The Evolution of U.S. Retail Concentration*

Dominic A. Smith† — Bureau of Labor Statistics

Sergio Ocampo‡ — Western University

April 13, 2021

Abstract

Increases in concentration have been a salient feature of industry dynamics in the United States. This trend is particularly notable in the retail sector, where large national firms have displaced small local firms. Existing work focuses on national trends, yet less is known about the dynamics of concentration in *local* markets and the relationship between local and national trends. We address these issues by providing a novel decomposition of the national Herfindahl-Hirschman Index into an average local HHI and a cross-market component that accounts for firms in multiple locations. We measure concentration using new data on product-level revenue for all U.S. retail stores and find that despite local concentration increasing by 34 percent between 1992 and 2012, the cross-market component explains 99 percent of the rise in national concentration, reflecting the expansion of multi-market firms. We estimate an oligopoly model of retail competition and find that the increase in markups implied by rising local concentration had a modest effect on retail prices.

JEL: L8

Keywords: Retail, Local Markets, Concentration, Herfindahl-Hirschman Index

*Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1975 (CBDRB-FY20-P1975-R8604). We gratefully acknowledge advice from Thomas Holmes, Teresa Fort, Amil Petrin, Joel Waldfogel, Emek Basker, Emily Moschini, Brian Adams, Jeff Thurk, and Juan Herreño. We also thank participants of seminars at the George Washington University, Norwegian Business School (BI), Bureau of Labor Statistics, Federal Reserve Board, Federal Reserve Bank of Cleveland, Kent State University, 2019 SED, Federal Trade Commission, and Center for Economic Studies. This paper is based upon work supported by the Doctoral Dissertation Fellowship at the University of Minnesota.

†Email: smith.dominic@bls.gov; Web: https://www.bls.gov/pir/authors/smith.htm

‡Email: socampod@uwo.ca; Web: https://sites.google.com/site/sergiocampod/

1 Introduction

U.S. retailing has become substantially more concentrated. Between 1997 and 2007 alone, the share of sales going to the 20 largest firms increased from 18.5 percent to 25.4 percent (Hortaçsu and Syverson, 2015) and the national Herfindahl-Hirschman Index (HHI) in retail doubled. Increases in concentration have been accompanied by rising markups, which raises concerns about increasing market power (Hall, 2018; De Loecker, Eeckhout, and Unger, 2020). However, much of the growth of national concentration is caused by firms opening new establishments (Cao, Hyatt, Mukoyama, and Sager, 2019), a change that would not necessarily increase market power in retail.

In this paper, we show that national industry-level concentration trends in the U.S. contain no information on changes in local product markets, which is the relevant market definition in retail because consumers primarily choose between local stores selling a given product. We study the relationship between national and local trends using new data on sales by product category for all U.S. retail establishments. We find that less than 1 percent of the change in national concentration is explained by increases in local concentration. The remaining 99 percent of the change comes from consumers in different markets increasingly buying from the same firms, a phenomenon that we call cross-market concentration.

We find this result by implementing a new decomposition of national concentration as measured by the HHI. We interpret the HHI as the probability that two dollars spent at random are spent at the same firm. The decomposition exploits the law of total probability to separate the national HHI into a weighted average of the probability that two dollars *spent in the same market* are spent at the same firm (local concentration) and the probability that two dollars *spent in different markets* are spent at the same firm (cross-market concentration). Local concentration is weighted in the decomposition by the probability that two dollars are spent in the same market *regardless of the firms* at which they are spent, a measure of how concentrated spending is across markets. We call this measure collocation.

Applying the decomposition to U.S. data makes it clear that changes in national concentration must be driven by changes in cross-market concentration reflecting the increasing importance of multi-market firms. The distribution of retail sales across locations in the U.S. implies a low weight on local concentration—the collocation term is less than 2 percent throughout our sample—capturing the fact that even the largest retail markets in the U.S. are too small to affect national concentration. Because of this, a firm can only be large at the national level if it is present in many markets. In this sense, the trends in national concentration contain no information about the competitive environment in local markets.

Thus, understanding what is happening in local markets requires data on local sales. We assemble the necessary data using the Census of Retail Trade (CRT). The data cover all U.S. retail stores which allows us to measure local concentration for the entire United States. Our data span 1992 to 2012 which allows us to document the distribution of changes in concentration over 20 years. Crucially, the data contain sales by product category which not only allows us to properly define product markets, but also to handle retailers that sell multiple products by appropriately assigning their sales across markets.

We use these data to document three new facts on concentration in the retail sector. First, the decomposition of national HHI shows that cross-market concentration accounts for 99 percent of the change in national concentration, with local concentration accounting for less than 1 percent. Second, both the national and local HHI increase between 1992 and 2012, but at different rates, with the national HHI increasing faster than the local HHI. National and cross-market concentration more than tripled, from 1.3 percent to 4.3 percent. Local concentration increased by one-third, from 6.4 percent to 8.5 percent. Third, the majority of markets and product categories feature increasing concentration. The local HHI increased in 70 percent of commuting zones increasing their concentration between 1992 and 2002. The local HHI also increased for seven of the eight major product categories in retail between 1992 and 2012, with Clothing being the exception.[1]

We take into account online and other non-store retailers and find that they have a small effect on local concentration because they account for less than 10 percent of CRT sales throughout our sample. Establishing the exact effect of non-store retailers on local concentration is challenging because the CRT does not contain the location of sales for non-store retailers. Nevertheless, we obtain bounds for the effect of introducing non-store retailers by assigning their national sales to local markets using a range of assumptions on how concentrated their local sales would be. Under most assumptions local concentration would slightly decrease relative to our main results.

We improve over previous measures of retail concentration that rely on industry-based classifications of retail markets by defining markets on a product basis.[2] Industry-based measures do not account for the increasing importance of multi-product retailers in the general merchandising subsector which, by definition, sell the same products as retailers in other industries. For example, Walmart is in the general merchandising subsector (3-digit NAICS 452) but competes with grocery, clothing, and toy stores.[3] In fact, general merchandisers account for more than 20 percent of sales in Electronics & Appliances, Groceries, and Clothing, demonstrating that competition across industries is a relevant feature of retail markets.

We link the broad increase in local concentration to increases in retail markups. We find that increasing local concentration raised retail markups by 2 percentage points between 1992 and 2012. To do this, we use a simple model of local retail competition based on the work of Atkeson and Burstein (2008) and Grassi (2017) to ask how much markups would be expected to increase due to the observed increases in local concentration. The model is

[1]The eight major product categories are Clothing, Furniture, Sporting Goods, Electronics & Appliances, Health Goods, Toys, Home Goods, and Groceries. These categories account for 82 percent of retail sales throughout the sample.

[2]In Appendix D we document differences between industry- and product-based measures of concentration. These measures are conceptually different, as they have different definitions of a market.

[3]Walmart reports SIC code 5331 to the Security and Exchange Commission, which corresponds to NAICS 452990 (Securities and Exchange Commission, 2020). References to specific firms are based on public data and do not imply the company is present in the confidential data.

tractable enough to derive an explicit link between the local HHI and average markups at the product level. We exploit this link to estimate the model with available data from the CRT and the Annual Retail Trade Survey (ARTS), making it possible to study the historic relationship between concentration and markups despite the lack of long series on prices and costs for U.S. retailers.

The effect of local concentration on consumers through markups is likely to be limited. If increases in concentration are caused by low-cost firms increasing their market share, prices may fall despite increases in markups (Bresnahan, 1989). In fact, the 2 percentage point increase in markups due to local concentration is small relative to the 34 percent decrease in relative retail prices observed in the same period. Even if lower costs could have been achieved without increased local concentration, prices would have fallen only 1 percentage point more over 20 years.

Our main findings documenting the evolution of U.S. retail concentration complement previous work that has found increasing concentration in retail and other sectors of the economy.[4] Our results based on administrative data corroborate the strong increase in national retail concentration and show that local concentration has also increased, albeit at a lower rate. The local trends are in line with Rinz (2020) and Lipsius (2018) who find increasing local labor market concentration in the retail sector using the Longitudinal Business Database.

We provide new series of national and local retail concentration by individual product categories, which better reflect the nature of competition inside retail. Our product-based measures differ from the retail sector results in Rossi-Hansberg, Sarte, and Trachter (2020) and Benkard, Yurukoglu, and Zhang (2021) which both find decreasing local concentration. Rossi-Hansberg et al. (2020) base their results on data from the National Establishment Time Series (NETS), which has issues tracking establishments over time

[4]See Basker, Klimek, and Van (2012); Foster, Haltiwanger, Klimek, Krizan, and Ohlmacher (2016); Hortaçsu and Syverson (2015); Grullon, Larkin, and Michaely (2019); Autor, Dorn, Katz, Patterson, and Van Reenen (2020); Ganapati (2020).

making it problematic for measuring trends (Crane and Decker, 2020). Moreover, the NETS groups most general merchandise stores into a single eight-digit SIC code (53119901), ignoring competition between these stores and those in other industries.[5] On the other hand, Benkard et al. (2021) study the brand of products that consumers purchase, finding that brand concentration is high, but that both national and local concentration decrease over time. Both retail and brand concentration are important for consumers. However, they are conceptually different and can move independently from each other. We measure concentration in the retail firms from which consumers buy, while Benkard et al. (2021) measure concentration in the brands that consumers purchase, which may be available from different retailers.

We also contribute to work documenting changes in the structure of the retail sector by showing that national concentration does not reflect trends in local concentration. Instead, increasing national concentration reflects consumers in different markets shopping at the same firms. Thus, we help highlight the role of the expansion of large firms in explaining changes in the U.S. firm size distribution (Cao et al., 2019; Hsieh and Rossi-Hansberg, 2019). These large retail firms, particularly Walmart and Target, have been shown to lead to the closing of small stores (Jia, 2008; Haltiwanger, Jarmin, and Krizan, 2010), the closing of grocery chains (Arcidiacono, Bayer, Blevins, and Ellickson, 2016), and lower retail employment in local labor markets (Basker, 2005). We show that local concentration has indeed risen, but less sharply than national concentration.

More broadly, we contribute to work documenting the increasing importance of large firms as reflected by trends in national concentration. These trends have been related to the decline in the labor share (Autor et al., 2020), the decline of churn and reallocation of aggregate activity to large established firms (Decker, Haltiwanger, Jarmin, and Miranda, 2014, 2020), lower long-term growth due to lower innovation as competition decreases (Aghion, Bergeaud, Boppart, Klenow, and Li, 2019), and concerns about market power

[5] In Appendix E we show that the differences between our results are equally caused by differences in data, methodology, and market definition.

and rising markups (Hall, 2018; Traina, 2018; Edmond, Midrigan, and Xu, 2019; De Loecker et al., 2020). However, many of these concerns would operate through local markets, particularly in labor and retail markets. For instance, higher local employment concentration has been shown to negatively impact wages (Berger, Herkenhoff, and Mongey, 2019; Jarosch, Nimczik, and Sorkin, 2020; Azar, Berry, and Marinescu, 2019; Rinz, 2020). Despite these concerns, our results show that local concentration has had a limited effect on retail markups between 1992 and 2012, and does not explain the increase in markups during this period.

