Concentration growth

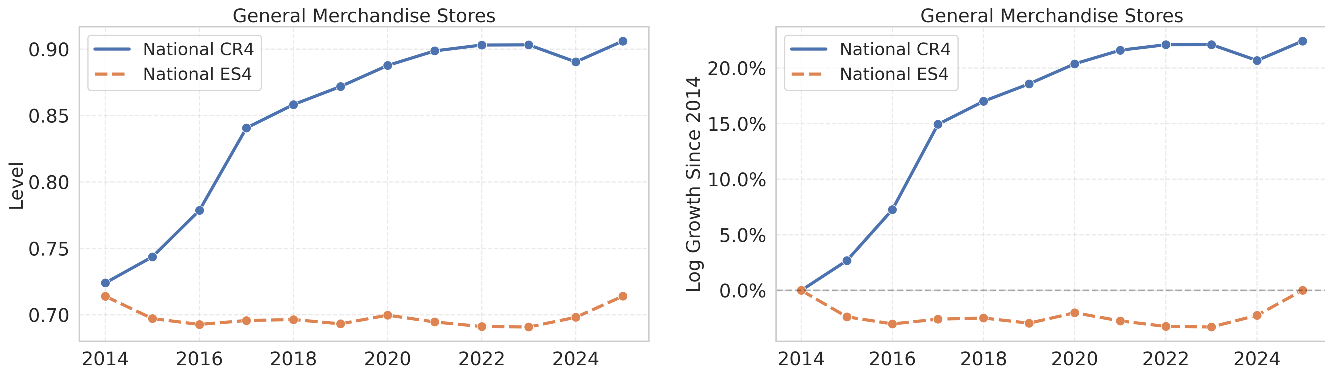
Notes: The figure reports concentration measures from 2014 to 2025. CR4 denotes the four-firm concentration ratio, and ES4 denotes the average effective share of the top four firms. Effective shares are constructed from the historical panel, with customers defined at the card level. To eliminate composition effects, the ES4 for each category is constructed by measuring the growth rate for each firm in the top four in a particular year and then chaining this growth rate over time. After constructing category-level concentration measures, we average across categories using fixed 2024 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 growth relative to 2014; for each measure $X$, the plotted value in year $t$ is $\log(X_t) - \log(X_{2014})$.

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 share is essentially flat.

More generally, the average trends in Figure 13 mask substantial heterogeneity across categories. The attenuated relationship between trends in sales shares and trends in effective shares is even more apparent when looking within particular categories. Table A10 compares 2014 to 2025 category-level changes in conventional concentration, measured by national CR4, to changes in effective concentration, measured by ES4 for the same firms. In most categories, switching from CR4 to ES4 moves the measured trend closer to zero, reinforcing the conclusion that changes in aggregate sales concentration do not translate into proportionate changes in effective shares.

The distinction between conventional and effective market shares is also related to the divergence between national and local concentration trends documented by Rossi-Hansberg, Sarte and Trachter (2021). In their data, national sales concentration rises while local sales concentration is relatively flat. One fundamental difficulty in interpreting differences in trends across

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.

geographic levels is that it is not obvious ex ante which geography is most relevant. National markets may be appropriate for some categories, firms, or customers, while much narrower local markets may be appropriate for others. Effective shares provide a complementary customer-level perspective by using realized shopping patterns rather than imposing a fixed geographic boundary. The fact that effective-share trends are much more muted than national concentration trends suggests that rising national concentration has not translated into a comparable rise in customer dominance.

Taken together with the evidence on growth margins, these results imply that the rise of "superstar" firms has not been met with a similar 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 appear extremely large using conventional market shares look much less exceptional using effective shares. Similarly, effective shares have risen much less than aggregate concentration over the last decade. Interpreting these patterns through the lens of the variable-markup framework that maps market power to firm "size" suggests little evidence of a broad-based increase in size-based product-market power in recent years.

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.

**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 Appendix

## A1.1 Data Description Details

### A1.1.1 Details of the Transaction data to TransUnion Credit Records

First, the TransUnion data had to be cleaned to ensure each card matched to only one person each month. This procedure was created by the credit card company, which we followed for years that had yet to be de-duplicated. Each person-card couple is noted with what type of account it is; examples include an individual account or an authorized user account. A hierarchy of account types was created and used for de-duplication. The hierarchy is as follows: an individual account, the maker of the account (i.e., the primary account holder, but has a guarantor), joint contractual liability for an account, an authorized user/guarantor/participant, an undesignated account type, and finally a terminated or deceased account.

Then, the TransUnion data was merged with the transaction-level data on the card-quarter level for 2021 and 2022, and on the card-month level for 2023 onward. TransUnion data was provided quarterly data for 2021-2022, and monthly data from 2023 on.

After the initial merge, a one-to-one match between cards and people was forced over the entire dataset, as some cards had varied owners over time. First, the person who appeared as the card owner for the longest period was designated the cardholder. In the case of a tie, the person with the maximum number of transactions was selected. Finally, in the case of another tie, the person who spent the most on the card was then selected as the cardholder.

Finally, some cards that appeared in the transaction data for a longer period of time than they appeared in the TransUnion data. Thus, for all cards where this was the case, we did a card-level merge between TransUnion's person identifier and the unmatched months of transaction level data. Then, each newly matched month's demographics data was filled in with information from the person's nearest month in TransUnion.

## A1.2 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).[54] 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

[54]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).

\$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.

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.[55] For each entry event, we identify the 5-digit ZIP that had the most customers visit the 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).[56] 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.

We report results for various robustness checks for the baseline results in Table 2 in Appendix Table A8. In one specification, we replace the incumbent's market share with a customer-specific analog, $\Omega_{i,t-1}^{\text{market}}$, which we define as the average (spending-weighted) local market shares of the establishments where that customer shops in $t-1$, an alternate way of thinking about the market share that is relevant for consumers' decision-making. 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. The baseline sample of entry events is heavily tilted towards restaurants, since this category has a large number of establishments and high churn, so we also run specifications excluding this category. Across all specifications, the results remain consistent: customers with higher effective shares are less likely to switch to shopping at newly entering firms.

