NBER WORKING PAPER SERIES

LONG-TERM VERTICAL CONTRACTS, GEOGRAPHY, AND THE PERSISTENCE OF BRAND SHARES

Robert Clark
Jean-François Houde
Xinrong Zhu

Working Paper 33906
http://www.nber.org/papers/w33906

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2025

We thank Jean-Pierre Dubé, Markus Reisinger, Stephan Seiler, Raphael Thomadsen, Christopher Whaley, and numerous seminar participants for helpful comments. Researcher(s) own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. Main estimates and analyses in this paper based on SymphonyIRI Group, Inc. data are by the author and not by SymphonyIRI Group, Inc. All remaining mistakes are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w33906

Long-term Vertical Contracts, Geography, and the Persistence of Brand Shares
Robert Clark, Jean-François Houde, and Xinrong Zhu
NBER Working Paper No. 33906
June 2025
JEL No. L42, L81

ABSTRACT

Previous research highlights persistent differences in national brand market shares across regions, driven by geographic consumer preferences and early entry advantages (Bronnenberg et al. 2007, 2009). This study investigates the extent to which long-term vertical contracts can explain the observed geographic dispersion and persistence of national brand dominance. Using NielsenIQ data from the same product categories as Bronnenberg et al. (2007), we follow Abowd et al. (1999) to decompose the variance in brand shares into Brand × Retailer and Brand × Market effects. Results show that 50% of the variance is explained by retailer effects, compared to 20% by market effects. This suggests that some retailers use vertical contracts that favor a specific brand across all markets in which they compete, such as exclusive-dealing, slotting allowances, or category-captaincy contracts, and that these play a key role in sustaining national brand dominance documented in prior literature. By emphasizing how retailer-specific factors drive observed patterns, we complement demand-side explanations and shift attention to supply-side factors that could affect competition.

Robert Clark
Queen's University
clarkr@econ.queensu.ca

Jean-François Houde
University of Wisconsin - Madison
Department of Economics
and NBER
houde@wisc.edu

Xinrong Zhu
Imperial College London
x.zhu@imperial.ac.uk

1 Introduction

For many consumer products, a small number of widely distributed brands dominate sales. These brands, often referred to as *top national brands*, are presumed to consistently lead market shares across all markets and stores where they are sold. However, as noted by Dubé and Rossi (2019), the dominance of these leading national brands can exhibit significant variation, both across countries (Sutton 1991) and even across regions within the same country (Bronnenberg et al. 2007). Bronnenberg et al. (2009, 2012) attribute this variation in brand shares across markets to persistent geographic differences in consumer preferences, as well as to the advantages gained by brands that were early entrants in specific markets.

While geographic and historical factors are important, retailers may also play a crucial role in shaping the dispersion of brand shares. Vertical relationships between retailers and manufacturers can influence brand performance, as long-term contracts often provide preferential treatment to specific brands through mechanisms such as slotting allowances (Marx and Shaffer 2007, Hristakeva 2022b), exclusive dealing (Asker 2016), vertical rebates (Conlon and Mortimer 2021), full-line forcing contracts (Ho et al. 2012), or category captaincy (Viswanathan et al. 2020, Zhu 2024).[1,2] For instance, in category captaincy arrangements, a leading manufacturer is tasked with managing the product placement, pricing, and assortment for all brands within a given category, often to the benefit of their own brand.[3] The spatial distribution of retailers—especially regional ones concentrated in specific markets—may also contribute to brand share dispersion. Together, these factors can impact the spatial distribution of brand shares and the heterogeneity of market outcomes across top national brands.

[1]Slotting allowances refer to manufacturers' upfront payments that are independent of the retailers' quantity purchases to buy up scarce shelf space. A vertical rebate is a discount or financial incentive provided by a manufacturer to a retailer based on the retailer's purchases or sales performance.

[2]Contractual arrangements can lead to similar asymmetries in other settings too. For instance, health insurers implement narrow and tiered networks to reduce costs by prioritizing or contracting with specific providers. See for instance, Sorensen (2003), Dafny et al. (2015), Ho and Lee (2019), and Prager (2020) for examination of narrow networks, and Ho and Lee (2024) for tiering.

[3]Many companies reportedly engage in category captaincy arrangements with retailers. The following facts and quotes from the industry press suggest that the practice is being used in grocery retail industry: General Mills has won Category Captain Awards in the yogurt category for 2011, 2014, and 2018 by Progressive Grocer (Progressive Grocer 2011, 2015b, 2018). Danone "has built relationships with retailers, constituting a competitive advantage over new entrants or smaller players," according to analysis by the financial services firm, Morningstar, Inc (Morning Star Analysis Report 2020). Flowers Foods delivered impactful bench-marking, promotion analysis, assortment studies and outstanding in-store execution (Progressive Grocer 2015a). Groupo Bimbo developed the right assortment..., which allowed consumers to see more selection, while delivering the freshest product available and driving down waste (Progressive Grocer 2015a).

In this paper we revisit the relationship between geography and brand shares to determine how much of the observed spatial dispersion of national brand-shares is associated with market-specific unobserved heterogeneity (such as consumer preferences and brand history), and how much with retailer-specific factors (such as the presence of long-term vertical contracts). Our main contribution is to provide evidence of the important role of retailers in explaining the observed share asymmetries. Understanding the sources of dispersion in market shares – whether it stems from consumer preference and loyalty towards specific brands (brand-level switching costs) or retailer-level contractual relationships (retailer-level switching costs) – is critical for both theoretical and empirical studies.[4] Attributing persistent market share patterns to consumer preferences and habit formation can confound demand-side brand loyalty with supply-side retailer practices (e.g., exclusive agreements, slotting fees, and category captaincy), which could constrain consumer choices, distort competition, and may lead to brand concentration within retailers. Diagnosing asymmetric vertical contracts is also important for antitrust authorities, as illustrated by recent FTC allegations that a soft-drink manufacturer's practice of offering promotional payments to a particular big-box retailer and not to others in exchange for favorable shelf-space is anticompetitive.

To perform our analysis, we construct a sample of the same 25 product categories as in Bronnenberg et al. (2007) from NielsenIQ Scanner Data. Our sample spans the period from 2013 to 2016, and covers 109 retailers and 199 markets, as designated by NielsenIQ. Following Bronnenberg et al. (2007), we focus on the two brands with the highest national market shares in each category. Using our sample, we first confirm that the degree of share dispersion across markets is similar to Bronnenberg et al. (2007), but that an even greater share dispersion exists among retailers. The average interquartile range of top brands' shares is 0.79 across retailers compared to 0.64 across markets, highlighting the significant role of retailers in driving share dispersion. This significant variation in brand performance across retailers can, in part, be attributed to differences in vertical contracting and retail distribution strategies. Some retailers reportedly establish long-term vertical contracts with specific brands and utilize centralized distribution systems. At these retailers, the leading brand consistently maintains high market shares across all locations. In contrast, other retailers adopt a more decentralized approach, with product assortments that are more closely tailored to local market preferences-an observation that aligns with the findings of Bronnenberg et al. (2007).[5] We also provide descriptive evidence that, within retailer, shares

[4] See Bronnenberg and Dubé (2017) for a review of the brand loyalty literature, and Dubé et al. (2009, 2010) for specific examples.

[5] Walmart and Albertsons exemplify two distinct approaches to retail management through their use of centralized and decentralized systems, respectively. Walmart operates with a highly centralized supply model (`https://www.allthingssupplychain.com/the-amazing-supply-chain-of-walmart/`). In contrast, Albertsons employs a

of leading brands are very persistent over time. Focusing on the retailer-category pairs in which the leading brand has a significant share advantage in year t, the probability that in year $t+1$ a retailer's leading brand in a given category is the same as in t is 95%, roughly the same as the probability of persistence at the market level.

The descriptive evidence highlights significant heterogeneity across retailers in shares and assortments. This motivates our main research question: To what extent do retailers contribute to share dispersion and concentration among top national brands? Identifying these effects is challenging, as regional retailers (comprising 58% of those in the NielsenIQ data) often operate across neighboring markets, creating a positive correlation between retailer and market share distributions.

To overcome this challenge, we adapt the "two-way fixed effect" model from the labor literature, commonly used to disentangle worker and firm effects in wage dispersion (Abowd et al. 1999) (AKM). Our model decomposes the total variance in quantity shares, assortments, availability, and prices of the top two national brands into three components: (i) a Brand $\times$ Retailer match effect, (ii) a Brand $\times$ Market match effect, and (iii) the correlation between these two effects. This approach allows us to account for both retailer and market effects within a unified framework while explicitly modeling their correlation.

The results from the decomposition indicate that the Brand $\times$ Retailer component consistently accounts for more than 50% of the total variation in quantity shares, product assortments, availability, and prices of the top two national brands. In contrast, the Brand $\times$ Market component explains less than 20% of the observed variation in these outcome variables. Additionally, the variance decomposition of category concentration (measured by HHI) shows that the retailer dimension accounts for five times more variation than the market dimension. These results suggest that more than half of the difference in top national brands' market shares is due to differences in how retailers sell these brands, including their assortment and pricing choices, even under similar market conditions.

Using the fixed effects estimated with the AKM model, we revisit findings in Bronnenberg et al. (2009) to analyze and compare the magnitude of retailer and market heterogeneity. We find that accounting for retailer heterogeneity significantly reduces the geographic demand variation documented in prior studies. By using Brand $\times$ Market fixed effects as a proxy for market demand heterogeneity, we confirm the "city of origin effect," where a brand's market share declines with distance from its origin (Bronnenberg et al. 2009). However, this effect reduces by half when controlling for Brand $\times$ Retailer fixed effects, indicating that much of the observed spatial share dispersion is driven by retailer-brand-specific factors rather than geographic demand differences alone. We also observe a positive correlation between the

decentralized structure (https://scw-mag.com/news/1098-albertsons-companies/).

proximity of brands' cities of origin and retailers' primary markets and the retailer-brand effects. This correlation is robust to the inclusion of market-brand effects.

Together, these results suggest that many multi-market retailers enter into long-term vertical arrangements that favor brands that are historically dominant in their core markets. For instance, consider the following (fictional) example of a multi-market retailer using a centralized procurement strategy, similar to those employed by Target and Walmart. Due to the presence of local brand preferences and/or contractual frictions, this retailer is more likely to rely on a vertical arrangement that favors brands that have historically been dominant in its largest market. Moreover, the retailer can leverage its size by having a uniform assortment across all markets and negotiating better terms with upstream producers (a buy-power effect). The first feature leads to the positive correlation between brand history and local market shares documented in Bronnenberg et al. (2009), while the second feature leads to a strong retailer-brand effect that we associate with asymmetric vertical contracts. Importantly, the source of the spatial correlation between history and local shares is caused by the presence of asymmetric vertical contracts and uniform assortments, rather than brand loyalty alone.

We then provide empirical evidence that the retailer-brand-specific factors reflect the existence and impact of retailer-brand vertical relationships. Specifically, we analyze the Brand $\times$ Retailer fixed effects derived from the AKM model as a proxy for retailer-brand matched heterogeneity. These effects are notably larger in retailers with higher revenues and broader geographic reach, retailers with smaller private label segments, and categories characterized by less differentiated national brands. The fixed effects exhibit high correlation over time, indicating stable share dispersion within retailers. Together, this evidence highlights the role of vertical relationships in reinforcing market power within retailers, as documented in Draganska and Jain (2005), Scott Morton and Zettelmeyer (2004), Ellickson et al. (2018), and helps to explain the persistent market share advantages of certain top brands in specific markets.

Our findings are quantitatively different from the prior literature on the variability of brand shares. Bronnenberg et al. (2009) report that the retail component accounts for just 20% of the variance of national brand shares (vs 51% for the market component). We attribute the difference between their findings and ours at least in part to the difference in sample period, since Bronnenberg et al. (2009) focus on 1992 to 1995, while our primary sample spans from 2013 to 2016. As discussed above, and in further detail in the analysis below, there exist some retailers that tailor their assortment to local markets as implied in Bronnenberg et al. (2009), but there is also a high prevalence of retailers with long-term vertical contracts with specific brands. Hwang and Thomadsen (2016) employ variance decomposition methods (ANOVA and COV), and find some variation in the shares of major national brands across markets and retailers. However, compared to their findings, we observe a much stronger retailer effect. Through comprehensive replication analyses, we attribute the difference in results primarily to differences in data sources and their geographic coverage in particular. The IRI data (2004–2005) used by Hwang and Thomadsen (2016) featured a set of retailers with more limited geographic coverage than those in the more recent NielsenIQ data. For example, retailers in the IRI data span an average of 2.44 markets, whereas NielsenIQ retailers are present in 8 markets. This limited geographic coverage in the IRI data impacts identification in variance decomposition models.

Our study makes several contributions to the literature. First, it advances research on competition and market structure in retail industries. A key contribution of our analysis is clarifying whether dispersion in market shares arises from brand loyalty (brand-level switching costs) or vertical relationships (retailer-level switching costs). While prior work emphasizes geographic variations in brand shares driven by consumer preferences and early entry advantages (Bronnenberg et al. 2009, 2012), as well as national trends in concentration tied to dominant firms (Rossi-Hansberg et al. 2021), we demonstrate that retailer heterogeneity – specifically vertical contracts and retailer-level switching costs – explains at least half of the documented variation in demand across markets. By emphasizing how retailer-specific factors drive observed patterns, we complement demand-side explanations and shift attention to supply-side factors.

We also highlight the role of vertical relationships for the sustained within-retailer share concentration and across-retailer dispersion, offering insights into how these relationships can impact broader market share distributions and providing implications for competition policy. There are two broad areas of concern. The first is related to exclusive contracts. While such contracts can represent an efficient way to generate both buyer power and competition between brands, they can also be used to exclude rivals if retailers have high costs of switching brands.[6] In this sense, exclusive dealing disrupts brands' incumbency advantages and raises barriers to entry in markets where a dominant brand is the early entrant. Second, there are concerns that category captaincy contracts between retailers and brands can facilitate collusion. Clark et al. (2024) show how such arrangements contributed to the formation of a cartel that linked retailers and manufacturers in Canada's bread industry.

Our research also contributes to the existing literature on vertical relationships and their impacts on the retail sector. A great deal of research has focused on the impact of vertical relationships on prices within a supply chain (Villas-Boas 2007, Bonnet and Dubois 2010,

[6] See for instance Aghion and Bolton (1987), Rasmusen et al. (1991), Segal and Whinston (2000), and Asker and Bar-Isaac (2014).

Dubois et al. 2015, Luco and Marshall 2020). A sparser body of literature focuses on the effects of vertical contracts on product availability and assortment (Viswanathan et al. 2020, Hristakeva 2022b, Zhu 2024). We use the AKM model and the fixed effects extracted from it to provide empirical evidence consistent with the prevalence of vertical practices, where retailers give preferential treatment (in terms of product assortment and pricing) to specific brands over others. Our paper also relates to research on retailer-level marketing decisions (DellaVigna and Gentzkow 2019, Hitsch et al. 2019). We show that the retailer leading brand's share distributions are stable over time and across markets. This suggests that retail prices, assortment, and shelf allocations are determined at the retailer level.