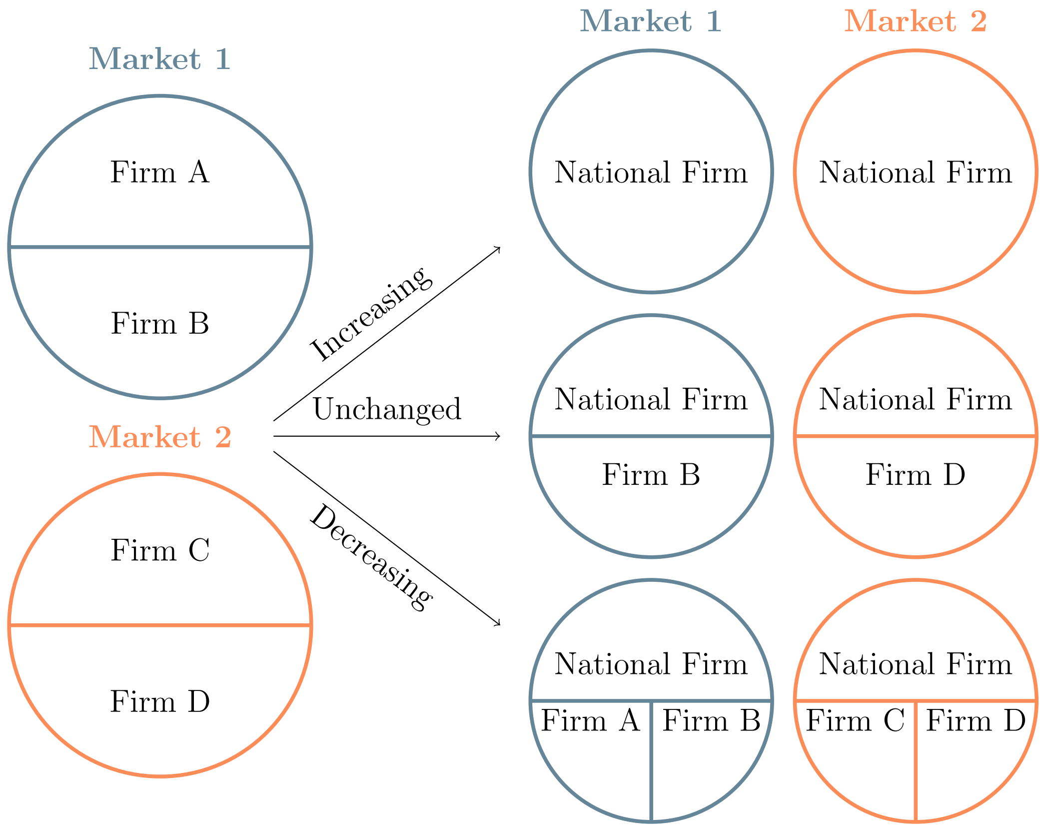
The rest of the paper proceeds as follows. Section 2 develops a decomposition of national concentration into local and cross-market concentration. Section 3 describes the data, including how we construct store-level sales by product. Section 4 measures national and local concentration and establishes the main facts about their evolution since 1982. Section 5 discusses the effects of local concentration on markups. Section 6 concludes.

2 National and Local Concentration

This section explains the relationship between national and local concentration. Increasing national concentration can be accompanied by increasing local concentration, but it may also be accompanied by decreases in local concentration. In fact, not much can be learned from the dynamics of local concentration if we only have information about national trends as is shown in Figure 1. National concentration can increase by having firms expand across markets, without affecting the layout of individual markets (row 2). Alternatively, the expansion of large firms can drive out competitors in local markets, increasing national and local concentration (row 1), or it can bring up more—potentially smaller—competitors, decreasing local concentration (row 3). The total effect on national and local concentration depends on how firms in individual markets respond.

The example in Figure 1 highlights the two mechanisms affecting national concentration

Figure 1: Effect of Increasing National Concentration on Local Concentration



Notes: The figure shows hypothetical market structures after the entry of a national firm in a two market economy that starts with four local firms.

that we study in this paper: changes in local concentration and changes in cross-market concentration. The first mechanism links changes in the composition of local markets and concentration at the national level. As local markets become more/less concentrated, so does the aggregate economy. The second mechanism links national concentration to the presence of multi-market firms. As firms expand across markets, they capture a larger share of national sales, in turn increasing national concentration. Note that, as shown in Figure 1, changes in cross-market concentration need not be accompanied by changes in local concentration. In what follows we make these ideas precise by developing a new decomposition of national concentration into local and cross-market concentration.

Our primary measure of concentration is the firm Herfindahl-Hirschman Index (HHI) for a given product category.We denote by i an individual firm and by j a product, so that

s_i^{jt} represents the sales share of firm i in product j at time t. More generally, we define subscripts and superscripts such that s_a^b is the share OF a IN b. The national HHI in a year is defined as the sum of the product-level HHIs, weighted by the share of product j's sales in total retail sales, s_j^t:

$$HHI^t = \sum_{j=1}^{J} s_j^t HHI_j^t, \quad \text{with} \quad HHI_j^t = \sum_{i=1}^{N} \left(s_i^{jt}\right)^2, \tag{1}$$

while the HHI of location ℓ and product j in year t is calculated as

$$HHI_{\ell j}^t = \sum_{i=1}^{N} \left(s_i^{j\ell t}\right)^2. \tag{2}$$

The national HHI for product j measures the probability that two dollars, x and y, chosen at random, are spent at the same firm.[6] We use the law of total probability to derive a decomposition of the HHI into two terms, based on whether the two dollars are spent in the same or different markets. The decomposition is given by

$$\underbrace{P(i_x = i_y)}_{\text{National HHI}} = \overbrace{P(\ell_x = \ell_y)}^{\text{Collocation}} \underbrace{P(i_x = i_y | \ell_x = \ell_y)}_{\text{Local HHI}} + \overbrace{P(\ell_x \neq \ell_y)}^{1 - \text{Collocation}} \underbrace{P(i_x = i_y | \ell_x \neq \ell_y)}_{\text{Cross-Market HHI}}, \tag{3}$$

where i_x is the firm at which dollar x is spent and ℓ_x is the location of the market in which dollar x is spent, and likewise for y.

Equation (3) has three components. The first component, $P(\ell_x = \ell_y)$, which we term collocation, captures the probability that two dollars are spent in the same location.[7] The second component, $P(i_x = i_y | \ell_x = \ell_y)$, is an aggregate index of local concentration, with

[6] In what follows, the j and t superscripts are dropped on all variables for convenience. In this context a market is characterized by its location, ℓ, as the product is fixed.

[7] The collocation term is $P(\ell_x = \ell_y) = \sum_{\ell=1}^{L} (s_\ell)^2$, where s_ℓ is the share of location ℓ in national sales.

local concentration measured as in equation (2).[8] This captures the extent to which consumers in a local market shop at the same firm. The third component, $P(i_x = i_y|\ell_x \neq \ell_y)$, which we call cross-market concentration, captures the probability that a dollar spent in different markets is spent at the same firm:

$$P(i_x = i_y|\ell_x \neq \ell_y) = \underbrace{\sum_{\ell}\sum_{n\neq\ell} \frac{s_\ell s_n}{1-\sum_p s_p^2}}_{\text{Weights}} \underbrace{\sum_{i=1}^{N} s_i^\ell s_i^n}_{\text{Cross-Market}} \ . \tag{4}$$

The cross-market concentration index between two markets (say ℓ and n) is given by the product of the shares of the firms in each location (the probability that two dollars spent one in each location are spent in the same firm). The pairs of markets are then weighted by their share of sales and are summed.

The collocation term plays a crucial role in determining the impact of local concentration in national measures. A low collocation term implies that local concentration can only have a limited effect on national trends, leaving the cross-market term as the driver of the national index. We will show later that this is in fact the case, which should come as no surprise because the U.S. has many markets and even the largest markets represent only a small fraction of total U.S. sales.

To implement the decomposition presented in equation (3), we need to measure concentration in each local market for a given product as well as link the activities of firms across markets. Doing this requires detailed data on establishment-level revenue by product for all firms in the U.S., which we describe in the next section.

[8]In the decomposition each local market is weighted by the conditional probability that the two dollars are spent in location ℓ given that they are spent in the same location: $s_\ell^2/(1-\sum_p s_p^2)$. These weights give more importance to larger markets than the more usual weights s_ℓ—the share of sales (of product j) accounted for by location ℓ (at time t). We present aggregated series for local concentration in Section 4 that use the latter weights. Appendix A derives these results in detail.

3 Data: Retailer Revenue for All U.S. Stores

This section describes the creation of new data on store-level revenue for 18 product categories for all stores with at least one employee in the U.S. retail sector. These data allow us to construct detailed measures of concentration that take into account competition between stores selling similar products in specific geographical areas.[9]

3.1 Data Description

We use confidential U.S. Census Bureau microdata that cover 1992 to 2012 (U.S. Census Bureau, 1992-2012). The source of the data is the Census of Retail Trade (CRT), which provides revenue by product type for retail stores in years ending in 2 and 7. We use CRT data on product-level revenue and information on each store's location to define which stores compete with each other. Importantly, a store's local competition will include stores in many different industries inside the retail sector because stores of different industries can sell similar products. This is particularly relevant for stores in the general merchandising subsector. The data we create here are uniquely equipped to deal with cross-industry competition. We combine the CRT data with the Longitudinal Business Database (LBD) (Jarmin and Miranda, 2002), which contains data on each store's employment and allows us to track stores over time.

3.2 Sample Construction

The retail sector is defined based on the North American Industrial Classification System (NAICS) as stores with a 2-digit code of 44 or 45. As such, it includes stores that sell final goods to consumers without performing any transformation of materials. We use the NAICS codes available from the CRT as the industry of each store. The sample includes all stores with positive sales and valid geographic information that appear in official CRT

[9] We use store and establishment as synonyms.

and County Business Patterns statistics that sell one of the product categories used in this study.[10]

Table 1 shows summary statistics for our sample. Even though the number of establishments and firms fluctuates over time, there is an overall decrease in both counts between 1992 and 2012. Notably, the decrease in firms is double that of establishments. This trend is consistent with the growing importance of multi-market firms in rising cross-market concentration that we show in Section 4. Despite these trends, employment increases over time, representing about 9 percent of U.S. employment over the whole sample period.[11]

Table 1: Sample Summary Statistics

	1992	1997	2002	2007	2012
Establishments	908	942	913	912	877
Firms	593	605	589	566	523
Sales	1,004	1,368	1,657	2,062	2,195
Employment	9.91	11.60	11.89	12.78	12.31

Notes: Establishment and firm numbers are expressed in thousands. Sales and employment numbers are expressed in millions. The numbers are based on calculations from the Census of Retail Trade and the Longitudinal Business Database.

3.3 Creation of Product-Level Revenue

We construct product-level revenue data for all U.S. stores, allowing us to assign a store in a given location to markets based on the types of products it carries. To do this, we exploit the CRT's establishment-level data on revenue by product line (e.g., men's footwear, women's pants, diamond jewelry). We then aggregate product line codes into 18 categories

[10]We exclude sales of gasoline and other fuels, autos and automotive parts, and non-retail products because franchising makes it difficult to identify firms. In our main results we exclude non-store retailers because sales from these stores are typically shipped to different markets than their physical location. We explore the implications of this assumption in Section 4.4.