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. 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.

[55] 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 relevant 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.

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

### A1.3 Validation of Entry Events

We use two methods to verify the estimated dates of firm entry in our transaction data. Between the two, we identify a total of 220 "verified" entering establishments.

1. **Yelp data**: For some of the entrants identified in the transaction data, we use both exact and fuzzy matching to find the date of the store's first review in the Yelp Academic dataset, a subset of Yelp's data with information on businesses' locations and review histories for some establishments. We consider an entry date in our transaction data to be "verified" with the Yelp Academic table if the date of the first transaction activity is less than four months away from the date of the entrant's first review on Yelp.
2. **Manual searches**: For other entering firms that are not present in the Yelp Academic table, we attempt to manually confirm their opening dates with a general Google search. We look for indicators such as the date of the store's grand opening announcement on its Facebook page or other social media accounts (if one exists), the first review for that store on Yelp (the whole site, not the subsample mentioned above), or news articles mentioning the opening of the new store.

# 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 $\theta$ 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 $\omega_j^{\text{eff}}$ 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 $\omega_{ij}$ is the within-category choice probability $\rho_{ij}$, 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, $\phi_j$ is a common quality shifter, $\mu_{ij}$ is a permanent customer–firm match component, and $\zeta_{ijt}$ is a transitory taste shock.[57] We assume that $\zeta_{ijt}$ follows a generalized extreme-value structure, with parameter $\eta$ governing substitution across firms within the category.[58] The term $\mu_{ij}$ captures persistent differences across households in the attractiveness of firms. Because $\mu_{ij}$ is fixed over time while $\zeta_{ijt}$ 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

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

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

[57] 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.

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

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 $\theta$. 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 $\rho_{ij}$ is the discrete-choice analogue of the customer-level spending share $\omega_{ij}$ of the earlier nested-CES benchmark; it is the latent object that pooled transaction data can recover under repeated purchase opportunities.[59]

As in the nested-CES benchmark, we derive closed-form customer- and firm-level demand elasticities, with the customer-level choice probability $\rho_{ij}$ now playing the role of the spending share $\omega_{ij}$, 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 $\rho_{ij}$ playing the role of the customer-level spending share $\omega_{ij}$.

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

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

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

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

[59] Fix prices and quality shifters $(p_j, \phi_j)$ and consider repeated purchase opportunities for household $i$ within the category. If the customer–firm match components $\mu_{ij}$ are fixed over time and the transitory taste shocks $\zeta_{ijt}$ 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 $\rho_{ij}$. 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.

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 $\omega_j^{\text{eff}}$ 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.[60]

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

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

[60] 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.

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 A9, 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.[61] Aggregating these customer-level elasticities then introduces both a firm-specific average price-sensitivity term and a price-sensitivity-weighted effective share.[62] In Appendix Table A9, 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.

## 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:[63]

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

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

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

[62] 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.

[63] For simplicity, in this section we use $\omega_{ij}$ 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.

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 $\omega_{ij}$ 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 $\omega_{ij}$ 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 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, $\omega_{ij}$ 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.

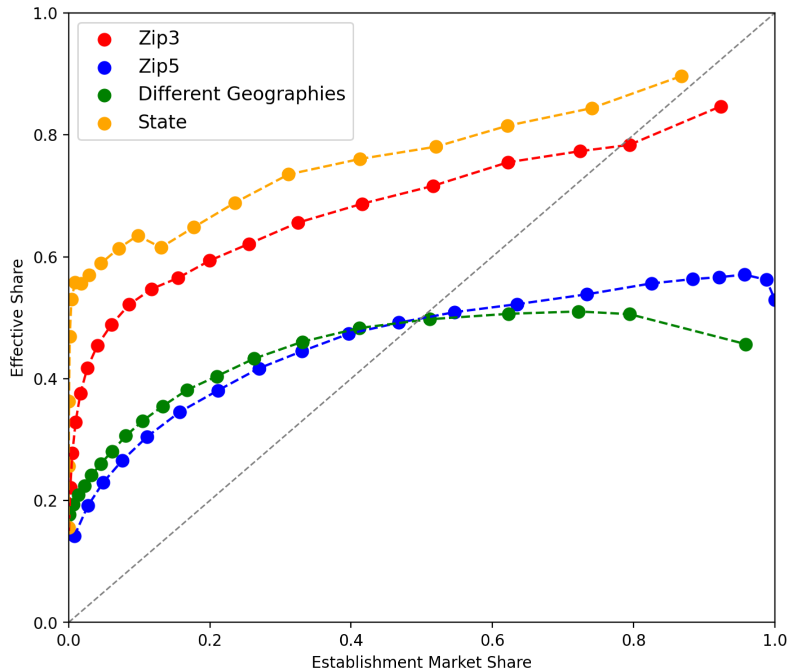
With heterogeneous customers, however, the relevant concentration object is different. Letting $y_{ij}$ denote spending of customer $i$ at firm $j$, $Y_j \equiv \sum_i y_{ij}$, $Y_i \equiv \sum_j y_{ij}$, and $Y \equiv \sum_i Y_i$, the sales-weighted sum of effective shares can be rewritten as

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

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.