Finally, we are also related tangentially to the growing literature on endogenous product choice, which includes studies such as Draganska et al. (2009), Fan (2013), Eizenberg (2014), Fan and Yang (2020, 2024), and Hollenbeck et al. (2024). In these papers retailers or manufacturers make assortment choices.

The remainder of the paper is structured as follows: Section 2 outlines the data source, sample construction, key data features, and outcome measures. Section 3 provides empirical evidence on share dispersion across the retailer dimension. Section 4 introduces the AKM variance decomposition model and presents the results. Section 5 analyzes the fixed effects extracted from the AKM model, offering insights into the role of long-term vertical contracts. Finally, Section 6 concludes the paper.

2 Data and Summary Statistics

In this section, we first describe our data sources and outline the process of sample construction. Next, we highlight key geographic characteristics of retailers, which are crucial for distinguishing between market-level and retailer-level share dispersion. Finally, we introduce our primary outcome measures and present summary statistics.

2.1 Data Sources and Sample Construction

Our primary data source is the NielsenIQ Retail Scanner ("RMS") dataset provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The RMS data record weekly quantity sold, revenue, and product information of roughly four million UPCs (universal product code) from about 35,000 stores and 200 markets across the U.S.[7] Product information includes UPC, UPC description, brand name,

[7]Market definition in NielsenIQ data is DMA (designated market area). A DMA typically embodies a single metropolitan area. The NielsenIQ dataset defines markets using DMAs (Designated Market Area), which are often tied to major cities of the US (in some cases larger than one city). Hereafter we use "market" and "DMA"

package size, and a rich set of product characteristics. The dataset also provides geographic information about stores and the retailers to which they belong.[8]

To precisely construct a product's brand structure, we combine NielsenIQ Scanner data with GS1 data, which provide a unique mapping between the UPC and parent company names. The GS1 dataset covers 10,871 parent companies. Forty percent of the branded UPCs from the NielsenIQ data can be directly matched with their parent company in GS1. The rest are identified using the first six digits from the UPC.[9] Although many of the brands are jointly owned by a small number of large conglomerates, we treat each parent brand in our sample as a separate entity. For example, Yoplait in the yogurt category is owned by General Mills but operates with its own management team. Therefore, we consider Yoplait as a parent brand. In the data, a "parent brand" encompasses all the UPCs sold under a given brand name. Each UPC identifies an individual item produced by a manufacturer. The UPCs for a six-pack of strawberry yogurt, a single container of strawberry yogurt, and a single blueberry yogurt from the same manufacturer would each be unique and distinct from one another. We adopt a product definition that is more aggregated than UPCs, and less aggregated than parent brand. Specifically, a *product* is a combination of a brand code and package size, both of which are provided by NielsenIQ. The brand code is an identifier for sub-brand (e.g., Kashi Go Lean from Kellogg's, or Dannon Light n Fit from Dannon). Therefore, UPCs that share the same category - parent brand (e.g., cereal - Kellogg's), the same sub-brand (Kashi Go Lean), and the same size, are considered to be the same product.

We follow Bronnenberg et al. (2007) & Bronnenberg et al. (2009) and construct a sample from 25 categories based on Table 3 of Bronnenberg et al. (2009).[10] The complete list of these 25 categories is provided in Table A1 in Appendix A.[11] Our primary sample spans four years, from 2013 to 2016, offering more recent and geographically extensive data compared to prior studies such as Bronnenberg et al. (2007, 2009) and Hwang and Thomadsen (2016). Additionally, for analyzing share variation over time, we construct a longer sample covering the period from 2006 to 2016.

interchangeably.

[8]NielsenIQ records the parent company that owns a chain (*parent_code*), and the retailer banner itself (*retailer_code*). The names of the retailers are not disclosed by NielsenIQ in the data. In our analysis, we use the *retailer_code* as retailer identifier. We also conduct robustness tests using *parent_code* instead of *retailer_code*, the conclusions remain very similar.

[9]The first six digits are a company's unique identification number in the UPC. Every company with this number assigns it to all of its products. `https://bytescout.com/blog/2013/10/upc-and-upc-e-purpose-advantages.html`.

[10]Table 3 of Bronnenberg et al. (2009) lists 28 categories sampled from 1992 to 1995. 25 out of the 28 categories are still recorded in NielsenIQ data for 2013 to 2016. In Appendix B.1, we replicate the quantity share dispersion results from Bronnenberg et al. (2007, 2009) using our more recent sample, restricting the analysis to the same markets examined in their study. We find similar share dispersion across markets for the top two national brands.

[11]Bronnenberg and Dubé (2017) expand the analysis, confirming that their findings apply to a broader set of categories.

2.2 Geographic Features of Retailers

To ensure a relatively homogeneous sample of retailers, we include only food grocery stores and mass merchandisers with grocery revenue that exceeds the first quartile of grocery revenue among the grocery stores in our sample. Table 1 summarizes the retailer characteristics in terms of geographic coverage. We classify retailers into three groups: (i) *local* retailers that operate within a single market, (ii) *regional* retailers that operate in more than one market but within the same census region, and (iii) *national* retailers that span multiple census regions. On average, retailers operate in about eight markets. 58% are classified as regional, while approximately 20% are categorized as national.

Table 1: Retailer Summary Statistics

	Local Retailers	Regional Retailers	National Retailers	All Retailers
Number of Stores	21.71	86.92	330	117.2
	(28.91)	(90.47)	(401.4)	(214.0)
Number of Markets	1	5.426	24.75	8.008
	(0)	(4.364)	(24.92)	(13.90)
Number of Regions	1	1	2.567	1.292
	(0)	(0)	(0.679)	(0.677)
N	83	209	67	359

Notes: This table provides summary statistics for the retailers in terms of geographic coverage. The last row reports the number of retailer-years, distinguishing the retailer in different years as a different retailer. We define local retailers as those that operate exclusively within a single DMA. Regional retailers encompass retailers that span multiple DMAs but remain within the same census region, while national retailers are those that operate across multiple census regions. Standard deviations are in the parenthesis.

A key feature of the data is the geographic correlation between retailers and markets. Many regional retailers operate across multiple adjacent markets, and local markets are often dominated by the same regional retailers. This results in a positive spatial correlation between retailers and markets. To measure this, we estimate a spatial autoregressive model:

$$\mathbf{y} = \beta_0 + \beta_1 \mathbf{W}\mathbf{y} + \mathbf{X}\beta_2 + \gamma_r + \boldsymbol{\epsilon},$$

where $\mathbf{y}$ represents market-retailer level outcomes describing the geographic distribution of a retailer (e.g., an indicator for the retailer's presence in a market or the number of stores a retailer operates in a market). $\mathbf{X}$ includes control variables such as market-level population, $\mathbf{W}$ is a spatial weighting matrix, and γ_r captures retailer fixed effects. We construct two spatial weighting matrices: $\mathbf{W_1}$, based on market contiguity, and $\mathbf{W_2}$, based on inverse distance between markets.

The estimation results, presented in Table 2, show large coefficients on the spatial lag for both outcome variables. This indicates that a retailer's presence in one market (or the number of its stores) is strongly positively correlated with its presence (or store count) in nearby markets.

Table 2: Spatial Auto-regressive Model Results

	Dependent variable	
Coefficient on spatial lag (β_1)	Presence	Number of Stores
W_1: contiguity	0.77	2.13
	(0.008)	(0.023)
W_2: inverse distance	0.53	1.96
	(0.011)	(0.032)

Notes: This table reports the estimated coefficients on spatial lag variables of the spatial auto-regressive model. We balance the sample by filling in zero for markets where the retailer does not have any stores present. The sample is at the retailer-market level of year 2016. Standard errors are in the parentheses.

In Appendix A.2, we also provide an example from our data illustrating that many local markets are concentrated and dominated by the same regional retailers in terms of market share. This geographic correlation can create a positive relationship between the share distributions of retailers and the markets they serve, making it challenging to disentangle retailer effects from market effects. These characteristics of the data informs our choice of the variance decomposition model discussed in Section 4.

2.3 Key Outcome Measures

We construct six key outcome measures that capture market shares, product assortments, prices, and market concentration. These measures reflect various dimensions of brand presence, market power, competition, and retailer strategies.

(1) Quantity Share: This measure represents a parent brand b's share of total sales (in equivalent units) within category c, at retailer r, in market m, during year t.[12] It is calculated as:

$$\text{Quantity Share}_{bcrmt} = \frac{\text{Sales}_{bcrmt}}{\text{Total Sales}_{crmt}},$$

(2) Fraction of UPCs: This assortment measure captures the proportion of unique product codes (UPCs) associated with a specific brand at a retailer, in a market, and during a given

[12] Brand sales are converted into "equivalent units," which are scaled measures of unit sales provided by NielsenIQ to adjust for different package sizes across products.

year. It is defined as:

$$\text{Fraction of UPCs}_{bcrmt} = \frac{\mathrm{N}^{upc}_{bcrmt}}{\mathrm{N}^{upc}_{crmt}},$$

where N^{upc}_{bcrmt} is the number of UPCs for brand b and N^{upc}_{crmt} is the total number of UPCs in category c at retailer r, market m, and year t.

(3) Fraction of Products: Based on our product definition, this measure reflects the proportion of a brand's products within a retailer, category, market, and year. It is calculated as:

$$\text{Fraction of Products}_{bcrmt} = \frac{\mathrm{N}^{prod}_{bcrmt}}{\mathrm{N}^{prod}_{crmt}},$$

where $\mathrm{N}^{prod}_{bcrmt}$ is the number of products for brand b, and N^{prod}_{crmt} is the total number of products in category c at retailer r, market m, and year t.

(4) Brand Availability: This binary indicator reflects whether a parent brand is sold at a retailer-market during a specific year. It captures retailer decisions about stocking and shelving, which are influenced by both supply and demand-side considerations.

(5) Herfindahl–Hirschman Index (HHI): This measure captures category concentration within a retailer and market. Two versions of HHI are computed: one including private label shares and another excluding them. The HHI is defined as:

$$HHI_{crmt} = \sum_{b \in crmt} \text{Share}^2_{bcrmt},$$

where Share_{bcrmt} is the quantity share of brand b in category c, retailer r, market m, and year t.

(6) Price: Sales weighted average yearly price for each parent brand at a given retailer and market, expressed as price per unit.

2.4 Summary Statistics

Following Bronnenberg et al. (2007, 2009), we focus on the top two brands that hold the largest national quantity shares within each of the 25 categories that we consider. The top two national brands display a significant overlap in their consumers' demographic characteristics (Table A2), suggesting a similarity in their product portfolio and market positioning. Table 3 reports several descriptive statistics summarizing the market-level shares of these 50 brands in the same way as in Bronnenberg et al. (2007). We summarize each brand's national mean quantity share, its cross-market share dispersion, and its cross-market share range. The cross-market share dispersion is calculated as the between-market standard deviation in market quantity shares divided by the national quantity share of a given brand.

The upper panel of Table 3 summarizes the statistics for all 50 brands, and the lower panel separately reports summary statistics for the 27 national leading brands with coverage in all 199 markets and the 23 brands without coverage in all markets. We observe significant share dispersion across markets for all brands. Moreover, the range in a brand's national quantity share is similar for brands with full national coverage and for brands without national coverage, which suggests that the high dispersion is not merely driven by brands having no presence in some markets.

Table 3: Description of the National Top Brands' Shares across 25 categories

	Average Share	*Dispersion*	*Range*	*Minimum*	*Maximum*
All Leading Brands (N=50)					
mean	0.241	0.579	0.556	0.053	0.609
sd	0.2	0.342	0.248	0.08	0.256
Leading Brands with Coverage in All Markets (27)					
mean	0.291	0.37	0.519	0.102	0.621
sd	0.165	0.142	0.204	0.085	0.228
Leading Brands without Coverage in All Markets (23)					
mean	0.187	0.824	0.595	0	0.595
sd	0.22	0.347	0.287	0	0.287

Notes: The sample is at the brand-market-year level. Dispersion is calculated as between-markets standard deviation in local quantity shares divided by national share.

Panel A of Table 4 reports summary statistics of quantity shares and product assortments, separately for the top and second national brands. The statistics are summarized at the brand-retailer-market-year level. We observe a significant combined national share of approximately 50% for the two brands across categories. In most categories, there is a clear duopoly market structure, where the top two national brands hold significant quantity share and provide similar product portfolios. Panel B of Table 4 reports descriptive statistics for HHI measures at the category-retailer-market level. The HHI at the retailer-market level of 3535.64 is notably higher than the average national level HHI of 2516.07. Furthermore, the average HHI calculated without private labels is higher than when they are included, indicating that at some retailers, private labels hold a significant share.

Table 4: Summary Statistics of Share and Product Assortment Variables

	Mean	*SD*	*Median*	*Minimum*	*Maximum*
A. Shares and Product Assortments					
Top National Brands					
Quantity Share	0.35	0.238	0.307	0	1
Fraction of UPCs	0.059	0.05	0.048	0	0.368
Fraction of Products	0.104	0.094	0.074	0	0.5
Second National Brands					
Quantity Share	0.142	0.156	0.1	0	1
Fraction of UPCs	0.064	0.104	0.036	0	1
Fraction of Products	0.102	0.129	0.065	0	1
B. Category Concentration					
HHI (with private labels)	3535.64	1691.77	3106.97	485.74	9999.96
HHI (without private labels)	4511.39	2638.03	3604.68	485.74	10000

Notes: Panel A summarizes each brand's quantity share, fraction of UPCs, and fraction of products. The sample is at the brand-retailer-market-year level. Panel B summarizes HHI measures and the sample is at the category-retailer-market-year level.

3 Descriptive Evidence Regarding Share Dispersion

Bronnenberg et al. (2007) and Bronnenberg et al. (2009) provide evidence of (i) persistent variation of shares within markets and geographic dispersion across markets, (ii) important variation across markets in the identity of leading brands, and (iii) spatial dependence that spans multiple markets. They attribute this variation to persistent geographic differences in brand-level marketing investments that influence consumer preferences, and to historic entry patterns. These findings implicitly assume that retailers adapt their assortment to local preferences. In this section, we confirm the existence of this sort of retailer, but we also discuss that there is a different sort of retailer that features centralized category-captaincy contracts with specific brands, and that some of the observed share asymmetry can be attributed to the existence of these vertical contracts.

Our analysis is split in two parts. In Subsection 3.1, we compare geographic patterns in market shares and share dispersion across markets and across retailers. We also provide evidence of the existence not only of retailers that adapt their assortment to local preferences as in Bronnenberg et al. (2009), but also of retailers with centralized category-captaincy contracts with specific brands. Then, in Subsection 3.2, we provide evidence of persistent brand shares within retailers over time, which can partly explain the persistence in geographic share dispersion documented in Bronnenberg et al. (2007).