[11]U.S. employment numbers come from Total Nonfarm Employees in the Current Employment Statistics (Bureau of Labor Statistics, 2019).

such that stores in industries outside of general merchandise and non-store retailers sell primarily one type of product.[12] For instance, stores in industries beginning with 448 (clothing and clothing accessory stores) primarily report sales in products such as women's dress pants, men's suits, and footwear, which are grouped into a Clothing category.

Aggregating product lines into categories allows us to accurately impute revenue by category for stores that do not report product-level data. The CRT asks for sales by product lines from all stores of large firms and a sample of stores of small firms. For the remainder, store-level revenue estimates are constructed from administrative data using store characteristics (e.g., industry and multi-unit status), which affects stores that account for 20 percent of sales. Appendix B provides the details of this procedure.

Our product-level revenue data accounts for the presence of multi-product stores. When a store sells products in more than one category, we assign the store's sales in each category to its respective product market. Consequently, a given store faces competition from stores in other industries. For example, an identical box of cereal can be purchased from Walmart (NAICS 452), the local grocery store (NAICS 448), or online (NAICS 454).[13]

Table 2 shows that cross-industry competition is pervasive in retail. On average, the main subsector for each product accounts for just over half of the product's sales. The remaining sales are accounted for by multi-product stores, particularly from the general merchandise and non-store retailer industries, which are included in the appropriate product markets based on their reported sales. The high sales shares of these multi-product stores makes industry classifications problematic when studying competition. Table C.1 reports the composition of sales for each product category, further distinguishing between general merchandisers and other multi-product retailers. Appendix D reports results by industry and compares them to our product measures. Industry-based concentration also increase

[12]Table B.2 lists all the product categories. We will focus on the eight "main" product categories that account for about 82 percent of sales of the stores in our sample for results for individual product categories. The remaining categories are individually small and have not been released due to disclosure limitations.

[13]The authors found a 10.8 oz box of Honey Nut Cheerios at Walmart, Giant Eagle, and Amazon.com on June 22, 2020.

at the national and local level, but the changes across products categories differ from the changes in their main subsector.

Table 2: Share of Product Category Sales by Main Subsector

	1992	2002	2012
Avg. Main Subsector Share	55.8	53.2	50.0
Max Main Subsector Share	79.8	73.1	72.4
Min Main Subsector Share	30.3	27.6	22.0

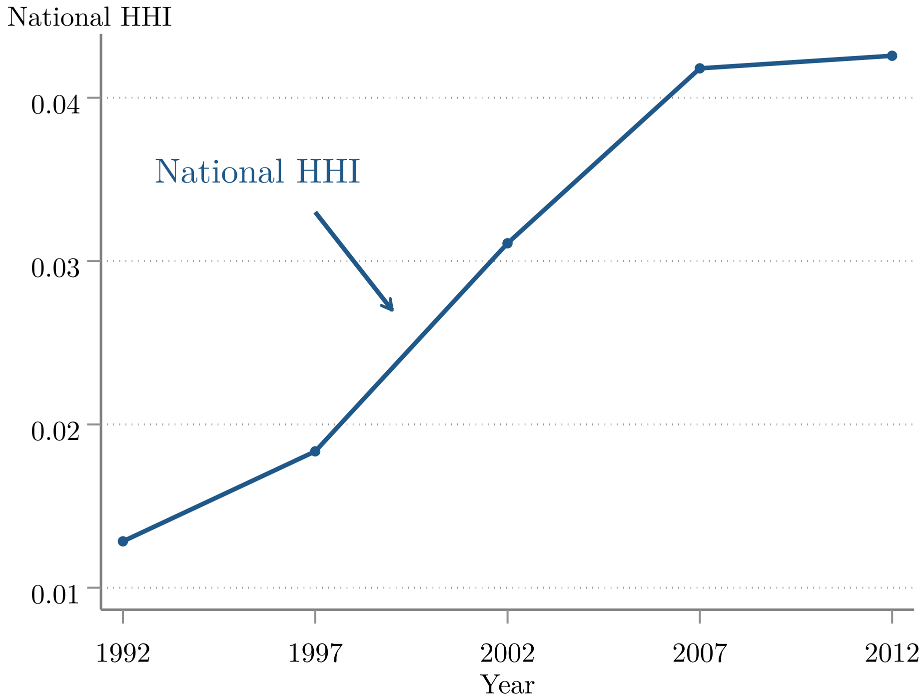
Notes: The numbers are based on calculations from the Census of Retail Trade. The average is the arithmetic mean across the eight main product categories of the share of sales accounted by establishments in the product's associated subsector.

4 Changes in Retail Concentration

In this section, we exploit the detailed microdata described in Section 3 to decompose national concentration in the U.S. retail sector into local and cross-market concentration.We calculate all concentration measures at the firm level by combining sales of the stores of a firm in each market. We show that local concentration has increased, although not as much as national concentration. Moreover, the decomposition reveals that national concentration is largely independent of local trends, with over 99 percent of the growth in national concentration accounted for by increasing cross-market concentration (consumers shopping at the same firms across markets).

Figure 2 plots national concentration in the U.S. retail sector as measured by the HHI defined in equation (1). Between 1992 and 2012, national concentration more than tripled. The probability that two dollars are randomly spent in the same retail firm went from 1.3 percent to 4.3 percent. Most of this increase occurred between 1997 and 2007. In fact, the national HHI was low and grew at a low rate in the years before 1997. In Appendix C we extend our sample to 1982 and show that national concentration increased by 1 percentage point in the 15 years between 1982 and 1997; in contrast it increased 2.3 percentage points in

Figure 2: National Concentration



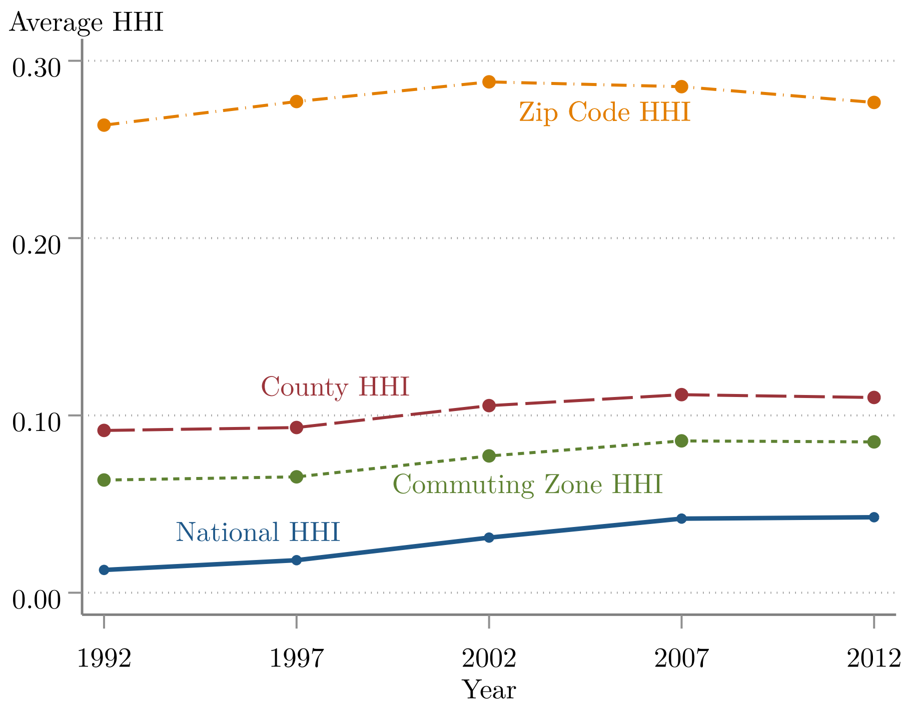
Notes: The numbers are based on calculations from the Census of Retail Trade. The numbers are sales weighted averages of the national Herfindahl-Hirschman Index (HHI) across product categories.

the 10 years between 1997 and 2007. Despite the striking increase in national concentration, Figure 2 provides almost no information on the underlying changes in local retail markets.

Figure 3 plots the level of national and local concentration between 1992 and 2012. Local concentration increases whether markets are defined by zip codes, counties, or commuting zones. Between 1992 and 2012, three of the four measures increased by about 2 percentage points, with the commuting zone HHI increasing by 34 percent from 0.064 to 0.085. But contrary to the national concentration index, local concentration did not accelerate its increase in the period after 1997. When we extend these results back to 1982, we find no change in the trends for local concentration until 2007, when all concentration measures plateau (Appendix C).

The national concentration results are consistent with previous industry-level work

Figure 3: National and Local Concentration



Notes: The numbers are based on calculations from the Census of Retail Trade. The figure plots the Herfindahl-Hirschman Index (HHI) for three different geographic definitions of local markets and national concentration. The local HHI is aggregated using each location's share of national sales within a product category. The numbers are sales weighted averages of the corresponding HHI across product categories.

using sales and employment for various sectors, including retail (Basker et al., 2012; Foster et al., 2016; Lipsius, 2018; Autor et al., 2020; Rinz, 2020; Rossi-Hansberg et al., 2020). The local concentration results are also consistent with studies on local labor market concentration that find increasing concentration in retail but decreasing local concentration overall (Lipsius, 2018; Rinz, 2020). Our results suggest that increasing local retail concentration may help explain the increases in markups documented in De Loecker et al. (2020). However, we show in Section 5 that local concentration implies only modest increases in markups for all product categories.

The picture that emerges from our data differs from the findings at the local level of Rossi-Hansberg et al. (2020), who find that local retail concentration has been steadily falling since 1992. Our results differ for multiple reasons. First, we use a different data set.[14] Second, different definitions of which stores are retailers are employed. Rossi-Hansberg et al. (2020) use Standard Industrial Classification (SIC) codes, while this paper uses NAICS.[15] Finally, the aggregate index of local HHI is calculated differently. Rossi-Hansberg et al. (2020) report the average change in the local HHI, weighting by the end-of-period sales/employment of each market, while we report the change in the average local HHI, weighting markets in each year according to that year's sales. This distinction matters because as markets become bigger, they also tend to become less concentrated. This mechanically gives more weight to markets where concentration is decreasing. The decomposition of Section 2 uses current period weights and avoids this bias. When we repeat our exercise using end-of-period weights, we find slight decreases in local concentration both at the industry and the product level. We expand upon differences between our studies in Appendix E.