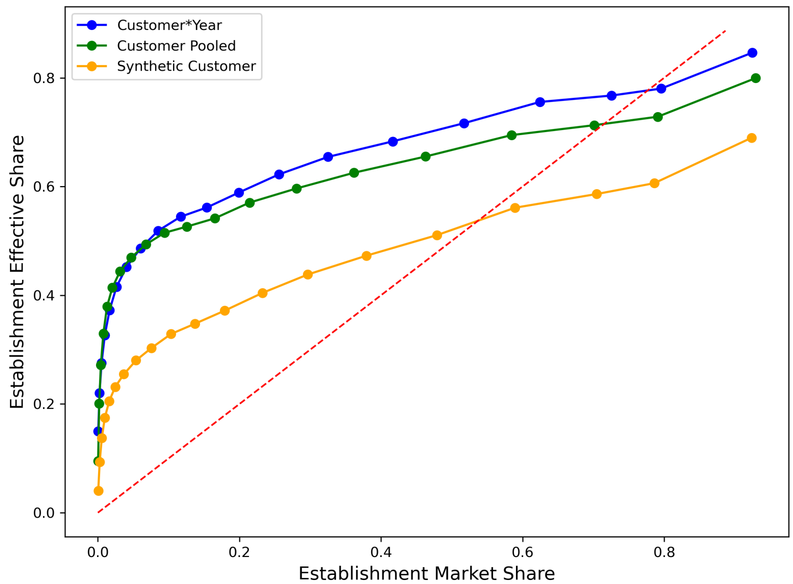
# A3 Appendix Tables and Figures

Figure A1: Robustness: Alternative Geographic Market Definitions



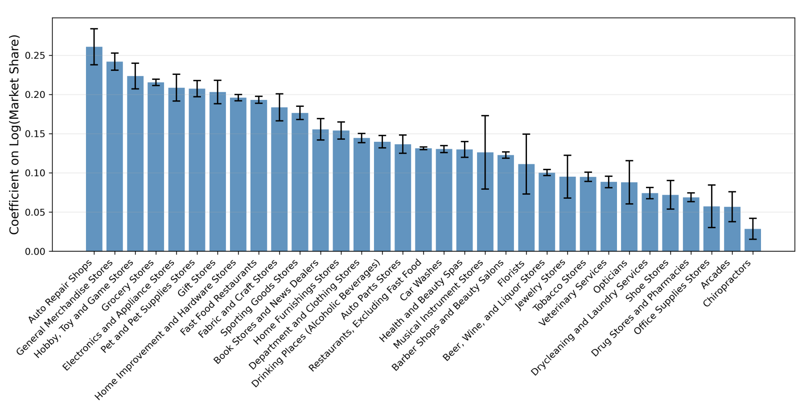

Notes: The dots are binscatter relationships with 20 bins, where the bins are defined by weighting firms by total establishment sales. Each bin plots the mean market effective share against the mean market share. Data is from 2025 and only includes in-person establishments with at least 200 customers. 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



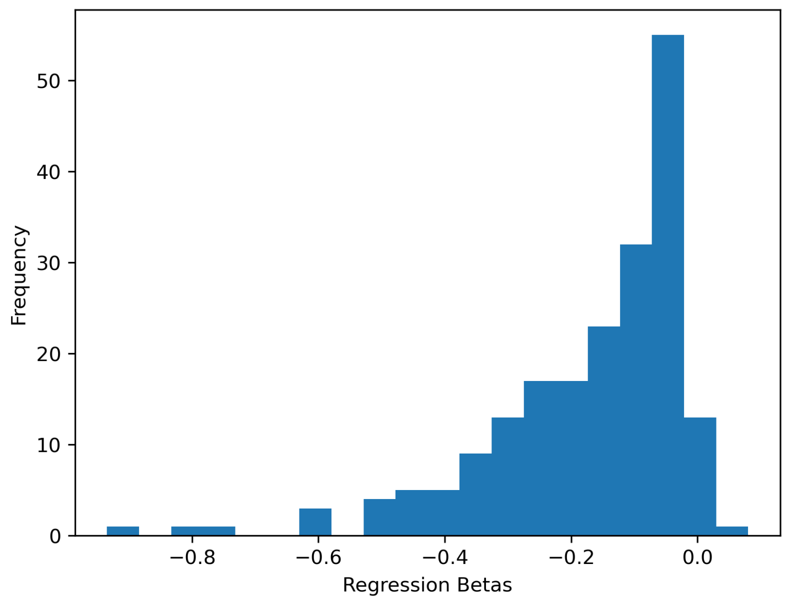

Notes: This figure shows binned scatterplots of effective shares and market shares computed using three distinct aggregation methods: pooling all spending within a year for each customer (our baseline approach), pooling all spending across the years 2021-2025 for each customer, or using our synthetic 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



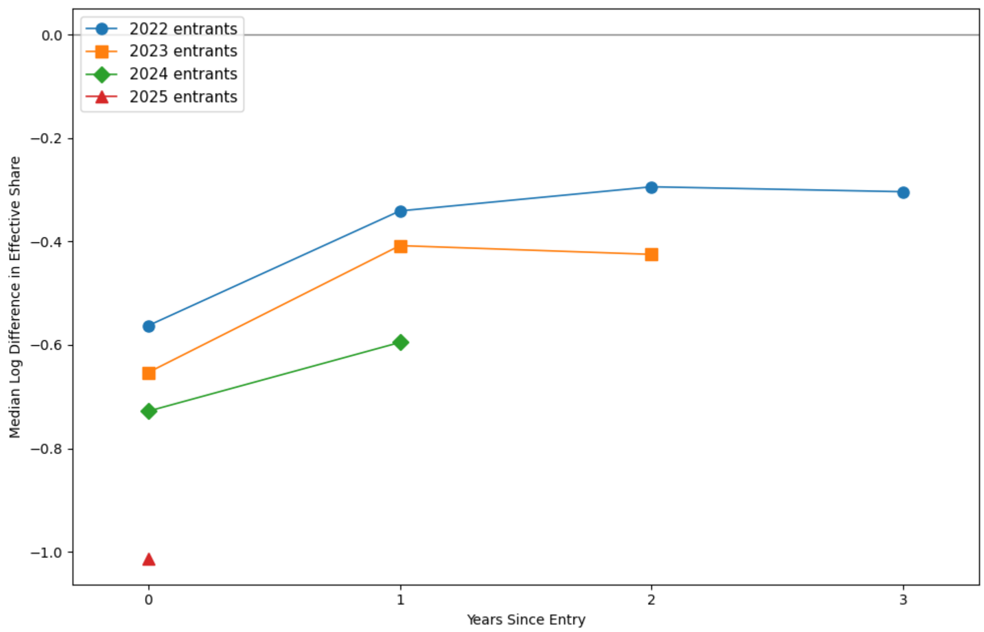

Notes: The figure reports category-specific estimates of the relationship between log firm effective market share and log firm market share. For each category, the plotted bar is the coefficient from a separate regression. Vertical black lines denote 95 percent confidence intervals. Categories are ordered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