3.1 Share Dispersion across Markets versus across Retailers

Following the methodology of Bronnenberg et al. (2007), we calculate market-level shares for the top two brands, and plot a brand's market-level share against its national share. The results are presented in Figure 1, with Panel (a) depicting top national brands and Panel (c) focusing on second national brands. Each vertical line in the graphs represents a national brand, and the dots along these lines indicate the brand's market-level shares across local markets. In both panels, the magnitude of cross-market share dispersion mirrors Figure 1 of Bronnenberg et al. (2007). Moreover, the observations are symmetrically distributed around the 45-degree line: for top national brands, approximately half of the markets exhibit higher within-market quantity shares relative to their national averages, while the other half show substantially lower shares.[13]

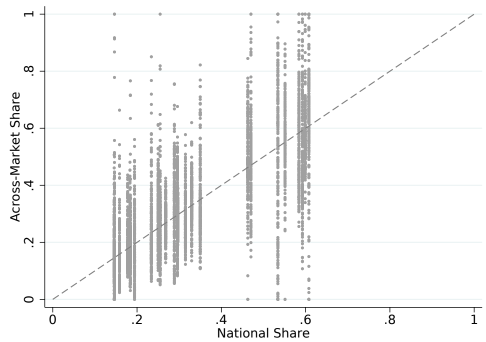
Next, we examine share dispersion *across retailers*. To do so, we calculate within-retailer share for the top two national brands, and plot the retailer-level share of the national brand against its national share. Panels (b) and (d) of Figure 1 present the results for the top national brands and the second national brands, respectively. Each vertical line corresponds to a national brand, while each dot along the line represents a retailer. Similar to Panels (a) and (c), there is a significant disparity between a brand's within-retailer share and its national share. This indicates that at approximately half of the retailers, the top national brands have a strong presence within the retailer, while in the other half, their share is notably lower than their national quantity share. This pattern also holds at national retailers (Figure B2).

Comparing dispersion *across retailers* versus *across markets*, we find that across-retailer dispersion is larger. The average IQR of quantity shares for the top national brands across retailers and categories is about 0.79, higher than that across markets and categories (0.64). Similarly, for the second national brands, the average IQR of quantity shares across retailers and categories is 0.51, also exceeding that across markets and categories (0.42).[14]

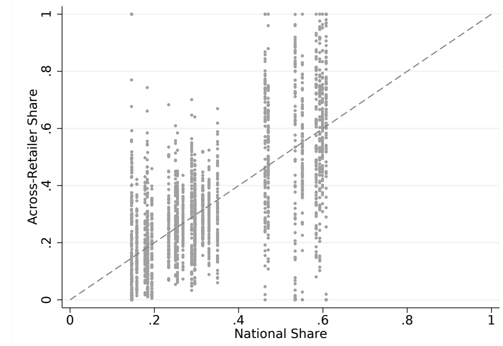
[13] As noted in Section 2, our sample covers more geographic markets and a more recent period of time than Bronnenberg et al. (2007). We replicate Panel (a) and (c) of Figure 1, using a sample of the same markets reported in their papers but from 2013-2016. The plot in Figure B1 confirms a comparable level of share dispersion across markets.

[14] To make the interquartile ranges comparable across markets and categories, the interquartile ranges are first divided by the median of quantity shares across markets within a category, then summarized across categories. We normalize IQR in a similar way for shares across retailers.

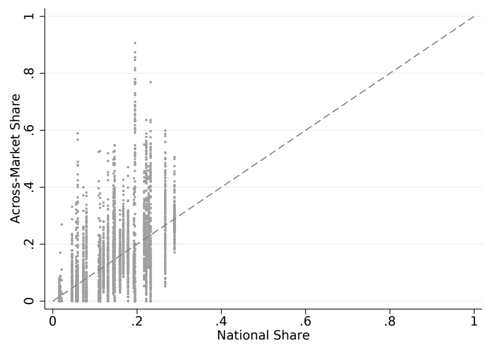
Figure 1: Across-Market share Dispersion versus Across-Retailer Share Dispersion



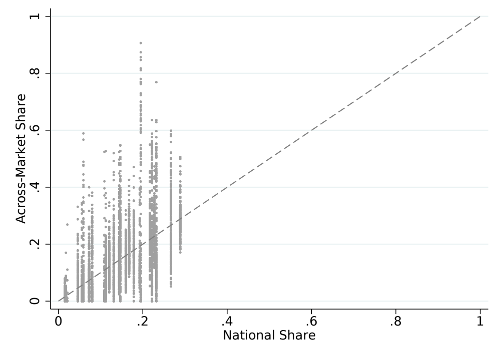
(a) Top National Brands: Across Markets



(b) Top National Brands: Across Retailers



(c) Second National Brands: Across Markets



(d) Second National Brands: Across Retailers

Notes: Each vertical line represents a brand, and each dot along the line denotes quantity share of the brand in a market (Panel a and c) or in a retailer (Panel b and d). The dashed line is a 45 degree straight line. For the top national brands, the average interquartile range across markets and categories is 0.64, while the average interquartile range across retailers and categories is 0.79. For the second national brands, the average interquartile range across markets and categories is 0.42, while the average interquartile range across retailers and categories is 0.51. Interquartile range is adjusted by median quantity share across markets (or retailers). These graphs use 2016 data.

This persistent geographic variation in brand performance can be driven in part by different retail distribution models and vertical contracting. Two models have emerged: decentralized systems and centralized frameworks (Chang and Harrington 2000). The former grants regional managers autonomy over assortment decisions, enabling local supplier partnerships that reflect market preferences. According to industry reports, retail cooperatives such as Alberstons are known to employ this approach, using a network of local divisions to accommodate regional demand.[15] Such systems often produce asymmetric brand performance across markets, as documented in studies of spatial share dispersion. For example, Bronnenberg et al. (2009) documents the importance of geography for explaining share vari-

[15] Albertson's decentralized model is explained in detail the following industry report: `https://www.thesnack.net/story/featured/safeway-albertsons-decentralization-who-is-affected/0070`.

ation using the top two brands in a particular category (Kraft and Unilever mayonnaise) and studying their shares in two different markets (Denver and Los Angeles). They show that Kraft dominates in Denver, while Unilever dominates in Los Angeles. They also show that shares of these two brands at Albertsons stores located in these two markets are characterized by the same patterns, implying that the retailer adapted their assortments to local preferences. See their Figure 1. Conversely, centralized distributors such as Costco use standardized procurement contracts and nationwide distribution centers.[16] Some retailers even rely on long-term vertical contracts such as category captaincy arrangements with particular brands resulting in consistent shelf presence, even in markets where consumer preferences might favor other brands. At these retailers, certain top national brands command high shares across all markets (as suggested by the dots above 45 degree line in Figure 1). However, we cannot directly relate these practices to specific examples in our data for reasons of confidentiality.

3.2 Persistence in Brand Shares within Retailer

Bronnenberg et al. (2009) also highlight enduring market share variation and the persistent dominance of national brands in certain geographic markets. In this section, we examine the within-retailer share persistence for the leading brand of each retailer. We calculate for each retailer and category the year-to-year transition probabilities, capturing switches between brands. The retailer leading brand in each category is defined as the brand with the highest share in the given category at the retailer. The sample is restricted to retailers present for at least two years and operating in multiple markets, resulting in a sample of 1,720 category-retailers. Our focus is on the transition probability within a retailer/category, conditional on the gap between the shares of the leading brand and the second brand within each retailer. Then retailers are divided into three equally sized groups based on the distribution of these share gaps across retailers that sold that category. On average, the thresholds of the share gap for these groups are 8%, 21%, and 40%.

Concentrating attention on the retailer-category pairs in which the leading brand has a significant share advantage in year t (largest share gap), the probability that in year $t+1$ a retailer's leading brand in a given category is the same as in t is 95%, roughly the same as the probability of persistence at the market level. For the middle share-gap group, the transition probability is 88%, while in the lowest share-gap group, it is 67%.

These results are intuitive. We would expect that in retailer/categories where the share-gap is narrow, there is strong likelihood that the leading brand flips from one period to

[16] https://ontheseams.substack.com/p/a-brief-primer-on-costcos-distribution.

the next. In contrast, when the gap is large, it is much more likely to persist, implying that within-retailer brand shares remain stable over time, with leading brands maintaining their position over time. This persistence, combined with the geographic distribution of retailers, may help to explain some of the persistence in market share variation documented in Bronnenberg et al. (2009). Of course, from this analysis we cannot determine whether this is driven by supply-side (contractual) or demand-side preferences.

Summary: Together, the descriptive evidence suggests that the share dispersion of top national brands across the retailer dimension is comparable to, or even greater than, the dispersion observed across the market dimension. Retailers vary significantly in how they sell and prioritize top national brands, with brand shares and dominance within a retailer remaining highly persistent over time. However, the dispersion analysis presented in this section is unconditional—that is, we document variation across retailers without accounting for market heterogeneity, and vice versa. Moreover, this section provides anecdotal evidence of two types of retailers – decentralized and centralized – that adapt their product assortment to match with those of the local markets in which they operate or not, respectively. However, the evidence cannot inform as to what fraction of the dispersion is due to the different types of retailers. To address both of these challenges, in Section 4, we employ a rigorous variance decomposition model to formally compare the contributions of market heterogeneity and retailer heterogeneity to share dispersion.

4 Variance Decomposition

4.1 AKM Decomposition Model

Our variance decomposition model draws on a body of research initiated by Abowd et al. (1999) (hereafter referred to as "AKM") in labor economics, which explores the sources of wage dispersion. AKM proposes a wage determination model that incorporates additive worker and firm fixed effects. This "two-way" fixed-effects model is expressed as:

$$y_{it} = \alpha_i + \psi_{\boldsymbol{J}(i,t)} + \nu_{it}, \tag{1}$$

where the log wage, y_{it}, of individual i in year t is specified as the sum of a worker component α_i, an establishment component $\psi_{J(i,t)}$, and an error term ν_{it}. Equation 1 can be estimated by ordinary least squares (OLS). After estimation, the variance of wages is decomposed in Equation 2, and the variance terms quantify how much each component contributes to the total variance of wages in the data.

$$\mathrm{Var}\left(y_{it}\right) = \mathrm{Var}\left(\alpha_i\right) + \mathrm{Var}\left(\psi_{\boldsymbol{J}(i,t)}\right) + 2\,\mathrm{Cov}\left(\alpha_i, \psi_{\boldsymbol{J}(i,t)}\right) + \mathrm{Var}\left(\nu_{it}\right). \tag{2}$$

We adapt the AKM model to measure the contributions of retailer effects and market effects in explaining variation in share, assortment, price, and dominance of the top two national brands across markets and retailers. The labor specification maps into our setting with brands taking the place of workers, and retailers and/or markets the place of firms.

An important distinction in our application lies in the fact that, on top of the two-way interaction between brands and retailers, we have a nesting structure characterized by the presence of retailers in multiple markets, and markets accommodating multiple retailers (see Subsection 2.2 and Appendix A.2). Furthermore, the descriptive evidence in Section 3 highlights that both the across-market and across-retailer share dispersion are brand specific and play a significant role in explaining the overall dispersion. Our approach views a brand as belonging to both the retailer and the market, and we fit the data using a "two-way match-effect" model with Brand × Retailer, and Brand × Market match-effects. Specifically, we estimate a model given by:

$$\mathrm{Y}_{brmt} = \alpha_b \times \gamma_r + \alpha_b \times \psi_{m(r)} + \epsilon_{brmt}, \tag{3}$$

where the unit of observation is at the level of brand-retailer-market-year. Y_{brmt} is the market outcome variable of brand b at retailer r in market m and year t. We analyze dispersion in the six outcomes listed in Subsection 2.3: (i) a brand's quantity share, (ii) brand availability, (iii) fraction of UPCs, (iv) fraction of products, (v) price, as well as (vi) dispersion in a category's concentration (HHI). $\alpha_b \times \gamma_r$ and $\alpha_b \times \psi_{m(r)}$ capture the match effects between brand and retailer, and brand and market, respectively. These two terms separately quantify the effects of retailer heterogeneity and market heterogeneity on the overall dispersion. In particular, the Brand × Retailer interaction terms flexibly measure the extent to which retailers shape a brand's market outcome dispersion, either due to differences in product distribution networks or vertical relationships across retailers. The Brand × Market effects capture taste differences toward a brand across markets. For HHI, we decompose the variation into Category × Retailer and Category × Market effects.

We then decompose the total variance in the market outcome variable into a component associated with the Brand × Retailer match effect, a component associated with the Brand × Market match effect, as well as a covariance term between the two match effects. The covariance term explicitly accounts for the correlation between retailers and markets in terms of their geographic and quantity share distributions. We consider the following decomposition:

$$Var(\mathrm{Y}_{brmt}) = \underbrace{Var(\alpha_b \times \gamma_r)}_{\text{Brand} \times \text{Retailer}} + \underbrace{Var(\alpha_b \times \psi_{m(r)})}_{\text{Brand} \times \text{Market}} + \underbrace{2Cov(\alpha_b \times \gamma_r, \alpha_b \times \psi_{m(r)})}_{\text{Covariance between Retailers and Markets}} + \underbrace{Var(\epsilon_{brmt})}_{\text{Residual}}. \quad (4)$$

The identification of the model follows the previous literature in labor, and is discussed in more details in Appendix C.1. A key requirement for identification is connectivity of the observations. In labor contexts, connectivity typically refers to workers transitioning between different firms (Abowd et al. 2004, Andrews et al. 2008). In our setting, this condition is satisfied by retailers operating across multiple markets and markets hosting multiple retailers. We apply the bias-correction method ("leave-one-out" algorithm) developed in Kline et al. (2020), which requires a "leave-one-out" sample to estimate the unbiased fixed effects. Appendix C.2 provides further details on this algorithm.

4.2 Results

Variance Decomposition. We estimate Equation 3 using the "leave-one-out" sample from the top two national brands in the 25 categories. For outcomes quantity share, fraction of UPCs, and fraction of products, we examine both the intensive margin and the extensive margin (i.e., brands not sold in a retailer-market-year are assigned a value of zero).

Table 5 reports the results of the AKM variance decomposition on the intensive-margin variables using the leave-one-out algorithm. The model fits the data well with relatively high R^2. The first column highlights the important role of the Brand $\times$ Retailer match effect, accounting for approximately 52% of the total variance in quantity shares, while the Brand $\times$ Market component contributes to around 20% of the overall variance. Columns (2) and (3) present variance decomposition findings for assortment depth variables, showing that the Brand $\times$ Retailer effects contribute to about 60% to 70% of the variance in assortments, while the Brand $\times$ Market component explains only 11%. Column (4) reports the results for prices. The Brand $\times$ Retailer component explains the majority of price variation (81.8%), while Brand $\times$ Market component only accounts for 3.4%. This finding is consistent with the uniform pricing phenomenon documented in DellaVigna and Gentzkow (2019).