4.1 Decomposing National Concentration

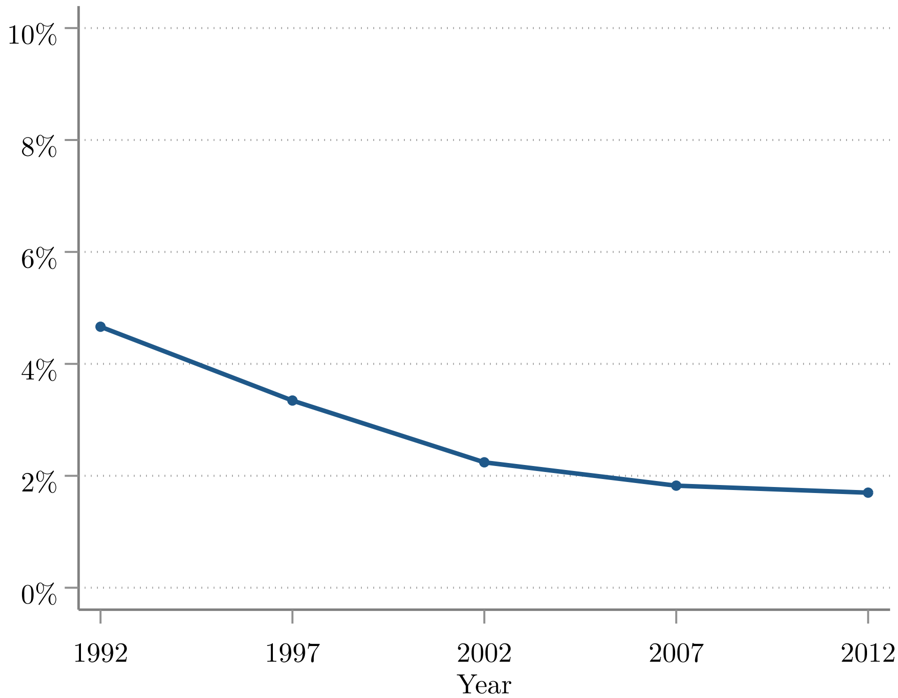
We now assess the contribution of local and cross-market concentration to national concentration, using the decomposition in equation (3). We focus on the 722 commuting zones that partition the contiguous U.S. as our definition of local markets in what follows. Commuting zones are defined by the U.S. Department of Agriculture such that the majority of individuals work and live inside the same one and provide a good approximation for the retail markets in which stores compete.[16] Choosing a larger geographical unit when defining retail markets decreases the level of local concentration and increases the contribution of local concentration to national concentration, relative to smaller geographical units like counties or zip codes.

[14] Rossi-Hansberg et al. (2020) use U.S. National Establishment Time Series (NETS) data.

[15] The primary difference between SIC and NAICS is that SIC includes restaurants in retail.

[16] It seems likely that if individuals live and work in a commuting zone, they do the majority of their shopping in that region. Calculating results in this way causes us to potentially overstate the role of local concentration in national trends relative to using smaller geographic units.

Figure 4: Share of Local Concentration Term in National Concentration



Notes: The numbers are based on calculations from the Census of Retail Trade. The share of local concentration is measured as the ratio of the local concentration term in equation (3) to the national Herfindahl-Hirschman Index (HHI). The local concentration term is the product of the collocation term and local HHI.

Figure 4 shows the contribution of local concentration to national concentration by year. Two things are clear. First, the contribution of local concentration to national concentration is small, never above 5 percent. This is because local concentration is weighted by the collocation term—the probability that two dollars spent in the U.S. are spent in the same market—which is small given the large number of markets in the country.[17] Second, the contribution of local concentration to national concentration has been falling over time as national concentration has been increasing. By 2012, local concentration accounted for just 1.7 percent of the level of national concentration.

The flip side of these results is the major role of cross-market concentration in shaping the national concentration index. National concentration has increased because consumers

[17]The collocation term is stable across time and is always less than 2 percent. See Appendix C.

in different locations are shopping at the same (large) firms; in fact, 99 percent of the change in national concentration is accounted for by changes in cross-market concentration.

4.2 Changes in Concentration across Markets

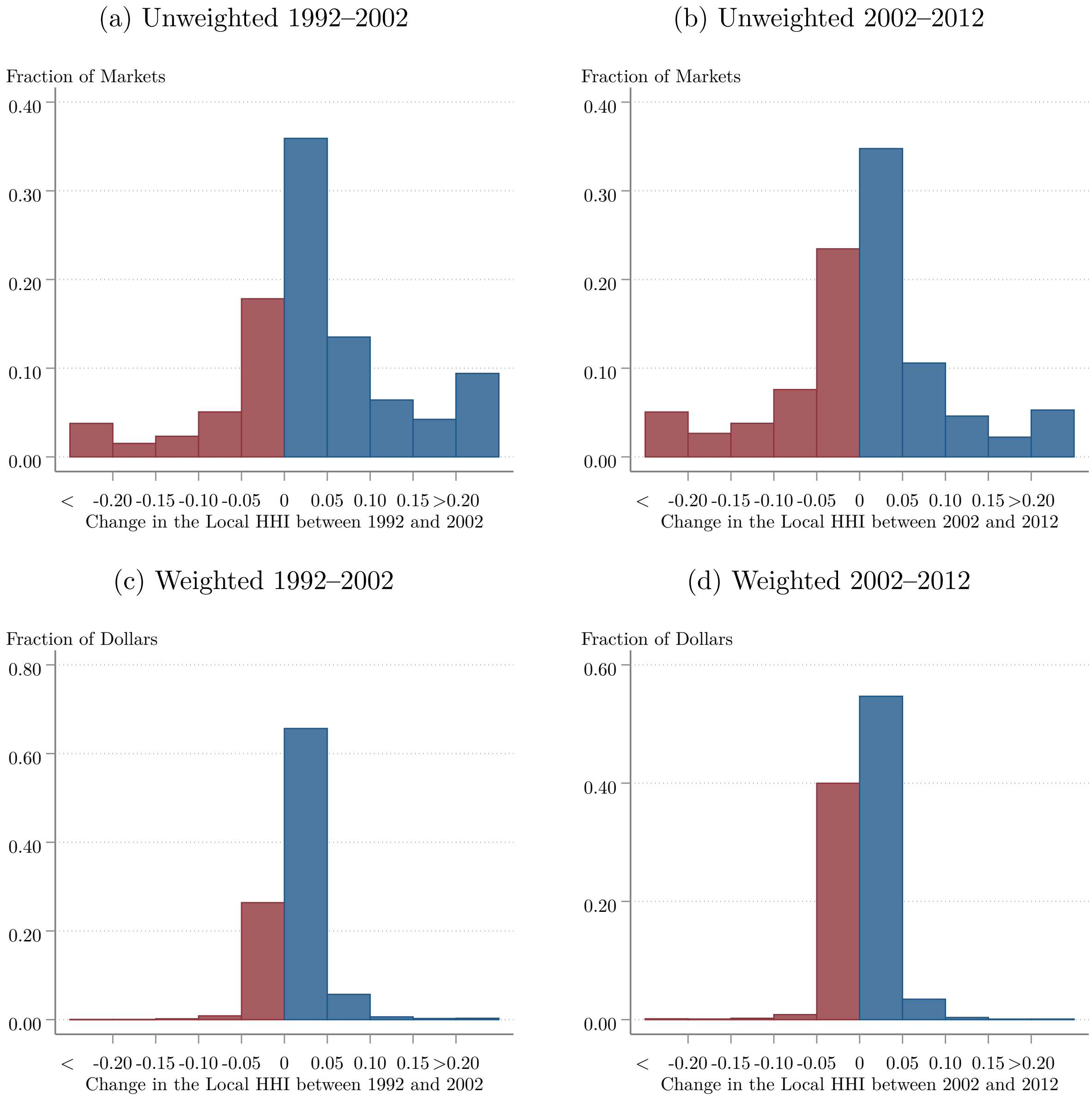
The increases in concentration have been broad based. Almost 60 percent of dollars spent in 2012 are spent in markets that have increased concentration since 2002. Figures 5b and 5d show the distribution of changes in concentration between 2002 and 2012. In just 10 years, 23 percent of markets had increases in concentration of over 5 percentage points. These changes are significant. One criterion used by the Department of Justice to determine when to challenge mergers is whether the local HHI will increase by 2 percentage points (Department of Justice and Federal Trade Commission, 2010).

The increases in local concentration were even more widespread in the decade from 1992 to 2002. Over 69 percent of markets accounting for 72 percent of retail sales increased their concentration. Figures 5a and 5c show the distribution of changes in concentration during this period. Most of the retail sales are concentrated in markets with relatively small increases in concentration (between 0 and 5 percentages point increases in the market's HHI), in both the 1992–2002 decade and the 2002–2012 decade. These markets account for 66 percent of retail sales in 2002 and for 55 percent in 2012.

4.3 Changes in Concentration across Products

Both local and national concentration increased for seven of the eight major product categories between 1992 and 2012, clothing being the exception. Figure 6 shows that these increases were significant for many products. Six of the eight categories had an increase in HHI between 3 and 4 percentage points, even though in 1992 the average level of the local HHI was relatively low; only Toys was above 0.1. Despite this common trend, there is substantial variation across product categories in the changes in concentration. Local concentration in Groceries increased by only 1.1 percentage points, and decreased in

Figure 5: Changes in Concentration across Markets

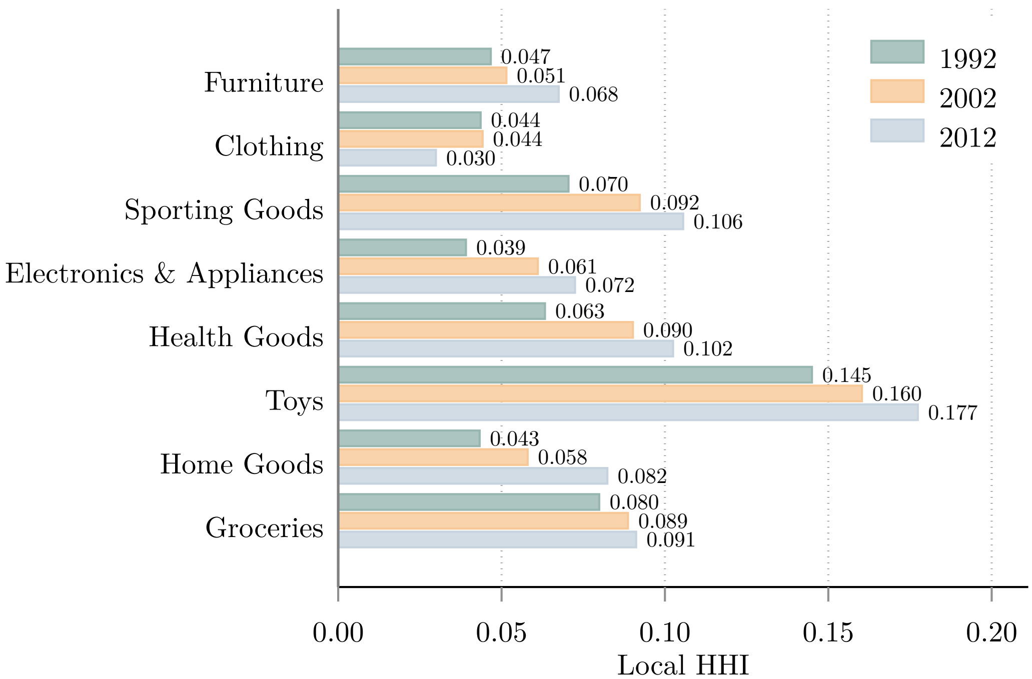


Notes: The numbers are based on calculations from the Census of Retail Trade. The top panels show the fraction of markets, commuting zone/product category pairs, with changes in concentration of a given size. The bottom panels weight markets by the value of sales in the product category. The columns report changes for the decades 1992 to 2002 and 2002 to 2012.

Clothing by 2012, while it almost doubled in Home Goods and Electronics & Appliances.

Figures 7a and 7b show results for national and cross-market concentration and make clear the tight link between these indexes. Both measures show the same patterns and levels because collocation is small and constant across products and time (Figure C.4), leaving little room for local concentration to impact national concentration. The increases in national and cross-market concentration are widespread and significant. Six of the categories experienced larger absolute changes in national and cross-market concentration even though the levels of national and cross-market concentration are markedly lower than those of local concentration.