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



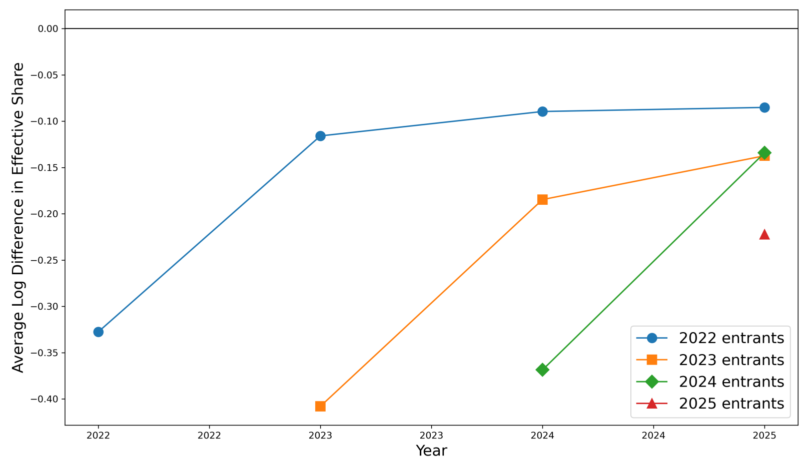

Notes: The figure plots the distribution of the coefficient on $\omega_{ij,e,-1}$ from the specification shown in Column 4 of Table 2 for 200 randomly selected entry events $e$.

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



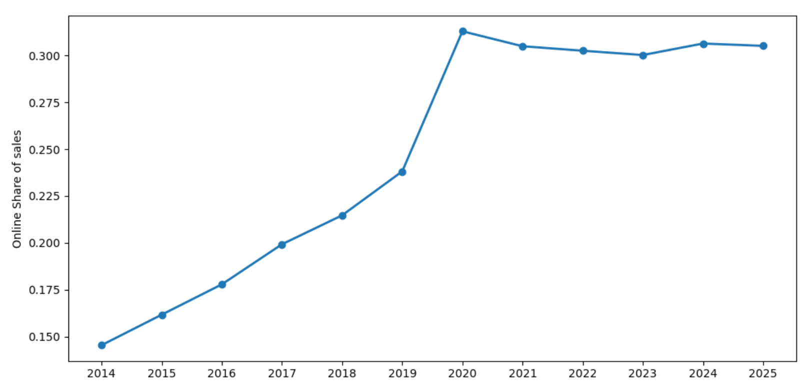

This figure follows cohorts of customers whose first observed purchase from a firm occurs in year $c$ and who continue to purchase from that firm in every year through 2025. For each cohort $c$ and each year $t \geq c$, we compute the log difference between the cohort's effective share in year $t$ and 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



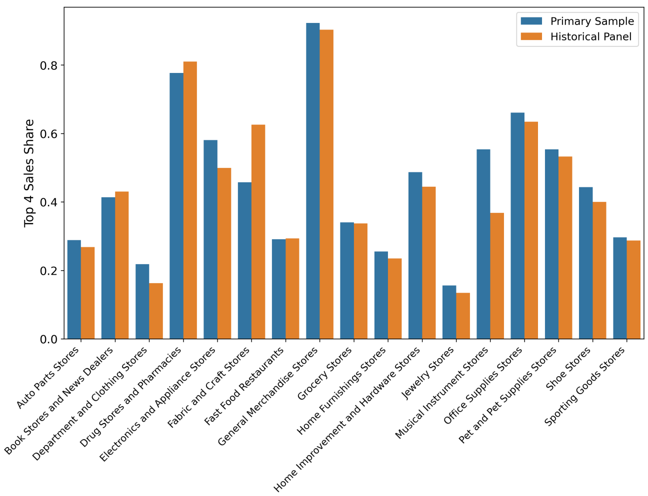

Notes: This figure presents differences in effective market share between newly entered and incumbent ZIP3 markets within the same firm. The sample consists of multi-establishment firms that operated in at least one ZIP3 market in 2021 and subsequently entered at least one additional ZIP3 market between 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

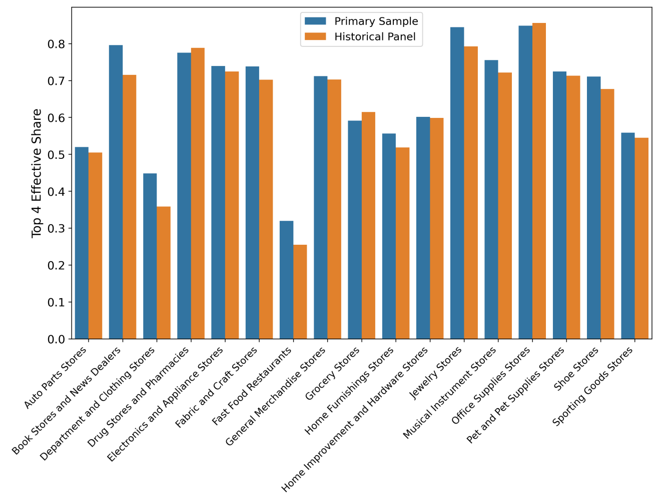



*Notes:* This figure plots the annual share of spending conducted online in the historical panel across retail categories from 2014 through 2025. Online spending share is measured as the card-not-present ("CNP") share of total sales, aggregated across retail categories in each year. The retail category 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.