Table 5: AKM Variance Decomposition—Intensive Margin

	(1) Quantity Share	(2) Fraction of UPCs	(3) Fraction of Products	(4) Price
Brand × Retailer	52.09	67.41	60.28	81.84
Brand × Market	20.49	11.4	11.12	3.44
2Cov	4.62	-3.79	-2.15	1.5
RMSE	0.073	0.028	0.036	0.04
R2	0.772	0.75	0.693	0.877

Notes: This table reports AKM variance decomposition results for the intensive-margin outcome variables. We report for each component the percentage of its variance relative to the total variance of the outcome variables.

Appendix C.3 provides additional results. Table C2 presents the AKM results for extensive-margin outcome variables, including a variance decomposition of brand availability. Table C3 reports the results for national retailers. The findings in both tables align closely with those reported in Table 5. We also estimate the AKM model using yearly samples spanning 2006 to 2016 and visualize the percentage of variation in fixed effects over time in Figure C2. Notably, the Brand × Retailer component consistently explains more than half of the overall share dispersion throughout this period. Finally, Table C4 presents the AKM decomposition results on the category concentration (HHI).[17] The variance of Category × Retailer effects accounts for 60% of the overall variance of HHI, while the Category × Market effect accounts for less than 10%.

Appendix C.4 presents a series of robustness checks, including the estimation of the variance decomposition model using OLS, substituting *parent_code* for *retailer_code* as the retailer identifier, and incorporating private label shares as a control variable.[18] The results remain robust across these specifications.

Comparison with Prior Work. Qualitatively, our findings are in line with those of Hwang and Thomadsen (2016), which uses a nested analysis of variance (ANOVA) and a components-of-variance (COV) to decompose shares of the national top two brands. However, some of their quantitative findings differ from ours. Their findings suggest that the market-level variation accounts for approximately 30% of the weekly share variation across stores, with retailer-level and store-level variations accounting for an additional 13% and 5%, respectively. In contrast, we find that 52% of the share variation is attributable to the

[17] For this analysis, we decompose the total variance in HHI at the retailer-market-year level into Category × Retailer, Category × Market and the covariance terms.

[18] In NielsenIQ data, *parent_code* generally indicates the parent company that owns a retailer, and the *retailer_code* indicates the retailer itself.

Brand × Retailer component, while the Brand × Market component accounts for 20% of total variation.

The contrast in results could be due to differences either in data sources, samples, or models. We provide a detailed comparison between our study and Hwang and Thomadsen (2016) in Appendix D. As a first step, we reconstructed the samples used in their study, creating a comparable sample from IRI 2004–2005, which is the dataset used in Hwang and Thomadsen (2016), and a sample from NielsenIQ 2005, which is the sample used in Hwang et al. (2010). We then estimated the AKM, ANOVA, and COV models using these samples alongside our primary sample from NielsenIQ 2013–2016.

The comparison highlights that the observed divergences are primarily driven by differences in the data sources, particularly in geographic coverage and sample connectivity. NielsenIQ data offers broader geographic coverage and features a higher prevalence of multi-market retailers and multi-retailer markets compared to IRI data. To evaluate sample connectivity, we calculate the maximum leverage—a measure recommended by Kline et al. (2020)—for both datasets.[19] The maximum leverage is 0.72 for IRI data and 0.5 for NielsenIQ data (2013–2016), indicating that NielsenIQ data (2013–2016) exhibits better connectivity properties. This enhanced coverage facilitates better identification of variance decomposition models. Moreover, the retailer network overlap in the NielsenIQ data is better than that of IRI. In the IRI data, either different retailers are surveyed, or the econometric results are disproportionately influenced by a few retailers that may not be representative (outliers). As a result, the identification condition for the "leave-one-out" algorithm (Kline et al. 2020) is better satisfied in the NielsenIQ data compared to the IRI data.

Summary To summarize, the findings in this section consistently show that the substantial variation in the top national brands' quantity shares, product assortments, prices, and category concentration across retailers and markets is primarily driven by the heterogeneous brand-retailer match effects. These effects arise from differences in how retailers sell and prioritize specific brands through assortment, availability, and pricing, potentially influenced by brand-retailer vertical relationships. Additional analysis on this topic is presented in Section 5.

5 Market vs Retailer Heterogeneity

In this section, we revisit the findings from Bronnenberg et al. (2007) and Bronnenberg et al. (2009), while taking into account the possible effects of long-term vertical contracts on retail

[19]See Appendix C.1 for a detailed illustration of connectivity, and Appendix C.2 for the leverage formula.

assortments and pricing.

To do so, we perform a more in-depth analysis of the fixed effects estimated from the AKM model. In Subsection 5.1, we extract these fixed effects and analyze their variation, providing further context to the AKM decomposition results presented in Subsection 4.2 and refining our interpretation of these factors. Then in Subsection 5.2, we examine the relationship between these components and other relevant variables, such as the distance to the city of origin and retailer characteristics. This approach of analyzing fixed effect components has been employed in other empirical studies, including Finkelstein et al. (2016) and Chetty and Hendren (2018).[20] Our results suggest that the market demand heterogeneity observed in previous research also captures substantial retailer heterogeneity, with nearly half of the demand variation attributable to retailer-specific factors. Additionally, this retailer heterogeneity appears to be driven by long-term vertical contracts that offer preferential treatment to one dominant brand over others.

5.1 Brand $\times$ Market and Brand $\times$ Retailer Fixed Effects

From the estimated AKM model (intensive margin), we extract the Brand $\times$ Market and Brand $\times$ Retailer fixed effects for share, price, and product assortments. These fixed effects capture conditional variation in market outcomes, either across markets (conditional on retailers) or across retailers (conditional on markets). These components of market outcome variables reflect underlying market conditions and structures. For instance, the Brand $\times$ Market fixed effects for shares capture market-specific heterogeneity in share dispersion, such as differences in demand across markets. Meanwhile, the Brand $\times$ Retailer fixed effects capture retailer-specific heterogeneity, including differences in vertical relationships. Understanding these market conditions and structures is not only of intrinsic interest but also important for informing policy implications, such as those related to competition.

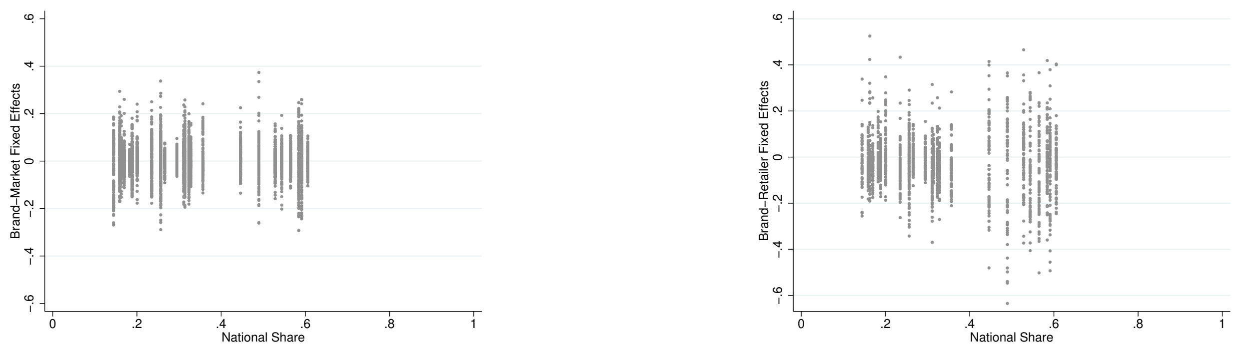
Figure 2 plots the Brand $\times$ Market and Brand $\times$ Retailer fixed effects estimated from the AKM model with quantity share as the outcome variable (Column (1) in Table 5). Similar to Figure 1, we plot the fixed effects of the top national brand in each category against the national market share of that brand. Each vertical line represents a category, and each dot represents either a market (in Panel (a)) or a retailer (in Panel (b)). Notably, even after simultaneously accounting for market and retailer heterogeneity in the model, the dispersion in Brand $\times$ Retailer fixed effects (Panel (b)) is greater than the dispersion in

[20] Finkelstein et al. (2016) decompose variation in insurance utilization into patient and place components to analyze the underlying mechanisms driving geographic differences in healthcare. Similarly, Chetty and Hendren (2018) break down variation in upward mobility across counties into county effects and parent ranking effects, using these estimates to identify area-level factors that influence economic mobility.

Brand × Market fixed effects (Panel (a)). Specifically, the IQR of Brand × Retailer fixed effects is 0.15—nearly twice the IQR of Brand × Market fixed effects, which is 0.08. This indicates that brand-retailer heterogeneity plays a more important role than brand-market heterogeneity in explaining share dispersion. This result contrasts with the patterns observed in Figure 1, where the unconditional share dispersion across markets appears nearly identical to that across retailers.

Figure 2: Brand × Market Fixed Effects versus Brand × Retailer Fixed Effects

(a) Brand × Market Fixed Effects (b) Brand × Retailer Fixed Effects



Notes: These figures plot Brand × Market fixed effects and Brand × Retailer fixed effects estimated from AKM model, with quantity share (intensive margin) as the outcome variable.

There are also meaningful correlations between within-retailer asymmetry in quantity share, assortment, and price. To demonstrate this, we calculate the differences in Brand × Retailer fixed effects for share, price, and assortment (fraction of UPCs) between the top two national brands within each retailer-category. These differences are standardized by their respective standard deviations. We then regress the difference in share fixed effects on the differences in price and assortment fixed effects. The results, presented in Table 6, show a positive coefficient for assortment differences and a negative coefficient for price differences, indicating that greater brand advantage in shares within-retailer is associated with larger product assortments and lower prices.

Table 6: Fixed Effects Correlation Regressions

	Diff Share FEs	Diff Share FEs	Diff Share FEs
Diff Assortment FEs	0.0415***		0.0477***
	(0.0131)		(0.0132)
Diff Price FEs		-0.0309**	-0.0378***
		(0.0123)	(0.0124)
Observations	7,156	7,156	7,156
R-squared	0.035	0.035	0.036
Year FEs	Yes	Yes	Yes
Category FEs	Yes	Yes	Yes
Mean Outcome	0.09	0.09	0.09

Notes: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

In summary, these results indicate that both the Brand × Market and the Brand × Retailer components reflect meaningful and interesting underlying economic primitives. Therefore, in the next section, we utilize these fixed effects in a series of analyses to offer insights about primitives such as market demand and long-term vertical relationships.

5.2 Insights on market demand and long-term vertical contracts

Persistence in Brand Asymmetry within Retailer. Many previous studies suggest that vertical relationships between retailers and manufacturers are typically long-term (Luo 2023). In this section, we examine the stability of within-retailer brand asymmetry over time. This analysis complements the examination of transition probabilities performed in Section 3.

To do so, we estimate an AR(2) regression on the estimated Brand × Retailer fixed effects to examine their correlation over the years from 2006 - 2016. The regression controls for retailer and category fixed effects. The results, presented in Table 7, show an elevated and significant coefficient for the one-year lag (0.839), indicating a strong persistence in brand asymmetry within retailers over time. This aligns with industry anecdotes indicating that vertical contracts are often long-term.

Table 7: AR(2) Regression on Brand × Retailer FEs

	Brand × Retailer FE (t)
Brand × Retailer FE (t-1)	0.839***
	(0.00771)
Brand × Retailer FE (t-2)	0.0180**
	(0.00755)
Constant	0.0126***
	(0.00372)
Observations	16,204

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Spatial Distribution of Brand Shares. Bronnenberg et al. (2009) demonstrate substantial and significant demand heterogeneity across markets. In this section, we confirm the presence of similar heterogeneity in our sample. However, we show that this heterogeneity is reduced by half after accounting for retailer heterogeneity.

We start by examining the persistence of a brand's market share across different markets and assessing the impact of distance from its city of origin on market share. As shown by Bronnenberg et al. (2009), a brand's proximity to its city of origin is positively correlated with its current market share. The city of origin represents the geographic area where a brand has operated the longest, while markets farther from this area typically correspond to regions the brand entered more recently. Brands systematically achieve higher market shares in regions they entered early — a phenomenon known as the "city of origin effect."

To examine demand heterogeneity in relation to the city of origin, we replicate Figure 3 from Bronnenberg et al. (2009) using the Brand × Market fixed effects from our AKM model, which captures market demand heterogeneity. Specifically, we identify the city of origin for each brand in our sample and, following a specification similar to Bronnenberg et al. (2009), estimate Equation 5 using data for the top national brands at national retailers:[21]

$$\Gamma_{bcm} = \sum_{k=0}^{10} \delta_k \mathbf{1}(\text{Dist}^k_{bcm}) + \gamma_c + \epsilon_{bcm}, \tag{5}$$

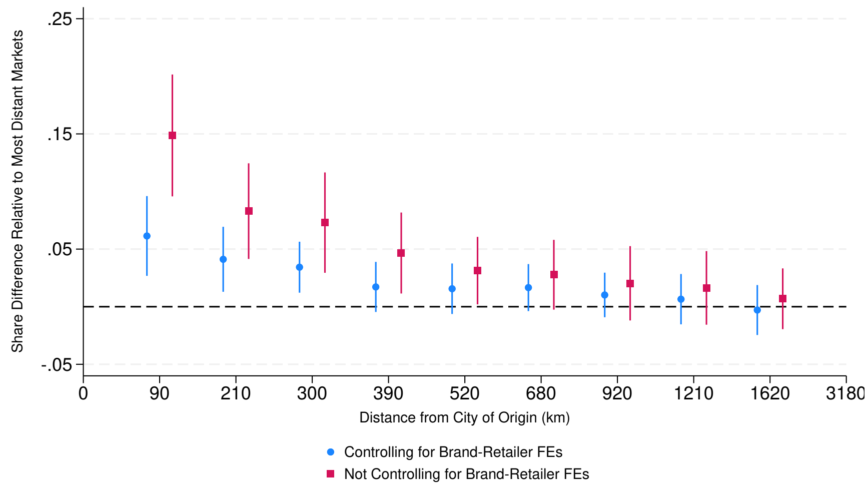
where Γ_{bcm} denotes Brand × Market FEs extracted from the model. We extract two versions of these fixed effects: Brand × Market FEs from the AKM model estimated *with* and *without* the Brand × Retailer FEs. γ_c are category fixed effects. We compute the distance between

[21]To ensure sufficient geographic coverage, we focus our analysis on 17 national retailers in 2016.

a market and a brand's city of origin, which ranges from 0 to 4,368 kilometers, and partition them into 10 intervals. For each interval, we create a dummy variable $\mathbf{1}(\text{Dist}^k_{bcm})$, which takes on a value of one if the distance from market m to the brand's city of origin falls into interval k. δ_{10}, which corresponds to the effect at distances over 1,620 kilometers, is normalized to zero.

The coefficients of interest, δ_k, capture the effect of distance from a brand's city of origin on its market share dominance. We visualize these coefficients in Figure 3, where the square dots represent δ_k derived from Γ_{bcm} without controlling for Brand $\times$ Retailer fixed effects, and the circle dots represent δ_k obtained from Γ_{bcm} with Brand $\times$ Retailer controls included.