Figure 6: Local Concentration across Product Categories

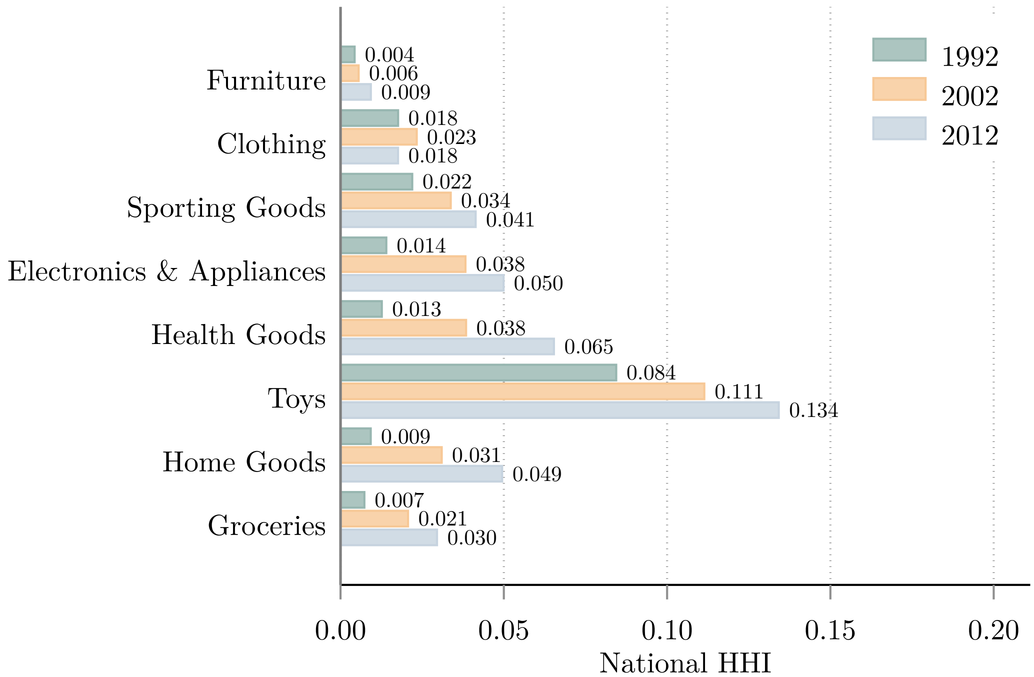


Notes: The numbers are local Herfindahl-Hirschman Indexes (HHI) by product weighted by market size from the Census of Retail Trade microdata.

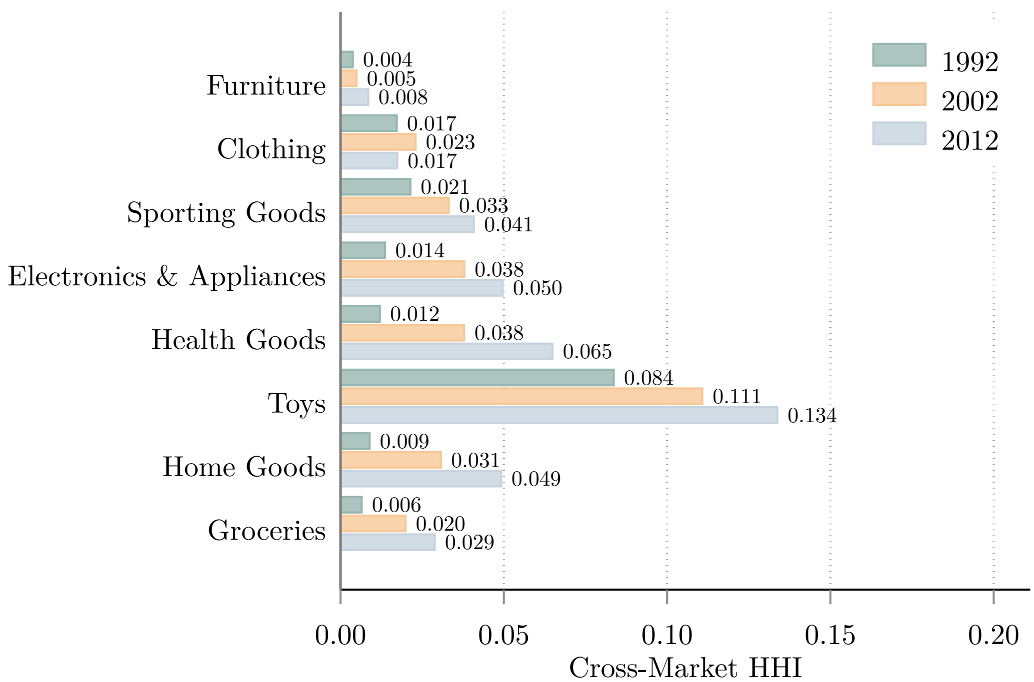
Finally, comparing Figures 6 and 7 shows that not all product markets evolved in the same way between 1992 and 2012. The markets for Furniture and Clothing changed very little, and both have relatively low levels of both local and national concentration. On the other hand, local markets for Groceries and Health Goods have become slightly more concentrated, while at the national level, concentration has increased more than fourfold, driven by increases in cross-market concentration.

Figure 7: National and Cross-Market Concentration across Product Categories

(a) National Concentration



(b) Cross-Market Concentration



Notes: The numbers are national and cross-market Herfindahl-Hirschman Indexes (HHI) by product weighted by market size from the Census of Retail Trade microdata.

4.4 Impact of Online and Other Non-Store Retailers

The previous results calculated local concentration using only brick-and-mortar retailers. In what follows, we consider the potential impact of non-store retailers on local concentration. The market share of non-store retailers has more than tripled between 1992 and 2012. However, the overall importance of non-store retailers remained limited through 2012. The initial sales share of non-store retailers is low, just 2.7 percent in 1992. This low share reflects the absence of online retailers and the limited role of other retailers that rely on mail order and telephone sales. The sales share of non-store retailers had risen to 9.5 percent by 2012, driven by an increase in online sales. The increase was uneven across product categories. Non-store retailers had significant market share in product categories, such as Furniture, Clothing, and Sporting Goods, but almost no market share in Groceries and Home Goods (see Appendix C.3).

The effect of online and other non-store retailers on local concentration depends on how their sales are distributed across and within markets. Unfortunately, the CRT does not record the location in which non-store retailers sell their products, making it impossible to determine the exact effect of these retailers on local concentration. Nevertheless, we can generate bounds for the effect of non-store retailers while being consistent with their behavior at the national level. To do this, we assume that the share of retail spending that goes to non-store retailers is constant across markets within a product category and is equal to the national sales share of non-store retailers in that category.

Having distributed the sales of non-store retailers across markets, we can construct a lower and upper bound for the local HHI. The total effect on concentration depends on the total market share of non-store retailers and how concentrated they are. The lower bound assumes that non-store retailers are atomistic, with the sales share of each non-store retailer equal to zero. The lower bound is

$$\underline{HHI} = (1 - s_{NS})^2 \, HHI_{BM}, \tag{5}$$

where s_{NS} corresponds to the sales share of non-store retailers and HHI_{BM} to the HHI of brick-and-mortar stores. In this case, non-store retailers decrease concentration by reducing the sales share of brick-and-mortar stores. The size of this decrease depends on the sales shares of non-store retailers in the product category. The upper bound assumes that all the sales of non-store retailers belong to a single stand-in firm. The upper bound is

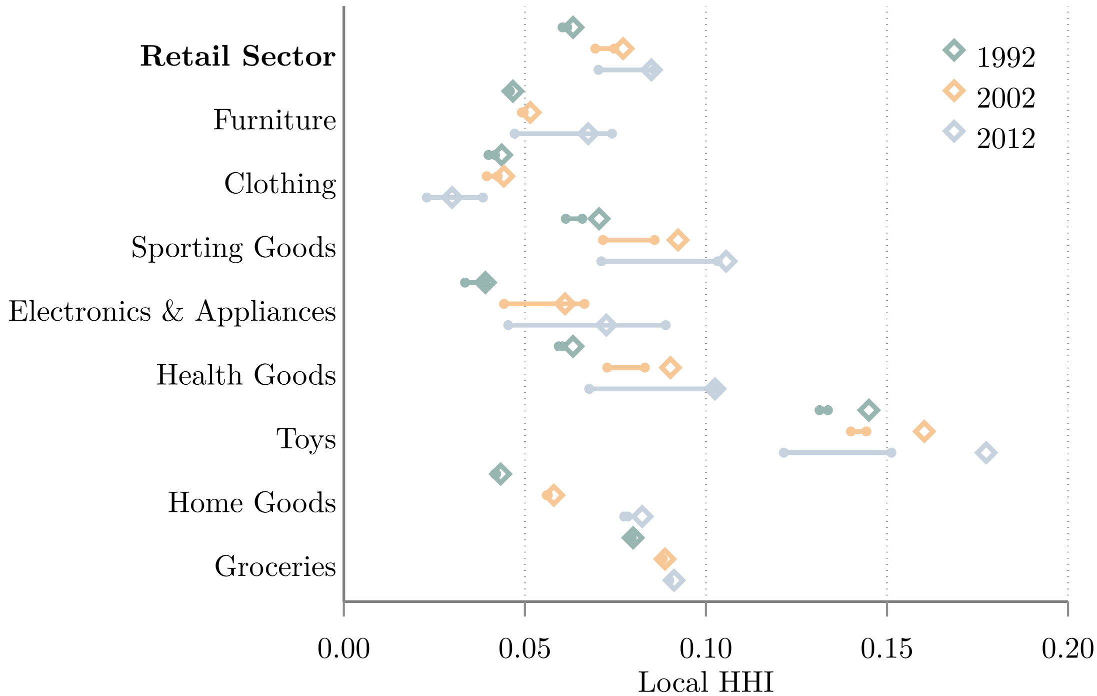
$$\overline{HHI} = (1 - s_{HS})^2 \, HHI_{BM} + s_{NS}^2. \tag{6}$$

This is an upper bound on concentration under the assumption that firms do not have both brick-and-mortar and non-store establishments, which is consistent with the data.

Figure 8 shows the bounds we construct for local concentration across product categories in the retail sector. As expected, including non-store retailers for categories like Home Goods or Groceries hardly affects the level of concentration because the market share of non-store retailers remains low throughout. The effects are larger for the other categories, especially for 2012. The diluting effect of accounting for non-store retailers dominates in most categories, with the bounds for local concentration lying below the estimated HHI for brick-and-mortar stores (marked by the diamonds in the figure). It is only in Electronics & Appliances, and to a lesser extent in Clothing, that the market share of non-store retailers is large enough for their inclusion to potentially increase concentration.