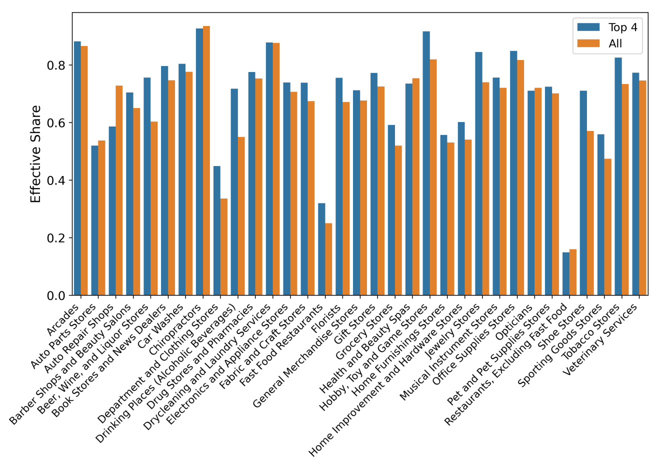
Figure A8: Validation of Historical Panel Data




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



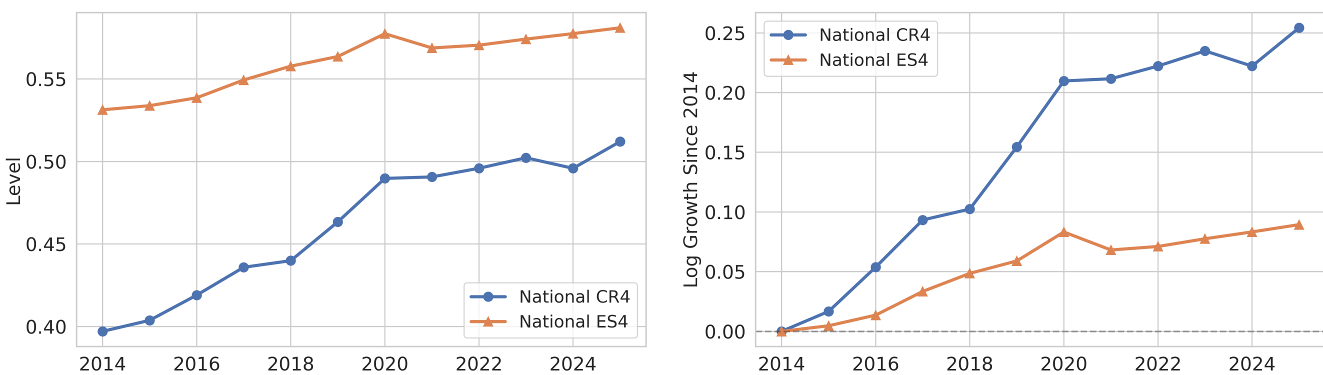

(b) Average of Top 4 Effective Shares: Primary vs Historical Sample




(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 2024: the four-firm concentration ratio (CR4) and the average effective share of the top four firms (ES4). 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 $\log(X_t) - \log(X_{2014})$.

Table A1: Summary Statistics for Various Analysis Samples

| | Customer*Year | Customer Pooled | Synthetic Customer | Transaction Data |
|---|---|---|---|---|
| **Customer Count** | 80.06 million | 36.15 million | 8.88 million | 1.21 billion |
| **Firm Count** | 2.45 million | 2.49 million | 2.67 million | 3.15 million |
| **Establishment Count** | 2.65 million | 2.68 million | 2.75 million | 3.56 million |
| **Total Sales** | 2.47 trillion | 2.08 trillion | 2.97 trillion | 8.74 trillion |
| **Online Share** | 39.71% | 38.71% | 38.94% | 34.02% |
| **Average Annual Spend** | 10,610.06 | 11,526.48 | 93,466.91 | 3,387.61 |

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 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

| Industry | Percent of People Dropped | | Percent of Spending Dropped | |
|---|---|---|---|---|
| | 5 Txns | 10 Txns | 5 Txns | 10 Txns |
| Shoe Stores | 88.2 | 97.9 | 66.8 | 87.4 |
| Pet and Pet Supplies Stores | 65.3 | 83.4 | 24.2 | 47.4 |
| Opticians | 94.8 | 99.5 | 86.0 | 97.0 |
| Musical Instrument Stores | 86.7 | 94.9 | 56.9 | 75.9 |
| Home Improvement and Hardware Stores | 47.9 | 67.5 | 12.3 | 26.0 |
| Florists | 94.3 | 98.9 | 71.3 | 84.5 |
| Car Washes | 74.2 | 90.2 | 44.2 | 68.4 |
| Home Furnishings Stores | 77.1 | 92.9 | 47.9 | 72.7 |
| Electronics and Appliance Stores | 86.4 | 96.9 | 66.8 | 87.5 |
| Auto Parts Stores | 80.1 | 93.2 | 50.9 | 70.8 |
| Sporting Goods Stores | 70.6 | 88.3 | 32.1 | 56.7 |
| Veterinary Services | 69.1 | 87.8 | 38.7 | 65.9 |
| Arcades | 89.1 | 96.7 | 63.1 | 78.6 |
| Office Supplies Stores | 88.3 | 97.4 | 60.9 | 79.9 |
| Drycleaning and Laundry Services | 79.3 | 90.7 | 36.2 | 55.5 |
| Grocery Stores | 13.0 | 24.6 | 1.0 | 3.0 |
| Auto Repair Shops | 90.6 | 98.8 | 75.3 | 94.3 |
| Drinking Places (Alcoholic Beverages) | 76.9 | 90.5 | 41.3 | 61.9 |
| Drug Stores and Pharmacies | 45.1 | 65.4 | 11.2 | 25.6 |
| Fabric and Craft Stores | 85.3 | 95.2 | 48.1 | 68.1 |
| General Merchandise Stores | 8.3 | 16.1 | 0.4 | 1.4 |
| Restaurants, Excluding Fast Food | 22.9 | 39.3 | 3.1 | 8.9 |
| Chiropractors | 66.8 | 83.4 | 36.6 | 57.4 |
| Book Stores and News Dealers | 84.4 | 95.2 | 53.9 | 75.9 |
| Beer, Wine, and Liquor Stores | 63.2 | 79.7 | 18.7 | 36.1 |
| Tobacco Stores | 72.7 | 85.9 | 22.4 | 40.0 |
| Department and Clothing Stores | 44.6 | 67.4 | 11.1 | 25.8 |
| Barber Shops and Beauty Salons | 64.3 | 85.1 | 27.2 | 54.3 |
| Fast Food Restaurants | 21.5 | 35.7 | 2.7 | 7.0 |
| Jewelry Stores | 96.3 | 99.1 | 81.0 | 92.1 |
| Hobby, Toy and Game Stores | 83.6 | 94.8 | 49.5 | 71.2 |
| Health and Beauty Spas | 79.1 | 91.9 | 47.0 | 71.2 |
| Gift Stores | 89.9 | 97.7 | 66.7 | 83.8 |