Figure 3: Effect of Distance from City of Origin on Market Share Dominance



Notes: These figures display the estimated δ_k coefficients from Equation 5, with whiskers representing 95 percent confidence intervals. To improve visualization, the circle and square markers for the same distance bins are intentionally offset and not perfectly aligned.

The square dots (the δ_k corresponding to Brand $\times$ Market FEs without controlling for Brand $\times$ Retailer FEs) show that a brand's market share decreases as the distance from its city of origin increases. The 95 percent confidence bands indicate that the effect of distance on market share dominance diminishes to zero at approximately 920 kilometers. This finding aligns with Bronnenberg et al. (2009), which also emphasizes that a national brand's market share decreases with greater distance from its origin.[22] This pattern highlights the presence of significant market heterogeneity in demand. However, once we control for the Brand $\times$ Retailer FEs, the coefficients for distance are reduced by approximately half (indicated by the circle dots). This substantial reduction suggests that a large portion of the variation

[22]While the overall pattern is consistent with their results, the specifics differ since the top brands in our sample are not identical to those in their study.

in Brand-Market fixed effects is explained by Brand-Retailer effects that capture retailer heterogeneity.[23]

Next, we adapt Equation 5 to investigate the relationship between within-retailer brand asymmetry and the distance to the city of origin. Rather than using Brand-Market fixed effects, we utilize Brand $\times$ Retailer fixed effects derived from the AKM model, both with and without controls for Brand-Market fixed effects.

To calculate the distance between a retailer and a brand's city of origin, we first identify the "retailer center," defined as the market where the retailer generates the highest revenue. We then measure the distance between each retailer center and the brand's city of origin.[24] Using this distance, we regress the Brand $\times$ Retailer fixed effects on distance bins, as specified in Equation 6. This regression is estimated at the brand-retailer level in 2016, with the coefficients (δ_k) identified by variation across retailers for the same brand.

$$\Gamma_{bcr} = \sum_{k=0}^{10} \delta_k \mathbf{1}(\text{Dist}_{bcr}^k) + \gamma_c + \epsilon_{bcr}. \tag{6}$$

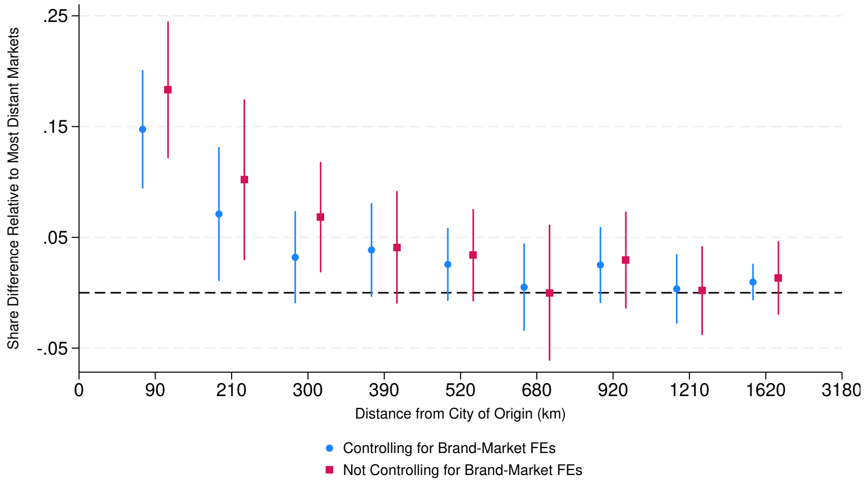
We plot δ_k in Figure 4, where square dots represent δ_k from the regression of Γ_{bcr} retrieved without Brand-Market controls and circle dots represent those with these controls. The square dots indicate that retailers with "core markets" closer to a brand's city of origin exhibit stronger preferential treatment for that brand compared to those farther away. This finding reflects brands' tendencies to establish relationships earlier and sustain them longer with retailers in or near their origin markets. Notably, the downward-sloping pattern of the coefficients remains virtually unchanged even after accounting for Brand $\times$ Market FEs that capture market heterogeneity, as shown by the circle dots.[25]

[23] Additionally, we examine this relationship between brand-market heterogeneity and distance over the years, by regressing the Brand $\times$ Market FEs on the distance to city of origin in each year. We find that the spatial share variation remain stable over time, and the retailer heterogeneity consistently absorbs half of the demand heterogeneity. More details are provided in Appendix E.

[24] The distance ranges from 0 to 4,428 kilometers.

[25] We also examine the relationship between within-retailer brand asymmetry and distance over the years, by regressing the Brand $\times$ Retailer FEs on the distance to brand's city of origin in each year. The retailer heterogeneity remains consistent, regardless of whether market demand heterogeneity is accounted for. Details are reported in Appendix E.

Figure 4: Effect of Distance from City of Origin on Brand Asymmetry Within-Retailer



Notes: These figures display the estimated δ_k coefficients from Equation 5, with whiskers representing 95 percent confidence intervals. To improve visualization, the circle and square markers for the same distance bins are intentionally offset and not perfectly aligned.

Retailer Heterogeneity and Vertical Contracts. Figure 2 highlights significant differences in how retailers prioritize the top national brands. These patterns, as noted in previous research, may stem from long-term vertical arrangements between retailers and major manufacturers. Such arrangements include practices like category captains (Viswanathan et al. 2020, Zhu 2024), vendor allowances (Hristakeva 2022b,a), exclusive dealing (Asker 2016), and vertical rebates (Conlon and Mortimer 2021), which often favor one large brand over its competitors.

In this section, we use the extracted within-retailer brand asymmetry measure (Brand $\times$ Retailer fixed effects) to test predictions from prior research on vertical contracts. Specifically, we examine the relationships between our brand asymmetry measure and retailer or category characteristics that prior literature has linked to within-retailer market power through vertical relationships. We assess whether greater dominance of the top brands is associated with these characteristics, offering insights into the role of vertical relationships in shaping brand performance within retailers.

Building on previous studies, we construct the following retailer-category-specific measures: (1) Product variety, as measured by the number of UPCs available within a retailer, following Draganska and Jain (2005); (2) Retailer size, which includes metrics such as the number of markets served and total retailer revenue, capturing the scale of the retailer's operations; (3) Private label segment, measured using the quantity share of private label products and the fraction of private label UPCs relative to the total number of UPCs;[26] (4)

[26]Prior research has demonstrated that private labels can act as a strategic negotiation tool for retailers in their

Product differentiation, proxied by the overlap in sizes between the top two brands, where a greater overlap (i.e., a higher number of common sizes offered by both brands) indicates lower product differentiation.

As the outcome variable, we calculate the absolute difference in estimated Brand $\times$ Retailer fixed effects between the top two brands within each retailer:

$$Y_{r,t} = \sum_{c=1}^{C} s_{r,c,t} \left| FE_{r,c,t,1} - FE_{r,c,t,2} \right|,$$

where $s_{r,c,t}$ is the revenue share of category c in retailer-year (r, t), $FE_{r,c,t,1}$ is the retailer-brand-year fixed effect for the top national brand, and $FE_{r,c,t,2}$ represents the same for the second national brand. A larger $Y_{r,t}$ indicates a more dominant position for one of the top two brands within a retailer. We analyze the relationship between $Y_{r,t}$ and retailer-category characteristics by estimating regressions that include year fixed effects as controls.

The results are presented in Table 8. Column (1) focuses on product variety and retailer size. It shows that product variety, measured by the number of UPCs, is negatively correlated with brand asymmetry, while retailer revenue is positively correlated. This suggests that, conditional on product variety, larger retailers exhibit stronger brand asymmetry. This finding aligns with the idea that these retailers can negotiate contracts that solidify the dominance of major manufacturers, often securing favorable terms from them (Scott Morton and Zettelmeyer 2004). The negative coefficient for the number of UPCs implies that retailers offering a wider variety of products tend to provide less preferential treatment to top brands. In the remaining columns, we incorporate other retailer characteristics into the analysis while controlling for the retailer size and variety variables. Columns (2) and (3) show a significantly negative correlation between brand dominance and the size of the retailer's private label segment. This suggests an association between larger private label segments and reduced prominence of national top brands in retailers' product assortments. These findings are consistent with industry observations and prior research, which emphasize the role of private labels as a bargaining tool. By leveraging private labels, retailers can strengthen their negotiating power against major national brands (Scott Morton and Zettelmeyer 2004, Ellickson et al. 2018). Column (4) reports a positive coefficient for the product differentiation measure (overlap in sizes), which suggests that higher brand asymmetry within a retailer is linked to less product differentiation between the top two brands (more overlapping sizes).

vertical relationships (Scott Morton and Zettelmeyer 2004, Ellickson et al. 2018).

Table 8: Relationship between Brand Asymmetry and Retailer Characteristics

	(1)	(2)	(3)	(4)	(5)
Log Number of UPCs	-0.0206***	-0.0208***	-0.0212***	-0.0142**	-0.0156***
	(0.00690)	(0.00533)	(0.00455)	(0.00658)	(0.00444)
Log Number of Markets	-0.00413	-0.00367	-0.00363	-0.00519	-0.00471
	(0.00387)	(0.00376)	(0.00352)	(0.00375)	(0.00353)
Log Retailer Revenue	0.0155**	0.0162***	0.0170***	0.0135**	0.0151***
	(0.00642)	(0.00507)	(0.00415)	(0.00609)	(0.00397)
Private Label Shares		-10.44***			5.318
		(3.082)			(4.203)
Fraction of Private Label UPCs			-0.239***		-0.279***
			(0.0525)		(0.0785)
Overlap in Sizes				2.532***	2.202***
				(0.534)	(0.496)
Observations	291	291	291	291	291
R-squared	0.199	0.283	0.345	0.351	0.459
Year FEs	Yes	Yes	Yes	Yes	Yes
Mean Outcome	0.101	0.101	0.101	0.101	0.101

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Additionally, there is a negative relationship between transition probability (as documented in Subsection 3.2) and brand asymmetry within retailers. A regression of transition probability on the absolute difference in Brand × Retailer fixed effects ($Y_{r,t} = \sum_{c=1}^{C} s_{r,c,t}|FE_{r,c,t,1} - FE_{r,c,t,2}|$) yields a coefficient of -0.164. This indicates that retailers with higher brand asymmetry and brand dominance are also more likely to maintain consistency with the dominant leading brand.

6 Conclusion

In this paper, we document significant asymmetries in the quantity shares of top national brands across retailers. Similar patterns are observed in product assortments, availability, prices, and category concentration. Even among some national retailers operating across multiple markets, the distribution of top national brand shares remains notably heterogeneous across retailers. This brand asymmetry is especially pronounced in larger retailers (in terms of revenue), in retailers with smaller private label segments and less variety, and in categories where national brands are less differentiated. Equally noteworthy is the persistence of within-retailer share concentration over time and across markets at some retailers, suggesting a sustained competitive advantage for certain brands within specific retailers.

Our findings highlight the role retailers play in shaping and sustaining the competitive landscape of retail grocery industry. The substantial market demand heterogeneity, or growing brand loyalty over time, as documented in previous research, can be partly attributed to retailers' preferential treatment of certain brands. This preferential treatment is often enabled and maintained by long-term vertical relationships. We present suggestive evidence that some retailers designate a "main supplier" for a category across multiple markets. These main suppliers benefit from widespread distribution across all establishments of the retailer, making it a dominant strategy for firms to secure a favorable position within a retailer. This favorable position is often achieved through vendor allowances, captaincy payments, or similar arrangements, typically favoring large manufacturers with substantial resources and established market presence. The intensified competition among upstream manufacturers can, in turn, enhance retailers' bargaining power. This interplay between retailer strategies, supplier competition, and market power of national top brands highlights the importance of further examining their implications for market efficiency and consumer welfare.

References

Abowd JM, Kramarz F, Lengermann P, Pérez-Duarte S (2004) Are good workers employed by good firms? A test of a simple assortative matching model for france and the united states. *Unpublished Manuscript* .

Abowd JM, Kramarz F, Margolis DN (1999) High wage workers and high wage firms. *Econometrica* 67(2):251–333.

Aghion P, Bolton P (1987) Contracts as a barrier to entry. *The American Economic Review* 77(3):388–401.

Andrews MJ, Gill L, Schank T, Upward R (2008) High wage workers and low wage firms: Negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171(3):673–697.

Asker J (2016) Diagnosing foreclosure due to exclusive dealing. *The Journal of Industrial Economics* 64(3):375–410.

Asker J, Bar-Isaac H (2014) Raising retailers' profits: On vertical practices and the exclusion of rivals. *American Economic Review* 104(2):672–686.

Bonnet C, Dubois P (2010) Inference on vertical contracts between manufacturers and retailers allowing for nonlinear pricing and resale price maintenance. *The RAND Journal of Economics* 41(1):139–164.

Bronnenberg BJ, Dhar SK, Dubé JP (2007) Consumer packaged goods in the United States: National brands, local branding. *Journal of Marketing Research* 44(1):4–13.

Bronnenberg BJ, Dhar SK, Dubé JPH (2009) Brand history, geography, and the persistence of brand shares. *Journal of Political Economy* 117(1):87–115.

Bronnenberg BJ, Dubé JP (2017) The formation of consumer brand preferences. *Annual Review of Economics* 9(1):353–382.

Bronnenberg BJ, Dubé JPH, Gentzkow M (2012) The evolution of brand preferences: Evidence from consumer migration. *American Economic Review* 102(6):2472–2508.

Chang MH, Harrington JE (2000) Centralization vs. decentralization in a multi-unit organization: A computational model of a retail chain as a multi-agent adaptive system. *Management Science* 46(11):1427–1440.

Chetty R, Hendren N (2018) The impacts of neighborhoods on intergenerational mobility II: County-level estimates. *The Quarterly Journal of Economics* 133(3):1163–1228.

Clark R, Horstmann I, Houde JF (2024) Hub-and-spoke cartels: Theory and evidence from the grocery industry. *American Economic Review* 114(3):783–814.

Conlon CT, Mortimer JH (2021) Efficiency and foreclosure effects of vertical rebates: Empirical evidence. *Journal of Political Economy* 129(12):3357–3404.

Dafny L, Hendel I, Wilson N (2015) Narrow networks on the health insurance exchanges: What do they look like and how do they affect pricing? A case study of Texas. *American Economic Review* 105:110–114.

DellaVigna S, Gentzkow M (2019) Uniform pricing in US retail chains. *The Quarterly Journal of Economics* 134(4):2011–2084.

Draganska M, Jain DC (2005) Product-line length as a competitive tool. *Journal of Economics & Management Strategy* 14(1):1–28.

Draganska M, Mazzeo M, Seim K (2009) Beyond plain vanilla: Modeling joint product assortment and pricing decisions. *Quantitative Marketing and Economics* 7(1):105–146.

Dubé JP, Hitsch GJ, Rossi PE (2009) Do switching costs make markets less competitive? *Journal of Marketing research* 46(4):435–445.

Dubé JP, Hitsch GJ, Rossi PE (2010) State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics* 41(3):417–445.