When non-store retailers are included, there is still a clear increase in local concentration between 1992 and 2002, although the levels are slightly lower. Moving from 2002 to 2012, the story becomes ambiguous, especially for product categories with a significant share of their sales going to non-store retailers. In many cases the bounds for 2012 contain the bounds for 2002, indicating that local concentration could either be increasing or decreasing depending on the concentration among non-store sales. At a national level, non-store retailers were not highly concentrated during this time period (Hortaçsu and Syverson, 2015). Thus, the increasing importance of non-store retailers is potentially decreasing local

Figure 8: Local Concentration and Non-Store Retailers



Notes: The numbers are based on calculations from the Census of Retail Trade. Diamonds mark local concentration for brick-and-mortar stores as measured by the Herfindahl-Hirschman Index (HHI) at the commuting zone level. The continuous lines cover the bounds on concentration including non-store retailers. We assume that sales of non-store retailers are distributed across local markets proportionately to the sales of brick-and-mortar retailers. The upper bound assigns all the sales of non-store retailers to a single stand-in firm. The lower bound assumes that non-store retailers are atomistic, with the sales share of each individual non-store retailer equal to zero.

market concentration between 2002 and 2012.

5 Markups and Local Concentration

In the previous sections, we showed that local concentration increased by 2.1 percentage points, on average, between 1992 and 2012. These changes can imply higher markups and ultimately affect consumer prices. However, studying this relationship is challenging because long series on prices and costs for U.S. retailers are not available. Nevertheless, linking changes in concentration to changes in prices is critical to assess the potential impact of concentration on consumers. To deal with data limitations, we use a standard model of Cournot competition based on the work of Atkeson and Burstein (2008) and Grassi (2017). This model provides us with an explicit link between the local HHI and average product markups. We find that increases in local concentration imply a 2.1 percentage point increase in markups between 1992 and 2012, roughly a third of the observed increase in markups during that period.

The model features strong parametric assumptions on firms to maintain tractability. In particular, we assume that 1) firms face isoelastic demand curves, with elasticities of demand varying by product but not by location, 2) firms operate a constant returns to scale technology, and 3) pricing decisions are taken at the market level, ignoring links between stores of the same firm across locations. Under these assumptions, the competitive environment faced by a firm is completely described by the firm's local market share. This allows us to link local concentration, as measured by the local HHI, to prices and markups. In this way, our model is limited by the extent to which the distribution of market shares captures the competitive environment in retail markets. In Appendices F.4.1 and F.4.3 we discuss how to relax some of the assumptions listed above and the effects on our results.

The model economy contains I firms operating in L different locations (representing commuting zones) where J different products are traded. Firms compete in quantities in a non-cooperative fashion and have market power in the local product markets in which they operate.[18] A market is characterized by a pair (j, ℓ) of a product and a location, with an isoelastic demand curve for each product. Firms produce using a constant returns to scale technology and differ only in their productivity, $z_i^{j\ell}$. The firm's marginal cost is $\lambda_i^{j\ell}$. A complete description of the model is in Appendix F.

The solution to each firm's problem is to charge a market-specific markup, $\mu_i^{j\ell}$, over

[18] In Appendix F.1.2 we solve the model for competition in prices and monopolistic competition.

the firm's marginal cost so that the price is $p_i^{j\ell} = \mu_i^{j\ell} \lambda_i^{j\ell}$.[19] The markup is characterized in terms of the firm's market share, $s_i^{j\ell}$, and the product's elasticity of demand, ϵ_j:

$$\mu_i^{j\ell} = \frac{\epsilon_j}{(\epsilon_j - 1)(1 - s_i^{j\ell})}. \tag{7}$$

Markups will be larger for firms with higher market shares and for products with a less elastic demand. Importantly, equation (7) allows us to estimate markups using only data on market shares and elasticities of demand.

The model provides an explicit link between local retail concentration and markups faced by consumers (Grassi, 2017). We use the firm-specific markups in equation (7) to derive closed-form expressions for markups in each market (μ_j^ℓ) as well as for the average markup of each product nationally (μ_j). Appendix F.2 presents the derivations. Both markups directly depend on the local HHI:

$$\mu_j^\ell = \frac{\epsilon_j}{\epsilon_j - 1} \left[1 - HHI_j^\ell\right]^{-1}, \tag{8}$$

$$\mu_j = \frac{\epsilon_j}{\epsilon_j - 1} \left[1 - \sum_{\ell=1}^{L} s_\ell^j HHI_j^\ell\right]^{-1}, \tag{9}$$

where HHI_j^ℓ is the HHI of product j in location ℓ and s_ℓ^j is the share of location ℓ in the national sales of product j. As the local HHI approaches zero, markups approach the Dixit-Stiglitz markups under monopolistic competition. As markets become more concentrated, average markups increase. The sensitivity of markups to increases in concentration is larger for products with a lower elasticity of demand.

[19] Recent work has indicated that firms charge similar and even the same prices across locations in building material (Adams and Williams, 2019) and groceries (Dellavigna and Gentzkow, 2019). Whether the phenomenon holds more broadly is a subject for further research. Appendix F.4.3 shows that uniform pricing depends on a weighted average of local market power. Thus, our assumption of pricing-to-market should have a small effect on aggregate conclusions but may have distributional impacts.

5.1 Estimation and Data

The two key ingredients for analyzing markups are firms' market shares by product in each location, $s_i^{j\ell}$, and the elasticity of substitution for each product, ϵ_j. We obtain the shares directly from the CRT and estimate the elasticities using equation (9). Specifically, we use the product HHIs calculated in Section 4.2 and gross margins by industry from the Annual Retail Trade Survey (ARTS).

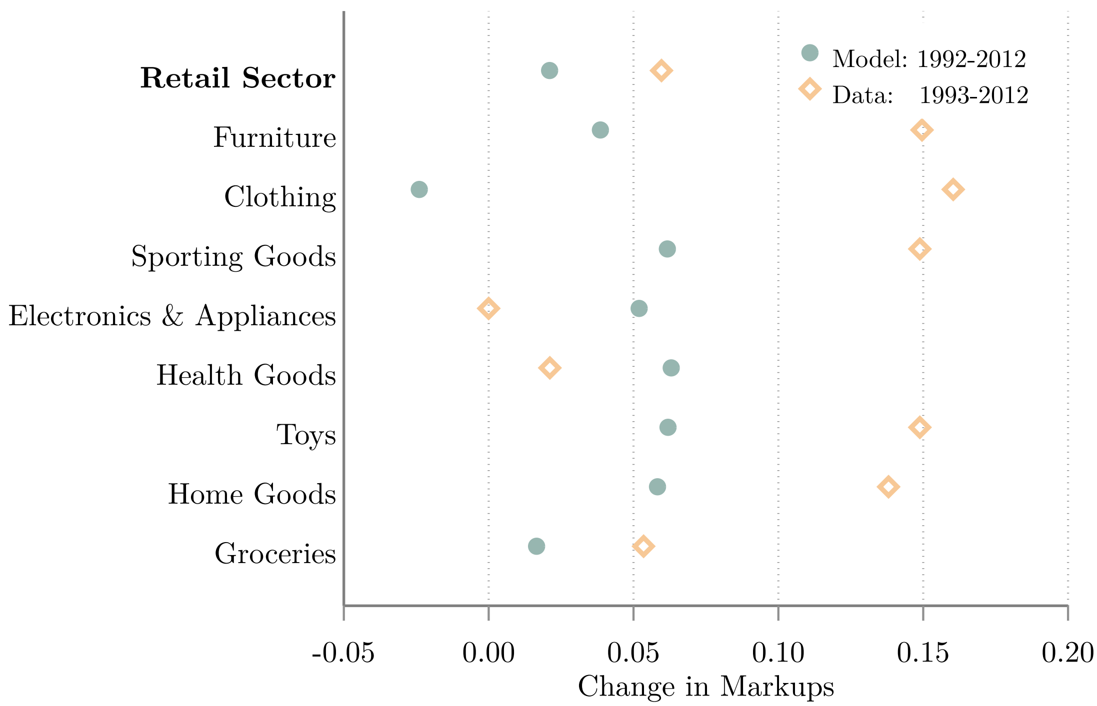
The ARTS provides the best, but not ideal, source to compare our results to because it computes markups using cost of goods sold and reports them for detailed industries. Markups using cost of goods sold are the most direct data analogue to markups in the model, as shown in Appendix F.2. However, there are still issues with comparing the industry-level results in the ARTS to our product-level results. Industry markups and product markups may move in opposite directions due to changes in composition. For example, if low margin clothing stores have been replaced by lower margin general merchandise stores, markups in the clothing industry would rise, while markups on clothing would decrease. Nonetheless, the ARTS allows us to construct markup measures that are informative about product markups by focusing on the industries that specialize in each product category.

5.2 Changes in Concentration and Markups

We conduct two exercises with the model. First, we fit the model to match product markups in 1992 given the observed levels of local concentration. Doing this provides us with estimates of the elasticities of substitution. Holding these estimates fixed, we can extend the model through 2012 and obtain the change in markups implied by the observed increase in local concentration. Second, we can fit the model to match observed markups for each Economic Census year by allowing the elasticities of substitution to be time varying. These exercises give us a measure of the relative role of local concentration in explaining the observed changes in markups.

The increase in local concentration implies an increase in retail markups of 2.1

Figure 9: Local Concentration and Markups



Notes: Diamonds mark the change in markups between 1992 and 2012 from the Annual Retail Trade Survey data for the main industry for each product category and a weighted average across products. Circles mark the change in markups implied by the change in local concentration given the model estimates for 1992.

percentage points between 1992 and 2012, but this falls short of the 6 percentage point increase in markups implied by the ARTS. Figure 9 shows that in all but two product categories, the observed increase in markups in the main industry for the product is higher than what is implied by the rise in product-level HHI. The changes in model markups in Figure 9 assume that the elasticity of demand faced by firms are constant over time and vary only because of changes in local HHI. However, many changes in the competitive environment of retail can be reflected in changes in these elasticities rather than changes in market concentration.

Table 3 shows the value of the elasticity of substitution needed to match the level of markups in each year. We find the lowest elasticities of substitution in Clothing and Furniture. These are categories that feature many different brands only available from a

Table 3: Estimated Elasticities of Substitution

Product Category	ϵ_j 1992	2002	2012
Furniture	2.54	2.32	2.33
Clothing	2.61	2.43	2.25
Sporting Goods	3.17	3.16	2.85
Electronics & Appliances	3.59	4.37	3.96
Health Goods	3.84	4.32	4.18
Toys	3.90	3.82	3.40
Home Goods	4.24	3.78	3.62
Groceries	5.44	4.57	4.87

Notes: The data are authors' estimates of product elasticities of substitution using industry markups from the Annual Retail Trade Survey and product-level local Herfindahl-Hirschman Indexes calculated from the Census of Retail Trade. The elasticities are the solution to equation (9).

small set of retail firms, leaving more room for differentiation than in products such as Toys and Groceries where different firms carry similar or even identical physical products.

To match the observed increase in markups, most product categories require a decrease in their elasticity of substitution. Of course the magnitude of the decrease depends on the initial level of the elasticities as markups respond more to changes for lower elasticities. For example, both Toys and Sorting Goods had an increase in markups of 8.7 percentage points that was not explained by the change in local concentration. For Toys, a decrease in elasticity of 0.5 was needed to explain this residual, while it was only 0.3 for Sporting Goods. The decreasing trend for the elasticities of substitution is consistent with the findings of Bornstein (2018), Brand (2020), and Neiman and Vavra (2020), who link the decrease to the rise of store and brand loyalty/inertia.

The exception to the trend of decreasing elasticities of substitution are Electronics and Appliances and Home Goods, which instead require an increase in their elasticities. Electronics and Appliances had no change in markups in the data, but based on the change in concentration, markups should have increased slightly over 5 percentage points.