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

| Industry | Person*Year Average | Person*Year Median | Synthetic Customer Average | Synthetic Customer Median | Person Pooled Average | Person Pooled Median |
|---|---|---|---|---|---|---|
| Shoe Stores | 7.7 | 6 | 15.3 | 11 | 11.2 | 8 |
| Tobacco Stores | 14.5 | 10 | 19.1 | 12 | 27.0 | 13 |
| Jewelry Stores | 16.3 | 6 | 11.9 | 7 | 12.9 | 7 |
| Home Furnishings Stores | 8.9 | 7 | 21.3 | 13 | 14.3 | 10 |
| Drycleaning and Laundry Services | 12.6 | 9 | 21.3 | 12 | 21.4 | 11 |
| Office Supplies Stores | 8.4 | 6 | 14.3 | 10 | 12.1 | 8 |
| Drug Stores and Pharmacies | 16.7 | 12 | 74.7 | 37 | 45.3 | 25 |
| Florists | 8.4 | 6 | 10.8 | 8 | 10.6 | 7 |
| General Merchandise Stores | 80.3 | 54 | 563.6 | 294 | 356.5 | 246 |
| Restaurants, Excluding Fast Food | 29.8 | 19 | 183.6 | 79 | 106.2 | 62 |
| Opticians | 6.6 | 6 | 11.5 | 9 | 8.6 | 7 |
| Drinking Places (Alcoholic Beverages) | 12.2 | 8 | 23.9 | 14 | 19.4 | 10 |
| Department and Clothing Stores | 16.2 | 10 | 70.0 | 37 | 41.5 | 22 |
| Veterinary Services | 9.9 | 8 | 20.4 | 14 | 18.5 | 12 |
| Barber Shops and Beauty Salons | 10.1 | 8 | 30.2 | 17 | 22.6 | 15 |
| Fabric and Craft Stores | 9.9 | 7 | 14.6 | 10 | 15.5 | 9 |
| Car Washes | 9.5 | 8 | 22.1 | 14 | 17.4 | 11 |
| Musical Instrument Stores | 9.7 | 8 | 11.1 | 8 | 14.9 | 9 |
| Home Improvement and Hardware Stores | 17.2 | 11 | 72.9 | 36 | 43.5 | 23 |
| Electronics and Appliance Stores | 8.6 | 6 | 17.3 | 12 | 11.6 | 8 |
| Chiropractors | 11.9 | 9 | 15.0 | 11 | 20.5 | 12 |
| Sporting Goods Stores | 10.7 | 8 | 30.2 | 17 | 20.4 | 12 |
| Pet and Pet Supplies Stores | 11.6 | 9 | 26.6 | 17 | 24.0 | 14 |
| Hobby, Toy and Game Stores | 9.4 | 7 | 16.1 | 11 | 15.1 | 9 |
| Arcades | 11.0 | 7 | 11.9 | 8 | 13.7 | 8 |
| Gift Stores | 8.7 | 6 | 14.2 | 10 | 12.1 | 8 |
| Auto Repair Shops | 7.0 | 6 | 16.0 | 12 | 10.4 | 8 |
| Book Stores and News Dealers | 10.2 | 7 | 17.8 | 11 | 15.2 | 9 |
| Health and Beauty Spas | 10.0 | 8 | 15.7 | 11 | 16.9 | 10 |
| Beer, Wine, and Liquor Stores | 16.9 | 10 | 46.3 | 23 | 36.5 | 16 |
| Fast Food Restaurants | 41.8 | 23 | 243.3 | 102 | 147.1 | 73 |
| Grocery Stores | 52.7 | 34 | 342.9 | 149 | 211.9 | 133 |
| Auto Parts Stores | 10.3 | 7 | 24.7 | 17 | 16.5 | 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. The first four columns have an activity filter applied that only includes 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. There is an activity filter of at least 5 transactions within the industry over 5 years.