Dubois P, Bonnet C, et al. (2015) Identifying two part tariff contracts with buyer power: Empirical estimation on food retailing. Technical report, CEPR Discussion Papers.

Dubé JP, Rossi PE (2019) *Handbook of the Economics of Marketing* (North Holland).

Eizenberg A (2014) Upstream innovation and product variety in the U.S. home PC market. *The Review of Economic Studies* 81(3):1003–1045.

Ellickson PB, Kong P, Lovett MJ (2018) Private labels and retailer profitability: Bilateral bargaining in the grocery channel. *Available at SSRN 3045372* .

Fan Y (2013) Ownership consolidation and product characteristics: A study of the US daily newspaper market. *American Economic Review* 103(5):1598–1628, URL http://dx.doi.org/10.1257/aer.103.5.1598.

Fan Y, Yang C (2020) Competition, product proliferation, and welfare: A study of the US smartphone market. *American Economic Journal: Microeconomics* 12(2):99–134, URL http://dx.doi.org/10.1257/mic.20180182.

Fan Y, Yang C (2024) Estimating discrete games with many firms and many decisions: An application to merger and product variety, NBER working paper, 30146, accepted, *Journal of Political Economy*.

Finkelstein A, Gentzkow M, Williams H (2016) Sources of geographic variation in health care: Evidence from patient migration. *The quarterly journal of economics* 131(4):1681–1726.

Hitsch GJ, Hortacsu A, Lin X (2019) Prices and promotions in US retail markets: Evidence from big data. Technical report, National Bureau of Economic Research.

Ho K, Ho J, Mortimer JH (2012) The use of full-line forcing contracts in the video rental industry. *American Economic Review* 102:686–719.

Ho K, Lee RS (2019) Equilibrium provider networks: Bargaining and exclusion in health care markets. *American Economic Review* 109(2):473–522.

Ho K, Lee RS (2024) Contracting over pharmaceutical formularies and rebates. Working paper, National Bureau of Economic Research.

Hollenbeck B, Hristakeva S, Uetake K (2024) Retailer price competition and assortment differentiation: Evidence from entry lotteries. Technical report, Working Paper.

Hristakeva S (2022a) Determinants of channel profitability: Retailers' control over product selections as contracting leverage. *Marketing Science* 41(2):315–335.

Hristakeva S (2022b) Vertical contracts with endogenous product selection: An empirical analysis of vendor allowance contracts. *Journal of Political Economy* 130(12):3202–3252.

Hwang M, Bronnenberg BJ, Thomadsen R (2010) An empirical analysis of assortment similarities across US supermarkets. *Marketing Science* 29(5):858–879.

Hwang M, Thomadsen R (2016) How point-of-sale marketing mix impacts national-brand purchase shares. *Management Science* 62(2):571–590.

Kline P, Saggio R, Sølvsten M (2020) Leave-out estimation of variance components. *Econometrica* 88(5):1859–1898.

Luco F, Marshall G (2020) The competitive impact of vertical integration by multiproduct firms. *American Economic Review* 110(7):2041–2064.

Luo B (2023) Manufacturer–retailer relationships and the distribution of new products. *Marketing Science* .

Marx LM, Shaffer G (2007) Upfront payments and exclusion in downstream markets. *The RAND Journal of Economics* 38(3):823–843.

McGahan AM, Porter ME (1997) How much does industry matter, really? *Strategic management journal* 18(S1):15–30.

Morning Star Analysis Report (2020) Danone's Narrow Moat Should Help the Group Withstand the Coronavirus. URL https://www.morningstar.com/stocks/market-overlooks-danones-prospects-margin-expansion, [Online; accessed 5. July. 2022].

Prager E (2020) Healthcare demand under simple prices: Evidence from tiered hospital networks. *American Economic Journal: Applied Economics* 12:196–223.

Progressive Grocer (2011) Category Captain: Yogurt: General Mills. URL https://progressivegrocer.com/category-captain-yogurt-general-mills, [Online; accessed 24. Dec. 2022].

Progressive Grocer (2015a) Winning Game Plans. URL https://progressivegrocer.com/category-captain-yogurt-general-mills, [Online; accessed 3. Sep. 2022].

Progressive Grocer (2015b) Winnng Game Plans. URL https://progressivegrocer.com/winnng-game-plans, [Online; accessed 24. Dec. 2020].

Progressive Grocer (2018) Category Captains Awards. URL https://progressivegrocer.com/2018-category-captains-awards-winners, [Online; accessed 28. Feb. 2021].

Rasmusen EB, Ramseyer JM, Wiley JS (1991) Naked exclusion. *The American Economic Review* 81(5):1137–1145.

Rossi-Hansberg E, Sarte PD, Trachter N (2021) Diverging trends in national and local concentration. *NBER Macroeconomics Annual* 35(1):115–150.

Scott Morton F, Zettelmeyer F (2004) The strategic positioning of store brands in retailer–manufacturer negotiations. *Review of industrial organization* 24(2):161–194.

Segal IR, Whinston MD (2000) Naked exclusion: Comment. *The American Economic Review* 90(1):296–309.

Sorensen AT (2003) Insurer-hospital bargaining: Negotiated discounts in post-deregulation connecticut. *The Journal of Industrial Economics* 51:469–490.

Sutton J (1991) *Sunk costs and market structure: Price competition, advertising, and the evolution of concentration* (MIT press).

Villas-Boas SB (2007) Vertical relationships between manufacturers and retailers: Inference with limited data. *The Review of Economic Studies* 74(2):625–652.

Viswanathan M, Narasimhan O, John G (2020) Economic impact of category captaincy: an examination of assortments and prices. *Marketing Science* .

Zhu X (2024) Inference and impact of category captaincy. *Management Science* .

Appendix

A Data

A.1 National Top Two Brands

Table A1 shows the list of categories in our sample, along with the top two national brands within each category. The table also provides information on the average quantity share of the brand over the sample period, the number of products in 2016, and the average price. The majority of the top two brands exhibit similarities in their national market share, product offerings, and average price.

Table A1: National Top 2 Brands in Each Category

Cateogy	Brand	Share	# of Products	Price
PIZZA SAUCE	Mizkan America, Inc.	0.16	4	1.75
PIZZA SAUCE	Riviana Foods Inc.	0.12	3	2
SAUCES - DIPPING	T. Marzetti Company	0.49	17	3.87
SAUCES - DIPPING	Litehouse Inc.	0.19	16	3.43
PICKLES - SWEET	Pinnacle Foods Group LLC	0.26	29	2.96
PICKLES - SWEET	Mount Olive Pickle Company Inc.	0.24	13	2.78
PICKLES - DILL	Pinnacle Foods Group LLC	0.24	28	3.31
PICKLES - DILL	Mount Olive Pickle Company Inc.	0.22	18	2.94
MAYONNAISE	Unilever Bestfoods North America	0.61	96	4.06
MAYONNAISE	Kraft Heinz Foods Company	0.15	47	3.49
SALAD DRESSING	Kraft Heinz Foods Company	0.18	93	2.32
SALAD DRESSING	The Hvr Company	0.18	44	4.08
MUSTARD	McCormick Company, Inc.	0.33	35	2.2
MUSTARD	Kraft Heinz Foods Company	0.11	38	3.15
RICE	Riviana Foods Inc.	0.19	129	4.58
RICE	Goya Foods, Inc.	0.11	34	6.44
CEREAL - READY TO EAT	The Kellogg Company	0.3	948	3.52
CEREAL - READY TO EAT	General Mills, Inc.	0.29	898	3.68
COFFEE	The J.M. Smucker Company	0.2	189	7.11
COFFEE	Starbucks Coffee Company	0.15	83	8.64
SOFT DRINKS	Coca-Cola USA Operations	0.32	535	2.93
SOFT DRINKS	Pepsi-Cola North America Inc.	0.27	687	2.63
MARSHMALLOWS	Kraft Heinz Foods Company	0.57	43	1.92
MARSHMALLOWS	Mazapan De La Rosa, S.A. De Cv	0.01	10	1.9

FRUIT SPREADS	The J.M. Smucker Company	0.45	25	3.13
FRUIT SPREADS	Welch Foods Inc	0.19	16	3.08
PIZZA-FROZEN	Nestle USA Inc.	0.36	418	5.12
PIZZA-FROZEN	The Schwan Food Company	0.21	492	4.07
PASTA	Nestle USA-Prepared Foods Division, Inc.	0.54	8	5.3
PASTA	Giovanni Rana	0.13	17	4.71
SAUSAGE-DINNER	The Hillshire Brands Company	0.17	72	3.49
SAUSAGE-DINNER	Johnsonville Sausage, L.L.C.	0.14	31	4.01
SAUSAGE-BREAKFAST	The Hillshire Brands Company	0.31	57	4.03
SAUSAGE-BREAKFAST	Johnsonville Sausage, L.L.C.	0.1	13	3.56
TOPPINGS	Conagra Brands, Inc.	0.62	11	3.31
TOPPINGS	Alamance Foods Inc	0.03	4	2.19
YOGURT	Danone US, LLC	0.27	74	2.36
YOGURT	Yoplait USA	0.24	87	1.29
CHEESE - COTTAGE	Kraft Heinz Foods Company	0.15	49	2.56
CHEESE - COTTAGE	Daisy Brand	0.07	4	3.08
CHEESE - CREAM CHEESE	Kraft Heinz Foods Company	0.62	79	3.15
CHEESE - CREAM CHEESE	Challenge Dairy Products Inc.	0.01	1	2.07
BUTTER	Land O'Lakes, Inc.	0.33	21	3.99
BUTTER	Challenge Dairy Products Inc.	0.07	8	4.14
BAKERY - BREAD	Bimbo Bakeries USA, Inc.	0.16	433	3.26
BAKERY - BREAD	Flowers Foods, Inc.	0.08	92	2.9
BAKERY-BAGELS	Bimbo Bakeries USA, Inc.	0.6	64	3.47
BAKERY-BAGELS	Pinnacle Foods Group LLC	0.05	27	3.28
BEER	Anheuser-Busch InBev	0.26	392	10.09
BEER	Sazerac Company, Inc.	0.18	57	12.37

These national top brands also cater to similar customer profiles. To examine this, we use the NielsenIQ consumer panel to construct a sample of households who have purchased the two brands during our sample period. This sample includes all purchases involving either of the two brands. To examine whether significant demographic differences exist among households purchasing the first national brand versus the second national brand, we construct an indicator variable that takes a value of one if a household purchases the first national brand and a value of zero if a household purchases the second national brand. We then regress the indicator variable on household demographic variables such as income, presence of children, age, education, and employment status of both female and male household heads. The sample is at the household-purchase level, and the regression includes market, category and month fixed effects. Table A2 reports the regression results in Column (1). Most of the demographic variables do not exhibit a statistically significant difference between the top and second national brands. Although there are slight differences in income, presence of children, female employment and male education, these differences are very small compared to the overall standard deviation of these variables in the sample, as reported in Column (2).

Table A2: Demographic Differences between Top Two National Brands

	$\mathbb{1}\{\text{Purchase Top 1}\}$	Standard Deviation
Log Income	0.00726***	0.663
	(0.00208)	
Presence of Kids	-0.00687***	0.456
	(0.00264)	
Female Age	9.30e-05	11.586
	(0.000204)	
Female Education	-0.000592	0.450
	(0.00257)	
Female Employment	0.00514**	0.496
	(0.00223)	
Male Age	0.000180	11.262
	(0.000202)	
Male Education	-0.00602**	0.470
	(0.00239)	
Male Employment	-0.00297	0.458
	(0.00287)	
Observations	238,476	
R-squared	0.004	
Market FEs	Yes	
Category FEs	Yes	
Month FEs	Yes	

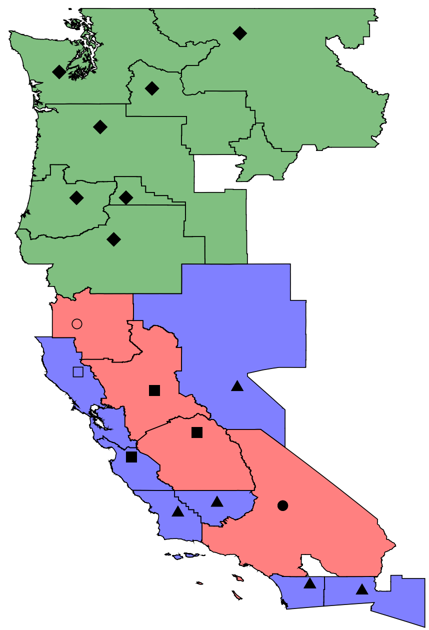
Notes: This table reports regression results of demographic differences for the national top two brands. The second column shows standard deviation of the demographic variables in the sample. *** p<0.01, ** p<0.05, * p<0.1.

A.2 Geographic Correlation between Retailers and Markets

Besides the fact that many of the retailers are regional and span multiple adjacent markets, many local markets tend to be concentrated and dominated by the same regional retailers in terms of market share. Figure A1 illustrates this pattern using a representative consumer goods category. The figure maps multiple markets, and the retailer with the highest revenue share in each location. To preserve anonymity, retailer locations are intentionally represented as distorted averages of the stores within a market. Each market is color-filled to reflect the top-selling brand at its dominant retailer.

The category demonstrates considerable concentration, with a one-firm index (C1) consistently exceeding 20% across markets. Two key patterns emerge: First, a limited number of retailers dominate clusters of neighboring markets, and a leading retailer tends to lead retail shares in a number of adjacent markets. Second, retailers that lead retail shares in nearby markets usually have an identical leading brand across all the markets, with the exception of one retailer represented by a square. These market structure features imply that market-level share distributions largely mirror those of dominant retailers. Furthermore, the fact that adjacent markets share the same "main retailer" and "leading brand" creates a "correlation" in market shares between markets and retailers, making it challenging to separate the effects of the market from those of the retailer.

Figure A1: An Example of Concentrated Retailers in Concentrated Markets



Notes: Each dot represents a retailer, with the shape of the dot indicating a unique retailer. The location of the dots is a distorted average of stores of a given anonymized retailer in a market. The filled colors of each market reflect the leading brands associated with the primary retailers in that market.

B Additional Descriptive Evidence

B.1 Replicate Bronnenberg et al. (2007)

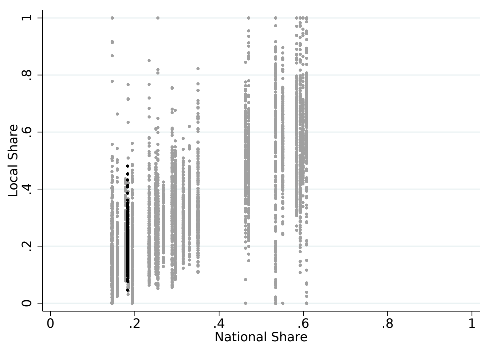
As noted in the main text, our sample covers more geographic markets and a more recent period of time than Bronnenberg et al. (2007). We replicate the main results about across-market share dispersion of top two national brands in Bronnenberg et al. (2007, 2009), using a sample of the same markets reported in their papers but from more recent years. Table B1 summarizes the distribution of the top brands' national quantity share (following Table 1 of Bronnenberg et al. (2007)), and Figure B1 plot the top brands local quantity shares against their national quantity shares (following Figure 1 of Bronnenberg et al. (2007)). Both the table and the figure confirm a comparable level of share dispersion to that observed in Bronnenberg et al. (2007, 2009).