This product category is a good example of the limitations of using the ARTS. The main subsector for Electronics and Appliances, 443, accounts for only 31 percent of sales of this product category (Table C.1), leaving significant opportunity for divergence between product markups and industry markups.

Altogether, our results suggest that changes local concentration explain about one-third of the increase in markups, raising them from 1.44 in 1992 to 1.46 in 2012. These increases are small relative to the 34 percent decrease in the relative price of retail goods during this period. The increases in markups and concentration may be the result of low-cost firms gaining market share, in which case the decrease in prices cannot be separated from the increase in concentration. Even if the implicit reduction in costs is realized without an increase in concentration, the decrease in prices would have been 35 percent.

6 Conclusion

Despite the attention given to the rise of national concentration in the U.S., less is known about the dynamics of local concentration and the relationship between observed national trends and the behavior of local markets. This paper helps to shed light on these issues by contributing in three related fronts. First, we decompose national concentration measures into a local component (national concentration rises as local markets become more concentrated), and a cross-market component (national concentration rises as the same firms are present in more markets, increasing their national market share). Second, we measure concentration at a granular level by compiling new Census microdata covering all U.S. retailers. Third, we estimate a model of oligopolistic competition that features an explicit link between the local HHI and markups to quantify the effect of concentration on retail markups.

We show that local concentration has almost no effect on national concentration measures. Instead, cross-market concentration explains most of the increase in national concentration observed since 1992. That is, national concentration is driven by consumers in different locations shopping at the same firms, highlighting the role of large multi-market retailers in explaining the dynamics of the retail sector.

Our measures of local concentration show broad increases across locations and products since the 1990s, although they are at lower rates than the increases in national concentration. We link these changes to markups and find that they explain roughly a third of the increase in markups observed in the retail industries associated to our main product categories. On the other hand, the large increases in national concentration that we document likely reflect the cost advantages enjoyed by large retailers. These cost advantages may be due to direct foreign sourcing (Smith, 2019), negotiating power with suppliers (Benkard et al., 2021), or investments in information and communication technologies (Hsieh and Rossi-Hansberg, 2019). These forces help explain why the relative price of retail goods has fallen by one-third between 1992 and 2012. Together our results suggest that the changes in product concentration in the retail sector have likely benefited consumers.

References

Adams, Brian and Kevin R. Williams, 2019. "Zone Pricing in Retail Oligopoly." *American Economic Journal: Microeconomics* 11(1):124–56, https://dx.doi.org/10.1257/mic.20170130.

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J. Klenow, and Huiyu Li, 2019. "A Theory of Falling Growth and Rising Rents." *National Bureau of Economic Research Working Paper Series* No. 26448, https://dx.doi.org/10.3386/w26448.

Arcidiacono, Peter, Patrick Bayer, Jason R Blevins, and Paul B Ellickson, 2016. "Estimation of Dynamic Discrete Choice Models in Continuous Time with an Application to Retail Competition." *Review of Economic Studies* 83(3):889–931, https://dx.doi.org/10.1093/restud/rdw012.

Atkeson, Andrew and Ariel Burstein, 2008. "Pricing-to-Market, Trade Costs, and International Relative Prices." *American Economic Review* 98(5):1998–2031, https://dx.doi.org/10.1257/aer.98.5.1998.

Autor, David H., David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen, 2020. "The Fall of the Labor Share and the Rise of Superstar Firms." *The Quarterly Journal of Economics* 135(2):645–709, https://dx.doi.org/10.1093/qje/qjaa004.

Azar, José, Steven T. Berry, and Ioana Elena Marinescu, 2019. "Estimating Labor Market Power." https://dx.doi.org/10.2139/ssrn.3456277 (accessed 2020-03-20).

Barnatchez, Keith, Leland Dod Crane, and Ryan A. Decker, 2017. "An Assessment of the National Establishment Time Series (Nets) Database." *Board of Governors of the Federal Reserve System: Finance and Economics Discussion Series* No. 2017-1, https://dx.doi.org/10.17016/FEDS.2017.110.

Basker, Emek, 2005. "Job Creation or Destruction? Labor Market Effects of Wal-Mart Expansion." *The Review of Economics and Statistics* 87(1):174–183, https://dx.doi.org/10.1162/0034653053327568.

Basker, Emek, Shawn D Klimek, and Pham Hoang Van, 2012. "Supersize It: The Growth of Retail Chains and the Rise of the "Big-Box" Store." *Journal of Economics and Management Strategy* 21(3):541–582, https://dx.doi.org/10.1111/j.1530-9134.2012.00339.x.

Benkard, C Lanier, Ali Yurukoglu, and Anthony Lee Zhang, 2021. "Concentration in Product Markets." https://anthonyleezhang.github.io/pdfs/concentration.pdf (accessed 2021-03).

Berger, David W, Kyle F Herkenhoff, and Simon Mongey, 2019. "Labor Market Power." *National Bureau of Economic Research Working Paper Series* No. 25719, https://dx.doi.org/10.3386/w25719.

Bornstein, Gideon, 2018. "Entry and Profits in an Aging Economy: The Role of Consumer Inertia." https://www.dropbox.com/s/8ud1oukigjqd9i6/JMP_Bornstein.pdf (accessed 2020-12-15).

Brand, James, 2020. "Differences in Differentiation : Rising Variety and Markups in Retail Food Stores." https://www.jamesbrandecon.com/s/Brand_JMP_Oct10.pdf (accessed 2020-12-15).

Bresnahan, Timothy F., "Empirical studies of industries with market power." In *Handbook of Industrial Organization*, vol. 2, chap. 17, pp. 1011–1057. https://dx.doi.org/10.1016/S1573-448X(89)02005-4.

Bureau of Labor Statistics, 2019. "Current Employment Statistics." https://www.bls.gov/ces/ (accessed 2020-03-20).

Cao, Dan, Henry R. Hyatt, Toshihiko Mukoyama, and Erick Sager, 2019. "Firm Growth Through New Establishments." https://dx.doi.org/10.2139/ssrn.3361451 (accessed 2019-09-23).

Crane, Leland Dod and Ryan A. Decker, 2020. "Research with private sector business microdata: The case of NETS/D&B." https://conference.nber.org/conf_papers/f142811.pdf (accessed 2020-01-04).

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger, 2020. "The Rise of Market Power and the Macroeconomic Implications." *The Quarterly Journal of Economics* 135(2):561–644, https://dx.doi.org/10.1093/qje/qjz041.

Decker, Ryan A., John C Haltiwanger, Ron S. Jarmin, and Javier Miranda, 2014. "The Role of Entrepreneurship in US Job Creation and Economic Dynamism." *Journal of Economic Perspectives* 28(3):3–24, https://dx.doi.org/10.1257/jep.28.3.3.

———, 2020. "Changing Business Dynamism and Productivity: Shocks vs. Responsiveness." *American Economic Review* 110(12):3952–90, https://dx.doi.org/10.1257/aer.20190680.

Dellavigna, Stefano and Matthew Gentzkow, 2019. "Uniform Pricing in U.S. Retail Chains." *The Quarterly Journal of Economics* 134(4):2011–2084, https://dx.doi.org/10.1093/qje/qjz019.

Department of Justice and Federal Trade Commission, 2010. "Horizontal Merger Guidelines." Tech. rep., Department of Justice, Federal Trade Commission. https://www.justice.gov/atr/file/810276/download (accessed 2019-09-23).

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, 2019. "How Costly Are Markups?" http://www.virgiliumidrigan.com/uploads/1/3/9/8/13982648/how-costly-markups-01-14-19.pdf (accessed 2019-09-23).

Fort, Teresa C and Shawn D Klimek, 2018. “The Effects of Industry Classification Changes on US Employment Composition.” *U.S. Census Bureau Center for Economic Studies Working Papers* No. 18-28, https://www.census.gov/library/working-papers/2018/adrm/ces-wp-18-28.html.

Foster, Lucia, John C Haltiwanger, Shawn D Klimek, C.J. Krizan, and Scott Ohlmacher, “The Evolution of National Retail Chains: How We Got Here.” In *Handbook on the Economics of Retailing and Distribution*, chap. 1, pp. 7–37. https://dx.doi.org/10.4337/9781783477388.

Ganapati, Sharat, 2020. “Growing Oligopolies, Prices, Output, and Productivity.” *American Economic Journal: Microeconomics* Forthcoming, https://www.aeaweb.org/articles/pdf/doi/10.1257/mic.20190029.

Grassi, Basile, 2017. “IO in I-O: Size, Industrial Organization, and the Input-Output Network Make a Firm Structurally Important.” https://sites.google.com/site/grassibasile/home/research/job-market-paper (accessed 2020-04-23).

Grullon, Gustavo, Yelena Larkin, and Roni Michaely, 2019. “Are US Industries Becoming More Concentrated?” *Review of Finance* 23(4):697–743, https://dx.doi.org/10.1093/rof/rfz007.

Hall, Robert E, 2018. “New Evidence of the Markup of Prices Over Marginal Costs and the Role of Mega-Firms in the US Economy.” *National Bureau of Economic Research Working Paper Series* No. 24574, https://dx.doi.org/10.3386/w24574.

Haltiwanger, John C, Ron S. Jarmin, and C. J. Krizan, 2010. “Mom-and-Pop Meet Big-Box: Complements or Substitutes?” *Journal of Urban Economics* 67(1):116–134, http://www.nber.org/papers/w15348https://doi.org/10.1016/j.jue.2009.09.003.

Hortaçsu, Ali and Chad Syverson, 2015. “The Ongoing Evolution of US Retail: A Format Tug-of-War.” *Journal of Economic Perspectives* 29(4):89–112, https://dx.doi.org/10.1257/jep.29.4.89.

Hsieh, Chang-Tai and Esteban Rossi-Hansberg, 2019. “The Industrial Revolution In Services.” *National Bureau of Economic Research Working Paper Series* No. 25968, https://dx.doi.org/10.3386/w25968.

Jarmin, Ron S. and Javier Miranda, 2002. “Data for: The Longitudinal Business Database (1976-2012).”, Federal Statistical Research Data Center [distributor]. https://www2.census.gov/library/working-papers/2002/adrm/ces-wp-02-17.pdf (accessed 2020-03).

Jarosch, Gregor, Jan Sebastian Nimczik, and Isaac Sorkin, 2020. “Granular Search, Market Structure, and Wages.” *National Bureau of Economic Research Working Paper Series* No. 26239, https://dx.doi.org/10.3386/w26239.

Jia, Panle, 2008. "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry." *Econometrica* 76(6):1263–1316, https://dx.doi.org/10.3982/ECTA6649.

Lipsius, Ben, 2018. "Labor Market Concentration Does Not Explain the Falling Labor Share." https://ideas.repec.org/p/jmp/jm2018/pli1202.html (accessed 2019-09-23).

Neiman, Brent and Joseph S. Vavra, 2020. "The Rise of Niche Consumption." *National Bureau of Economic Research Working Paper Series* No 26134, https://dx.doi.org/10.3386/w26134.

Rinz, Kevin, 2020. "Labor Market Concentration, Earnings , and Inequality." *Journal of Human Resources* Forthcoming, https://dx.doi.org/10.3368/jhr.monopsony.0219-10025R1.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter, "Diverging Trends in National and Local Concentration." In **Martin Eichenbaum and Erik Hurst**, editors, *NBER Macroeconomics Annual 2020, volume 35*, University of Chicago Press. http://www.nber.org/chapters/c14475.