Table A4: Transunion Industry Spending Shares

| **Industry** | **Studied Industries** | **Total Spend** |
|---|---|---|
| General Merchandise Stores | 0.3129 | 0.2776 |
| Grocery Stores | 0.1682 | 0.1492 |
| Restaurants, Excluding Fast Food | 0.0931 | 0.0826 |
| Home Improvement and Hardware Stores | 0.0713 | 0.0632 |
| Department and Clothing Stores | 0.0654 | 0.0580 |
| Fast Food Restaurants | 0.0365 | 0.0324 |
| Home Furnishings Stores | 0.0340 | 0.0302 |
| Sporting Goods Stores | 0.0235 | 0.0209 |
| Electronics and Appliance Stores | 0.0229 | 0.0203 |
| Drug Stores and Pharmacies | 0.0197 | 0.0175 |
| Veterinary Services | 0.0193 | 0.0172 |
| Auto Parts Stores | 0.0180 | 0.0160 |
| Auto Repair Shops | 0.0178 | 0.0158 |
| Beer, Wine, and Liquor Stores | 0.0129 | 0.0114 |
| Jewelry Stores | 0.0116 | 0.0103 |
| Barber Shops and Beauty Salons | 0.0108 | 0.0096 |
| Pet and Pet Supplies Stores | 0.0097 | 0.0086 |
| Opticians | 0.0072 | 0.0064 |
| Shoe Stores | 0.0070 | 0.0062 |
| Drinking Places (Alcoholic Beverages) | 0.0052 | 0.0046 |
| Health and Beauty Spas | 0.0048 | 0.0042 |
| Gift Stores | 0.0035 | 0.0031 |
| Hobby, Toy and Game Stores | 0.0033 | 0.0029 |
| Book Stores and News Dealers | 0.0030 | 0.0027 |
| Car Washes | 0.0028 | 0.0025 |
| Tobacco Stores | 0.0025 | 0.0023 |
| Florists | 0.0023 | 0.0021 |
| Office Supplies Stores | 0.0023 | 0.0020 |
| Fabric and Craft Stores | 0.0023 | 0.0020 |
| Drycleaning and Laundry Services | 0.0020 | 0.0018 |
| Chiropractors | 0.0019 | 0.0017 |
| Musical Instrument Stores | 0.0018 | 0.0016 |
| Arcades | 0.0005 | 0.0004 |
| Total | 1.0000 | 0.8879 |

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(SPC) | 0.617 | | 0.615 | 0.555 | | 0.555 |
| | (0.000) | | (0.000) | (0.000) | | (0.000) |
| log(Customers) | | -0.066 | -0.049 | | 0.049 | 0.000 |
| | | (0.000) | (0.000) | | (0.000) | (0.000) |
| $R^2$ | 0.498 | 0.010 | 0.503 | 0.790 | 0.581 | 0.790 |
| Partial $R^2$ | 0.498 | 0.010 | 0.503 | 0.505 | 0.011 | 0.505 |
| Observations | 2538299 | 2538299 | 2538299 | 2538299 | 2538299 | 2538299 |
| **Panel B: National Panel** | | | | | | |
| log(SPC) | 0.598 | | 0.588 | 0.522 | | 0.525 |
| | (0.000) | | (0.000) | (0.001) | | (0.001) |
| log(Customers) | | -0.144 | -0.092 | | 0.008 | -0.023 |
| | | (0.001) | (0.000) | | (0.000) | (0.000) |
| $R^2$ | 0.508 | 0.037 | 0.523 | 0.787 | 0.587 | 0.788 |
| Partial $R^2$ | 0.508 | 0.037 | 0.523 | 0.483 | 0.000 | 0.486 |
| Observations | 1806162 | 1806162 | 1806162 | 1806162 | 1806162 | 1806162 |

Notes: The dependent variable in all columns is the log effective share. The regressors of interest are log sales per customer, log(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: Firm-level relationship between effective shares and customer retention

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| $\log \omega_{j,t-1}$ | 0.083*** | | 0.123*** | 0.066*** | 0.080*** | 0.094*** |
| | (0.006) | | (0.008) | (0.001) | (0.007) | (0.007) |
| $\log \Omega_{j,-1}$ | 0.026* | | | 0.029*** | 0.013 | -0.012* |
| | (0.013) | | | (0.001) | (0.012) | (0.007) |
| log Sales / Customer | | 0.044*** | -0.041** | | | |
| | | (0.015) | (0.016) | | | |
| log # of Customers | | 0.013 | 0.017** | | | |
| | | (0.017) | (0.008) | | | |
| Firm FE | | | | X | | |
| Synthetic Customer Shares | | | | | X | |
| Nationally Defined Firm | | | | | | X |
| Observations | 1,790,761 | 1,790,761 | 1,790,761 | 1,722,742 | 1,790,761 | 1,239,234 |
| $R^2$ | 0.523 | 0.368 | 0.535 | 0.889 | 0.488 | 0.469 |
| Within $R^2$ | 0.298 | 0.071 | 0.316 | 0.198 | 0.247 | 0.229 |

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. 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 A7: 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) |
| $\omega_{i,j,-1}$ | 0.038*** | 0.039*** | 0.040*** | 0.058*** | 0.654*** | 0.615*** | 0.572*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.013) | (0.015) | (0.017) |
| $\Omega_{j,-1}$ | 0.015 | 0.018 | 0.011 | -0.002 | 0.013 | 0.021 | 0.004 |
| | (0.014) | (0.017) | (0.018) | (0.017) | (0.034) | (0.049) | (0.057) |
| Market FE | X | X | X | X | X | X | X |
| Observations | 8,480,411 | 6,992,653 | 6,408,490 | 5,198,619 | 8,480,411 | 6,992,653 | 6,408,490 |
| $R^2$ | 0.253 | 0.310 | 0.351 | 0.381 | 0.376 | 0.351 | 0.343 |
| Within $R^2$ | 0.006 | 0.007 | 0.007 | 0.016 | 0.225 | 0.164 | 0.127 |

Notes: Columns (1), (2), (3), and (4) show the relationship between 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 and is active in the category in the given time period (all years for column (4)). 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 A8: Robustness Checks for Consumer Responses to Firm Entry

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| $\log \omega_{ij,-1}$ | -0.028*** | -0.029*** | -0.028*** | -0.027*** |
| | (0.001) | (0.003) | (0.001) | (0.004) |
| $\log \Omega_{i,-1}^{\text{market}}$ | -0.001 | -0.008*** | | |
| | (0.001) | (0.002) | | |
| $\log \Omega_{j,-1}$ | | | -0.008*** | -0.014*** |
| | | | (0.001) | (0.002) |
| Include Restaurants? | X | | X | |
| Incumbents Included | All | All | Entrant's ZIP3 | Entrant's ZIP3 |
| Observations | 34,345,857 | 2,968,696 | 20,100,085 | 1,857,580 |
| $R^2$ | 0.061 | 0.086 | 0.057 | 0.088 |
| Within $R^2$ | 0.012 | 0.016 | 0.015 | 0.023 |