Table B1: Description of the Top-Selling Brands across 25 categories

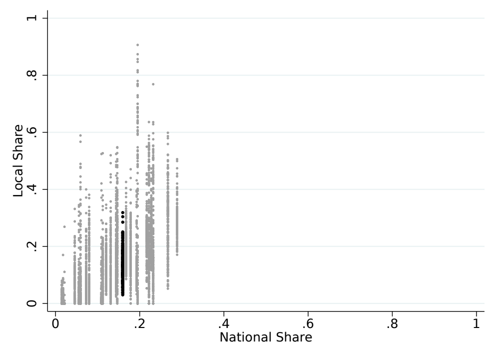
	Average Share	*Dispersion*	*Range*	*Minimum*	*Maximum*	*IQR*
all leading brands (N=50)						
mean	0.249	0.622	0.461	0.076	0.537	0.127
sd	0.202	0.417	0.221	0.095	0.245	0.085
leading brands with coverage in all markets (34)						
mean	0.298	0.441	0.441	0.112	0.553	0.120
sd	0.099	0.224	0.178	0.096	0.221	0.059
leading brands with coverage in all markets (16)						
mean	0.169	1.008	0.504	0	0.504	0.142
sd	0.126	0.474	0.296	0	0.296	0.125

Figure B1: National Quantity Shares Plotted against Local Quantity Shares across 25 Categories (with the Coffee Category Highlighted)

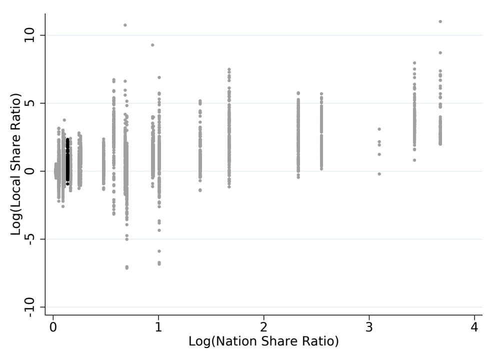
(a) Top National Brands



(b) Second National Brands



(c) Share Ratio of Top Brands

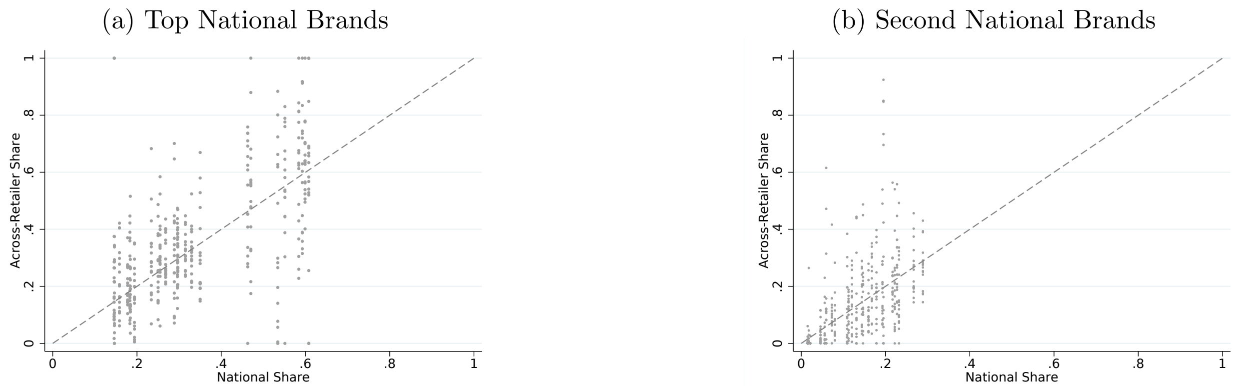


Notes: Coffee category is highlighted.

B.2 Share Dispersion of Top National Brands at National Retailers

This subsection describes the share dispersion of top two national brands at national retailers. Figure B2 plot the across-retailer shares of the top two national brands. The share dispersion across national retailers is also large. The average interquartile range across retailers and categories for top national brands at these retailers is 0.69, while for second national brands, it is 0.48.[27]

Figure B2: Across-Retailer Shares of Top and Second National Brand at National Retailers



Notes: Each vertical line of dots is a category, and each dot on a vertical line denotes quantity share of the top national brand in a market (Panel a) or in a retailer (Panel b). The dashed line is a 45 degree straight line.

C AKM Model: Identification, Algorithm, Additional Results

C.1 Leave-One-Out Identification and the Sample

Under a connectivity condition, the AKM model can be estimated by OLS. We use a simple example of one brand to illustrate connectivity.[28] Figure C1 shows a connected group that identifies all the retailer and market effects. The brand is sold in all three retailers and markets. Each market is connected to as least one other market in the group by a common retailer: for example retailer 1 is present in both market 1 and market 2, connecting these two markets. Similarly, each retailer is connected to at least one other retailer in the group by sharing a common market.[29]

However, Kline et al. (2020) prove that the OLS estimator of both variance and covariance terms are biased, and propose a bias correction method—the *leave-one-out estimator*. This method requires the data to be *leave-one-out connected*: that the firm effects remain estimable after removing any single observation, and vise versa for worker effects. This is a stricter requirement than the connectivity condition, which would strengthen the mobility

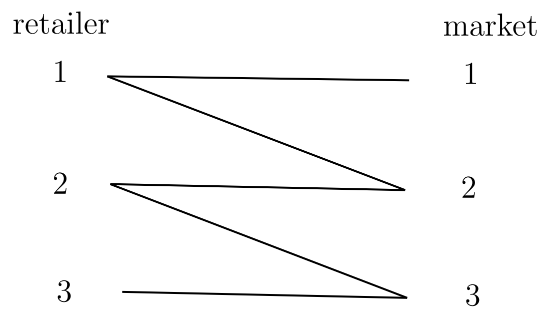
[27] Interquartile ranges are adjusted by median quantity share across retailers.

[28] Identification is the same for every brand, since brands are independent from each other.

[29] The connectivity condition can be satisfied in many circumstances. For example, for a given brand, if a market (market M) only has one retailer (retailer C) that sells that brand, but retailer C is also present in other markets, then market M is connected with the other markets. As a result, the Brand × Market fixed effect will be identified because of the multi-market retailer (retailer C). However, if this retailer C operates only in market M, then the Brand × Market fixed effect for market M will not be identified. So the only case that a Brand × Market effect is unidentified is when that brand is sold by only one local retailer in that market.

and connectivity in the sample, improving the identification of fixed effects. For example, retailer 3 in Figure C1 is not leave-one-out connected with other retailers. If the edge between retailer 2 and market 3 is removed (indicating that retailer 2 no longer operates in market 3), retailer 3 will no longer be connected to any other retailers. Similarly, market 1 will no longer be connected with any other markets if the edge between retailer 1 and market 2 is removed. Therefore, in our setting, leave-one-out connectivity is achieved by removing the retailers selling a given brand that are present in only one market (e.g., local retailers like retailer 3) and the markets with only one retailer selling that brand (e.g., monopoly markets like market 1).

Figure C1: A Simple Example of Connected Retailers and Markets



The final "leave-one-out" sample is at the brand-retailer-market-year level. Table C1 reports the mean and variance of the outcome variables (with zeros filled in for brands not sold in a retailer-market), the number of total observations, brand-retailer pairs, and brand-market pairs for both the full sample and the leave-one-out sample. The second panel of Table C1 imposes the leave-one-out requirement, keeping the brand-retailer and brand-market pairs that remain connected when any other pair is removed. Pruning the sample drops about 29% of the brand-retailer pairs and 23% of the brand-market pairs, but has little impact on the distribution of quantity shares: the mean and variance of quantity shares remain consistent with the full sample, with values of 24.57% and 5.26%, respectively, compared to 24.01% and 5.30% in the full sample.

Table C1: Summary Statistics of the Variance Decomposition Samples

	Share	Fraction of UPCs	Fraction of Products	Brand Availability	# of Observations	# of Brand-Retailers	# of Brand-Markets
Full Sample							
Local Retailers	0.217	0.0588	0.0931	0.796	4316	1917	1119
	(0.0525)	(0.00670)	(0.0116)	(0.163)			
Regional Retailers	0.235	0.0664	0.108	0.839	58968	3824	9039
	(0.0472)	(0.00750)	(0.0140)	(0.135)			
National Retailers	0.245	0.0554	0.0935	0.789	86216	1265	11185
	(0.0569)	(0.00623)	(0.0118)	(0.167)			
Pooled	0.240	0.0599	0.0994	0.809	149500	6348	11918
	(0.0530)	(0.00677)	(0.0127)	(0.155)			
Leave-One-Out Sample							
Local Retailers	0.145	0.0377	0.0617	0.506	1496	878	675
	(0.0440)	(0.00531)	(0.00997)	(0.250)			
Regional Retailers	0.240	0.0675	0.110	0.854	55674	3741	8472
	(0.0464)	(0.00743)	(0.0136)	(0.125)			
National Retailers	0.252	0.0561	0.0945	0.803	80007	1260	10047
	(0.0568)	(0.00621)	(0.0115)	(0.158)			
Pooled	0.246	0.0605	0.100	0.820	137177	5291	10527
	(0.0526)	(0.00673)	(0.0124)	(0.147)			

Notes: Variances are in the parentheses.

C.2 Leave-One-Out Algorithm

We follow the leave-one-out algorithm proposed by Kline et al. (2020) to estimate the AKM model. We first estimate the bias terms using an unbiased estimator of the error variance, and then calculate the new variance estimator as the OLS variance estimator plus the bias correction term. Borrowing notations from their paper, let $\theta = \beta' A \beta$ be the variance or covariance components from a linear model $y_i = x_i'\beta + \epsilon_i$, then the bias corrected variance estimator is the OLS estimator $\hat{\beta}' A \hat{\beta}$ subtracting a bias:

$$\hat{\theta} = \underbrace{\hat{\beta}' A \hat{\beta}}_{\substack{\text{“plug-in”} \\ \text{variance component}}} - \underbrace{\sum_{i=1}^{n} B_{ii} \hat{\sigma}_i^2}_{\text{error variance}} \tag{7}$$

where $B_{ii} = x_i' S_{xx}^{-1} A S_{xx}^{-1} x_i$; and where S_{xx} is the design matrix. $\hat{\sigma}_i^2$ is an unbiased estimator of the i-th error variance: $\hat{\sigma}_i^2 = y_i \left(y_i - x_i' \hat{\beta}_{-i} \right) = y_i \left(\frac{y_i - x_i'\hat{\beta}}{1 - P_{ii}} \right)$, and $P_{ii} = x_i' S_{xx}^{-1} x_i$ gives the leverage of observation i.

This bias-correction method is computationally heavy, since it involves calculation of the error variance for each observation, which requires inversions of large matrices (the design matrix). We use preconditioned conjugate gradients method to solve for B_{ii}, P_{ii} and $\hat{\beta}$.

C.3 AKM Decomposition — Additional Results

We estimate the AKM decomposition model on extensive margins of the market outcome variables, and report the results in Table C2. The relative magnitude of the percentages is highly consistent across all the outcome variables along the extensive margins: the Brand × Retailer match effect is the primary factor that explains the majority of variation in quantity shares, product assortment depths, and brand availability. It accounts for about 50% to 70% of the total variance, a magnitude two to three times greater than that attributed to the Brand × Market fixed effects.

Table C2: AKM Variance Decomposition—Extensive Margin

	(1) Quantity Share	(2) Fraction of UPCs	(3) Fraction of Products	(4) Brand Availability
Brand × Retailer	52.75	65.91	60.44	50.36
Brand × Market	21.48	9.84	11.03	17.48
2Cov	3.05	3.87	5.14	-2.09
RMSE	0.082	0.033	0.046	0.167
R2	0.773	0.796	0.766	0.658

Notes: This table reports AKM variance decomposition for the extensive-margin outcome variables. We report for each dimension the percentage of its variance relative to the total variance of the outcome variables.

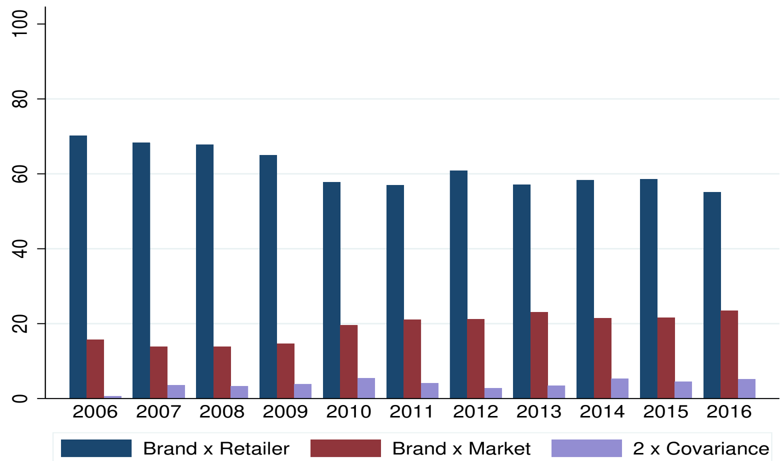
Table C3 reports the AKM decomposition results on national retailers.

Table C3: AKM Decomposition on National Retailers

	(1) Quantity Share	(2) Fraction of UPCs	(3) Fraction of Products
Brand × Retailer	50.37	65.15	56.3
Brand × Market	17.87	8.48	12.01
2Cov	0.3	-0.96	0.67
RMSE	0.099	0.032	0.046
R2	0.69	0.727	0.69

Notes: This table reports AKM variance decomposition for the intensive-margin outcome variables among the national retailers.

Figure C2: AKM Decomposition Results by Years



We estimate the AKM model to decompose the total variance in HHI at the retailer-market-year level into Category × Retailer, Category × Market and the covariance terms. The estimated percentages of variances are reported in Table C4. Column (1) employs the HHI measure calculated with private label shares, Column (2) uses the HHI excluding private label shares, and Column (3) uses the HHI calculated using the shares of the top two national brands.

Table C4: AKM Variance Decomposition—HHI

	(1) With private label	(2) Without private label	(3) Top Two Brands
Category × Retailer	59.53	60.38	54.34
Category × Market	9.64	6.5	8.64
2Cov	1.45	0.62	1.43
RMSE	0.075	0.10	0.11
R2	0.706	0.675	0.644

Notes: This table reports AKM variance decomposition for category-retailer-market HHI. We decompose the total variance in HHI into *Category* × *Retailer* and *Category* × *Market* components. We report for each component the percentage of its variance relative to the total variance of the outcome variables.

C.4 AKM Decomposition — Robustness

Table C6 and Table C5 report AKM variance decomposition results using OLS.