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 A9: Robustness Checks for Log Relationship between Effective and Market Shares

| | (1) Base | (2) Sales Wtd | (3) Firm Spec. | (4) Service | (5) Retail | (6) Demo Control | (7) MU | (8) Card | (9) Card & Debit |
|---|---|---|---|---|---|---|---|---|---|
| *Panel A: Local (ZIP3)* | | | | | | | | | |
| $\log \Omega_{j,t}$ | 0.173*** | 0.169*** | 0.144*** | 0.144*** | 0.205*** | 0.181*** | 0.173*** | 0.088*** | 0.070*** |
| | (0.000) | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) |
| $R^2$ | 0.710 | 0.773 | 0.643 | 0.665 | 0.721 | 0.738 | 0.710 | 0.831 | 0.842 |
| Observations | 546,205 | 546,205 | 352,781 | 361,822 | 184,383 | 528,289 | 546,205 | 753,452 | 997,446 |
| Within $R^2$ | 0.235 | 0.327 | 0.149 | 0.150 | 0.374 | 0.307 | 0.235 | 0.107 | 0.079 |
| *Panel B: National* | | | | | | | | | |
| $\log \Omega_{j,t}$ | 0.156*** | 0.080*** | 0.140*** | 0.140*** | 0.181*** | 0.166*** | 0.156*** | 0.082*** | 0.081*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) |
| $R^2$ | 0.645 | 0.687 | 0.610 | 0.630 | 0.627 | 0.671 | 0.645 | 0.800 | 0.807 |
| Observations | 396,714 | 396,714 | 263,336 | 271,686 | 125,028 | 391,586 | 396,714 | 548,674 | 744,969 |
| Within $R^2$ | 0.155 | 0.072 | 0.124 | 0.124 | 0.210 | 0.215 | 0.155 | 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. 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 A10: Four-Firm Concentration and Effective Share by Retail Category, 2014 and 2025

| | 2014 | | 2025 | |
|---|---|---|---|---|
| Industry | CR4 (%) | ES4 (%) | CR4 (%) | ES4 (%) |
| Auto Parts Stores | 25.5 | 52.3 | 31.0 | 55.4 |
| Book Stores and News Dealers | 41.6 | 72.3 | 43.2 | 69.4 |
| Department and Clothing Stores | 18.4 | 31.3 | 19.0 | 38.2 |
| Drug Stores and Pharmacies | 81.0 | 77.7 | 69.5 | 78.8 |
| Electronics and Appliance Stores | 49.2 | 70.4 | 45.0 | 72.9 |
| Fabric and Craft Stores | 57.1 | 61.8 | 63.6 | 74.2 |
| Fast Food Restaurants | 26.2 | 24.0 | 30.1 | 25.4 |
| General Merchandise Stores | 72.4 | 71.4 | 90.6 | 71.4 |
| Grocery Stores | 35.4 | 56.5 | 33.9 | 60.5 |
| Home Furnishings Stores | 21.1 | 54.6 | 25.3 | 58.3 |
| Home Improvement and Hardware Stores | 49.2 | 58.4 | 48.5 | 60.5 |
| Jewelry Stores | 16.2 | 90.0 | 13.8 | 85.3 |
| Musical Instrument Stores | 32.8 | 71.1 | 45.6 | 72.7 |
| Office Supplies Stores | 75.6 | 80.7 | 60.5 | 83.8 |
| Pet and Pet Supplies Stores | 48.7 | 74.6 | 53.8 | 74.0 |
| Shoe Stores | 44.0 | 67.4 | 37.3 | 66.2 |
| Sporting Goods Stores | 25.9 | 51.0 | 34.9 | 54.1 |

*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 A11: Median Log Difference in 2025 Effective Shares Between 2025 Entrants and Consistent Shoppers, By Industry

| Industry | Median |
|---|---|
| Grocery Stores | -0.60 |
| Department and Clothing Stores | -0.27 |
| Drug Stores and Pharmacies | -0.44 |
| Electronics and Appliance Stores | -0.20 |
| Home Furnishings Stores | -0.19 |
| General Merchandise Stores | -0.45 |
| Auto Parts Stores | -0.17 |
| Auto Repair Shops | -0.37 |
| Barber Shops and Beauty Salons | -0.49 |
| Beer, Wine, and Liquor Stores | -0.78 |
| Book Stores and News Dealers | -0.59 |
| Car Washes | -0.45 |
| Chiropractors | -0.23 |
| Drinking Places (Alcoholic Beverages) | -0.70 |
| Florists | -0.53 |
| Gift Stores | -0.50 |
| Health and Beauty Spas | -0.41 |
| Home Improvement and Hardware Stores | -0.14 |
| Jewelry Stores | -0.47 |
| Opticians | -0.28 |
| Pet and Pet Supplies Stores | -0.51 |
| Shoe Stores | -0.41 |
| Sporting Goods Stores | -0.19 |
| Tobacco Stores | -0.51 |
| Veterinary Services | -0.23 |
| Fast Food Restaurants | -0.48 |
| Restaurants, Excluding Fast Food | -0.52 |
| Drycleaning and Laundry Services | -0.36 |
| Fabric and Craft Stores | -0.46 |
| Hobby, Toy and Game Stores | -0.52 |
| Musical Instrument Stores | -0.46 |
| Office Supplies Stores | -0.34 |
| Arcades | -0.42 |

*Notes:* For each firm, we compute the sales-share weighted average log effective share separately for 2025 entrants and customers who shopped at the firm every year from 2021-2025, and take the difference for each firm. The table reports the median firm-level log difference in effective shares within each category. A negative value signifies a lower effective share for 2025 entrants versus consistent shoppers.

Table A12: Mapping of Merchant Category Codes (MCC) to Industry Groups and Census Industries

| MCC | MCC Name | Industry Group | Census Industry |
|---|---|---|---|
| 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.