Table C5: AKM Decomposition using OLS—Intensive Margin

	Quantity Share	Fraction of UPCs	Fraction of products
Brand × Retailer	54.08	69.11	62.55
Brand × Market	25.9	14.34	15.01
2Cov	2.15	-5.18	-4.44
RMSE	0.061	0.024	0.031
R2	0.821	0.783	0.731

Notes: This table reports AKM variance decomposition for the intensive-margin outcome variables using OLS.

Table C6: AKM Decomposition using OLS—Extensive Margin

	Quantity Share	Fraction of UPCs	Fraction of Products	Brand Availability
Brand × Retailer	53.83	66.85	61.5	52.11
Brand × Market	23.32	11.8	13.43	20.75
2Cov	2.46	3.29	4.46	-3.07
RMSE	0.073	0.029	0.04	0.146
R2	0.796	0.819	0.794	0.698

Notes: This table reports AKM variance decomposition for the extensive-margin outcome variables using OLS.

We also estimate the AKM decomposition using the *parent_code* as the retailer identifier instead of the *retailer_code*. Table C7 presents the results for the set of extensive margin variables in this estimation.

Table C7: AKM Decomposition with Parent_Code as Identifier

	Quantity Share	Fraction of UPCs	Fraction of Products
Brand × Retailer	57.72	70.37	62.63
Brand × Market	21.38	7.78	10.3
2 Cov	0.14	2.7	5.66
RMSE	0.074	0.03	0.041
R2	0.792	0.809	0.786

Notes: This table reports AKM variance decomposition for the extensive-margin outcome variables, using parent_code.

Lastly, we examine the explanatory power of retailer-level private label shares. We incorporate the retailer-market-year level private label shares on the right-hand side of Equation 3. The results, reported in Table C8, are also very similar to the main results.

Table C8: AKM Decomposition with Private Label Shares as Control

	Quantity Share	Fraction of UPCs	Fraction of Products
Brand × Retailer	52.11	66.28	61.15
Brand × Market	21.68	9.94	11.1
2 Cov	5.75	4.6	5.73
RMSE	0.079	0.033	0.046
R2	0.787	0.8	0.77

Notes: This table reports AKM variance decomposition for the extensive-margin outcome variables, controlling for retailer-market level private label shares.

D Comparison with Previous Studies

In this section, we compare our findings with those of Hwang and Thomadsen (2016) and Bronnenberg et al. (2007). Using IRI data from 2004-2005, Hwang and Thomadsen (2016) finds that approximately 30% of the weekly purchase share variation across stores is driven by market-level factors, with account-level (i.e. retailer-level) and store-level factors contributing an additional 13% and 5%, respectively. Bronnenberg et al. (2007), analyzing NielsenIQ data from 2006, identifies substantial market-level share dispersion. The differences in results are likely attributable to variations in the data sources and methodologies used. Table D1 summarizes the data, methods, and key findings from our study alongside these earlier works.

Table D1: Comparison of Methods and Findings across Studies

	This paper	**Hwang and Thomadsen (2016)**	**Bronnenberg et al. (2007)**
Data	NielsenIQ (2013-2016)	IRI (2004-2005)	NielsenIQ (2006)
Method	AKM variance decomposition	ANOVA and COV	R-squared decomposition
Finding	Brand-retailer accounts for 53% of share dispersion. Brand-market explains 24%.	Market-level variation accounts for 30% of share variation. Retailer-level explains 13%.	Large market-level share dispersion.

To better understand the sources of differences in the results, we assemble three different samples: IRI (2004-2005) and NielsenIQ (2006), both using the categories and brands from Hwang and Thomadsen (2016), and our main sample, NielsenIQ (2016).[30] These samples are further aggregated to different levels: brand-retailer-market-year, brand-retailer-market-quarter, brand-retailer-market-week, and brand-store-week.[31]

[30] The IRI sample we assembled consists of 51 markets, which is three more than reported in Hwang and Thomadsen (2016).

[31] The sample in Hwang and Thomadsen (2016) is at the brand-store-week level, while the sample in Bronnenberg et al. (2007) is at the brand-market-month level.

Using these samples, we estimate both the model employed in our paper (AKM) and the models used in Hwang and Thomadsen (2016) (ANOVA and COV). For ANOVA, we adopt the specification outlined in their study as follows and adjust to different samples of aggregation levels:

$$
\begin{aligned}
PS_{s,c,m,t} &= \gamma_t + \varepsilon_{s,c,m,t} \\
PS_{s,c,m,t} &= \alpha_m + \gamma_t + \varepsilon_{s,c,m,t} \\
PS_{s,c,m,t} &= \beta_{c,m} + \gamma_t + \varepsilon_{s,c,m,t} \\
PS_{s,c,m,t} &= \eta_{s,c,m} + \gamma_t + \varepsilon_{s,c,m,t}
\end{aligned} \tag{8}
$$

where s denotes store, c retailer, m market, and t time. $PS_{s,c,m,t}$ represents purchase share of a brand at a store, retailer, market and time. γ_t, α_m, $\beta_{c,m}$ and $\eta_{s,c,m}$ are time, market, retailer-market, and store fixed effects respectively. We sequentially estimate these regressions, and calculate how much adding a set of fixed effects at a finer level marginally increases the regression's R^2 as we move from the top equation to the bottom. A feature of ANOVA is that it attributes covariance between effects to the first effect introduced into the model (Hwang and Thomadsen 2016, McGahan and Porter 1997). Consequently, different ordering of a set of predictor variables can yield different results. Therefore, we estimate ANOVA by employing two models with identical fixed effects but different ordering. In Model 1, market effects are introduced first, followed by retailer-market effects. In contrast, Model 2 begins with retailer effects, followed by retailer-market effects.

For COV, we also follow Hwang and Thomadsen (2016) to decompose the total variance of purchase share at the store (s) - retailer (c) - market (m) - time (t) level into an across-market component (α_m), a within-market across-retailer component ($\beta_{c,m}$), a within-retailer-market across-store component ($\eta_{s,c,m}$), and a time component (γ_t). The estimation model is given by the random effect model in Equations 9, where σ^2 denotes variance.

$$
\begin{aligned}
PS_{s,c,m,t} = \alpha_m + \beta_{c,m} + \eta_{s,c,m} + \gamma_t + \varepsilon_{s,c,m,t} \\
\sigma^2\left(PS_{s,c,m,t}\right) = &\sigma^2\left(\alpha_m\right) + \sigma^2\left(\beta_{c,m}\right) + \sigma^2\left(\eta_{s,c,m}\right) \\
&+ \sigma^2\left(\gamma_t\right) + \sigma^2\left(\varepsilon_{s,c,m,t}.\right)
\end{aligned} \tag{9}
$$

Table D2 presents the results (reported as percentage variation) from AKM, ANOVA, and COV models using IRI data (2004–2005). Several key insights emerge: (1) The results at the brand-store-week level from closely align with those of Hwang and Thomadsen (2016), with the market level accounting for 29% of the variance and the retailer level accounting for 12%. (2) A comparison of ANOVA Model 1 and Model 2 reveals a significant difference in the proportion attributed to the retailer-market effect. The Brand × Retailer effects account for more variance in Model 2 than Model 1 since it is introduced first into Model 2 first. This highlights the key limitation of ANOVA that it is sensitive to the order in which effects are introduced into the model. (3) The AKM model decomposition indicates that the Brand × Market effects account for a larger share of variation in this IRI sample. It also suggests relatively large covariance between the retailer and market effects.

Table D2: Results from IRI 2004-2005 Data

	AKM	ANOVA Model 1	ANOVA Model 2	COV
Brand-retailer-market-year level				
Brand × Retailer	10.91	28.75	64.31	14.23
Brand × Market	53.38	69.3	33.73	81.05
2 Covariance	12.87	–	–	–
R2	77.2	98.46	98.46	95.47
Brand-retailer-market-quarter level				
Brand × Retailer	13.07	22.66	55.78	61.24
Brand × Market	54.47	63.13	30.01	9.7
2 Covariance	7.71	–	–	–
R2	75.2	88.02	88.02	80.09
Brand-retailer-market-week level				
Brand × Retailer	9.66	12.67	35.1	30.99
Brand × Market	36.98	41.36	18.93	35.73
2 Covariance	4.81	–	–	–
R2	51.5	58.01	58.01	84.97
Brand-store-week level				
Brand × Retailer	8.33	12.38	31.27	13.02
Brand × Market	26.84	28.69	9.8	29.56
2 Covariance	0.22	–	–	–
R2	35.4	49.99	49.99	45.68

Table D3 and Table D4 present results from NielsenIQ (2006) and NielsenIQ (2013-2016) respectively. In both samples, we find that the Brand × Retailer component explains more share variation than the Brand × Market component, and this conclusion is consistent across different models and levels of aggregation.

Table D3: Results from NielsenIQ 2006 Data

	AKM	ANOVA Model 1	ANOVA Model 2	COV
Brand-retailer-market-year level				
Brand × Retailer	83.64	75.61	87.69	72.54
Brand × Market	5.61	24.39	12.31	17.01
2 × Covariance	2.47	–	–	–
R2	91.7	100	100	89.54
Brand-retailer-market-quarter level				
Brand × Retailer	79.98	70.15	82.3	82.27
Brand × Market	6.28	22.79	10.65	8.66
2 Covariance	2.26	–	–	–
R2	88.5	93.75	93.75	91.6
Brand-retailer-market-week level				
Brand × Retailer	61.97	52.27	61.34	63.39
Brand × Market	5.06	16.01	6.94	4.83
2 Covariance	1.41	–	–	–
R2	68.4	71.51	71.51	71.25
Brand-store-week level				
Brand × Retailer	49.4	23.77	42.66	29.11
Brand × Market	6.36	24.58	5.69	19.97
2 Covariance	2.28	–	–	–
R2	58.1	57.78	57.78	52.82

Table D4: Results from NielsenIQ 2013-2016 Data

	AKM	ANOVA Model 1	ANOVA Model 2	COV
Brand-retailer-market-year level				
Brand × Retailer	52.89	42.87	58.13	58.21
Brand × Market	23.88	39.28	24.02	23.95
2 × Covariance	3.03	–	–	–
R2	79.8	83.42	83.42	84.62
Brand-retailer-market-quarter level				
Brand × Retailer	51.32	46.60	64.75	53.45
Brand × Market	22.78	38.88	20.72	31.34
2 × Covariance	0.09	–	–	–
R2	79.8	88.44	88.44	86.23
Brand-retailer-market-week level				
Brand × Retailer	20.91	30.97	52.21	44.28
Brand × Market	8.74	36.90	15.66	24.44
2 × Covariance	0.10	–	–	–
R2	63.8	71.84	71.84	75.09
Brand-store-week level				
Brand × Retailer	31.58	14.80	42.82	40.18
Brand × Market	17.42	36.30	8.28	14.59
2 × Covariance	0.073	–	–	–
R2	52.9	55.71	55.71	58.41

In sum, the differences in our results relative to Hwang and Thomadsen (2016) seems to be mainly due to data difference. One notable distinction between IRI and NielsenIQ is the extent of their geographic coverage. The IRI sample exhibits more limited geographic coverage. On average, retailers in the IRI sample operate in 2.44 markets, whereas retailers in our main NielsenIQ (2016) sample operate in 8 markets. This discrepancy can lead to differences in results, even when employing the same model.

To further explore this difference, we compute the connectivity measure recommended by Kline et al. (2020). As detailed in Appendix C.2, the leave-one-out algorithm calculates the leverage, P_{ii}, for observation i. The value of P_{ii} is less than one if the origin and destination firms of worker i are connected by a path that does not involve i (see Page 28, Example 4, of Kline et al.).[32] This calculation is performed for each observation; thus, a smaller maximum leverage indicates a better-connected sample. The maximum leverage is 0.72 for IRI data and 0.5 for NielsenIQ data (2013–2016), indicating that NielsenIQ data (2013–2016) exhibits better connectivity properties. In our setting, this reflects the presence of multi-market retailers and markets with multiple retailers, which better facilitate the identification of variance decomposition models.

[32]The leverage formula is introduced on Page 6 of Kline et al. (2020): $P_{ii} = x_i' S_{xx}^{-1} x_i$.

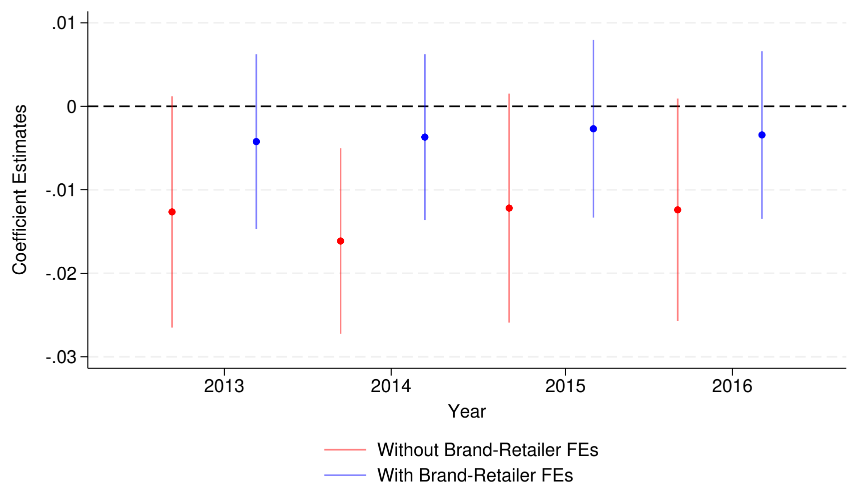
E Additional Analyses on Fixed Effects

We examine the relationship between brand-market heterogeneity and distance over the years, by regressing the Brand $\times$ Market FEs on the distance to city of origin in each year:

$$\Gamma_{bcmt} = \beta_t \,\text{Dist}_{bcmt} + \gamma_c + \epsilon_{bcmt},$$

We plot these coefficients in Figure E1, where the red dots represent β_t's from regressions of Γ_{bcmt} without including Brand-Retailer controls, and the blue dots represent results with these controls. A Wald test comparing β_t across years indicates no significant changes, suggesting that the spatial share variation remain stable over time. However, the coefficients decrease by half when Brand-Retailer effects are accounted for in every year, suggesting that the substantial market demand heterogeneity reported in earlier studies may partly capture unaccounted retailer heterogeneity.

Figure E1: Brand-Market Fixed Effects Over Years



We also examine the relationship between within-retailer brand asymmetry and distance over the years, by regressing the Brand $\times$ Retailer FEs on the distance to brand's city of origin in each year. The estimated coefficients over the years are visualized in Figure E2, with red dots indicating results without Brand-Market controls and blue dots indicating results with these controls. While the coefficients show a slight decrease over the years, hinting at a possible increase in retailer heterogeneity, Wald test statistics reveal no significant differences across years. Moreover, the close similarity in magnitude between the red and blue dots indicates that retailer heterogeneity remains consistent, regardless of whether market demand heterogeneity is accounted for.

Figure E2: Distance regressions over years: Brand × Retailer fixed effects