Securities and Exchange Commission, 2020. "Walmart Inc. EDGAR Results." https://www.sec.gov/cgi-bin/browse-edgar?action=getcompany&CIK=0000104169&owner=exclude&count=40&hidefilings=0 (accessed 2020-01-31).

Smith, Dominic, 2019. "Concentration and Foreign Sourcing in the U.S. Retail Sector." https://dominic-smith.com/smith_jmp2.pdf (accessed 2021-03).

Traina, James, 2018. "Is Aggregate Market Power Increasing? Production Trends Using Financial Statements." *Stigler Center for the Study of the Economy and the State: New Working Paper Series* No. 17, https://research.chicagobooth.edu/-/media/research/stigler/pdfs/workingpapers/17isaggregatemarketpowerincreasing.pdf.

U.S. Bureau of Economic Analysis, 2020. "Personal Consumption Expenditure Table 2." https://apps.bea.gov/iTable/iTable.cfm?ReqID=19&step=2#reqid=19&step=2&isuri=1&1921=underlying (accessed 2020-05-18).

U.S. Census Bureau, 1992-2012. "Census of Retail Trade - Base and Lines Files.", Federal Statistical Research Data Center [distributor]. https://www.census.gov/programs-surveys/ces/data/restricted-use-data/economic-data.html (accessed 2020-03).

Appendices

Table of Contents

A Concentration Decomposition

The Herfindahl-Hirschman Index for the retail sector is given by the sales-weighted average of the product-HHI:

$$HHI^t \equiv \sum_{j=1}^{J} s_j^t HHI_j^t. \tag{A.1}$$

The HHI for a given product can be decomposed into the contribution of local and cross-market concentration. This section provides additional details on the concentration decomposition. The decomposition starts from the probability that two dollars (x, y) are spent at the same firm (i), which gives the HHI at the national level:

$$HHI_j^t \equiv P(i_x = i_y; j, t) = \sum_{\ell=1}^{L} \sum_{i} \left(s_i^{j\ell t}\right)^2. \tag{A.2}$$

This probability can be divided into two terms:

$$\begin{aligned} P(i_x = i_y; j, t) = & \underbrace{\overbrace{P(i_x = i_y | \ell_x = \ell_y; j, t)}^{\text{Local Concentration}} \overbrace{P(\ell_x = \ell_y; j, t)}^{\text{Collocation}}}_{\textbf{Local Term}} \\ & + \underbrace{\overbrace{P(i_x = i_y | \ell_x \neq \ell_y; j, t)}^{\text{Cross-Market Concentration}} \overbrace{P(\ell_x \neq \ell_y; j, t)}^{\text{1 - Collocation}}}_{\textbf{Cross-Market Term}} \end{aligned} \tag{A.3}$$

When we report contribution of local and cross-market concentration for the retail sector, we report the sales-weighted average of these two terms across products.

The collocation probability is calculated as:

$$P(\ell_x = \ell_y; j, t) = \sum_{\ell}^{L} \left(s_\ell^{jt}\right)^2. \tag{A.4}$$

When we report the collocation for the retail sector, we report the sales-weighted average of collocation across prodcuts: $\text{Collocation}_t = \sum_j s_j^t P(\ell_x = \ell_y; j, t)$.

Local concentration is calculated as:

$$
\begin{aligned}
P\left(i_x = i_y | \ell_x = \ell_y; j, t\right) &= \sum_{\ell=1}^{L} \underbrace{P\left(\ell_x = \ell | \ell_x = \ell_y; j, t\right)}_{\text{Location Weights}} \overbrace{P\left(i_x = i_y | \ell_x = \ell, \ell_x = \ell_y; j, t\right)}^{\text{Local HHI}} \\
&= \sum_{\ell=1}^{L} \frac{(s_\ell^{jt})^2}{\sum_n (s_n^{jt})^2} \sum_{k=1}^{K} \left(s_k^{j\ell t}\right)^2 \qquad \text{(A.5)}
\end{aligned}
$$

This probability can be further decomposed into a term due to the average number of firms in each market (location) and a term due to the inequality of shares across firms within a market:

$$
\begin{aligned}
P\left(i_x = i_y | \ell_x = \ell_y; j, t\right) &= \sum_{\ell=1}^{L} s_\ell^{jt} \left(\frac{1}{N_\ell} + \sum_{k \in K_\ell} \left(s_k^{j\ell t} - \frac{1}{N_\ell} \right)^2 \right) \\
&= \underbrace{\sum_{\ell=1}^{L} s_\ell^{jt} \frac{1}{N_\ell}}_{\text{Average Number of Firms}} + \sum_{\ell=1}^{M} s_\ell^{jt} \sum_{i \in K_\ell}^{K} \underbrace{\left(s_k^{j\ell t} - \frac{1}{N_\ell} \right)^2}_{\text{Inequality of shares}}
\end{aligned}
$$

When we report the local HHI for individual product categories we also report the retail sector's average local HHI using sales weights instead of the weights implied by the decomposition to facilitate comparison to other research such as Rinz (2020) and Lipsius (2018):

$$
\mathrm{HHI}_t^{\mathrm{Local}} = \sum_{j}^{J} s_j^t \sum_{\ell}^{L} s_\ell^{jt} \sum_{i} \left(s_i^{j\ell t}\right)^2 \qquad \text{(A.6)}
$$

The cross-market term is calculated as:

$$
\begin{aligned}
P\left(\ell_x = \ell_y; j, t\right) P\left(i_x = i_y | \ell_x \neq \ell_y; j, t\right) &= (1 - \sum_{\ell=1}^{L} \left(s_\ell^{jt}\right)^2) \sum_{k=1}^{L} \sum_{\ell \neq k} \frac{s_k^{jt} s_\ell^{jt}}{1 - \sum_m^L \left(s_m^{jt}\right)^2} \sum_{i=1}^{I} s_i^{jkt} s_i^{\ell jt} \\
&= \sum_{k=1}^{L} \sum_{\ell \neq k} s_k^{jt} s_\ell^{jt} \sum_{i=1}^{I} s_i^{jkt} s_i^{j\ell t}.
\end{aligned}
$$

This calculation is the same in the results for product category because $1 - \sum_m^L (s_m^{jt})^2$ cancels in the calculation of the collocation term.

B Cleaning and Aggregating Product Lines Data

The Economic Census collects data on establishment-level sales in a number of product categories (Figure B.1 provides an example form). Many establishments have missing product line sales either due to them not responding to questions or because they do not receive a form.[20] In total, reported product lines data account for about 80 percent of sales. We develop an algorithm to impute data for missing establishments, which involves aggregating product line codes into categories such that we can accurately infer each establishment's sales by category with available information. For example, we aggregate lines for women's clothes, men's clothes, children's clothes, and footwear into a product category called clothing.

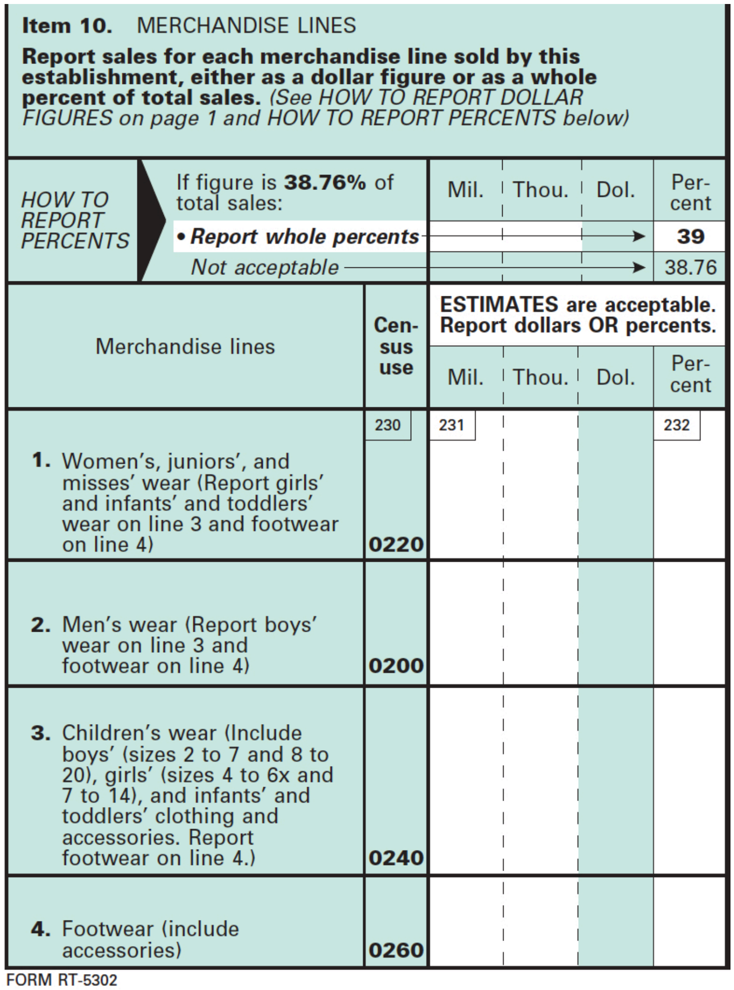
We then establish 18 product categories detailed in Table B.1. Of these 18 product categories, 8 categories that we label "Main" account for over 80 percent of store sales in the sample. The other 10 product categories are specialty categories that account for a small fraction of aggregate sales and are sold primarily by establishments in one specific industry. For example, glasses are sold almost exclusively by establishments in 446130 (optical goods stores). We create these categories so that establishments that sell these products are not included in concentration measures for the 8 main product categories.

B.1 Aggregating Product Lines

The first step of cleaning the data is to aggregate reported broad and detailed product line codes into categories. Some codes reported by retailers do not correspond to valid product line codes, and we allocate those sales to a miscellaneous category. The Census analyzes reported product line codes to check for issues and flags observations as usable if they pass this check. We include only observations that are usable and then map these codes to categories. We use the reported percentage of total sales accounted for by each product line instead of the dollar value because the dollar value is often missing. Typically

[20] Establishments of large firms are always mailed a form, but small firms are sampled.

Figure B.1: Sample Product Lines Form


Item 10. MERCHANDISE LINES

Report sales for each merchandise line sold by this establishment, either as a dollar figure or as a whole percent of total sales. *(See HOW TO REPORT DOLLAR FIGURES on page 1 and HOW TO REPORT PERCENTS below)*

HOW TO REPORT PERCENTS	If figure is **38.76%** of total sales:	Mil.	Thou.	Dol.	Per-cent
	• ***Report whole percents***				**39**
	Not acceptable				38.76

Merchandise lines	Cen-sus use	ESTIMATES are acceptable. Report dollars OR percents.			
		Mil.	Thou.	Dol.	Per-cent
1. Women's, juniors', and misses' wear (Report girls' and infants' and toddlers' wear on line 3 and footwear on line 4)	230 **0220**	231			232
2. Men's wear (Report boys' wear on line 3 and footwear on line 4)	**0200**				
3. Children's wear (Include boys' (sizes 2 to 7 and 8 to 20), girls' (sizes 4 to 6x and 7 to 14), and infants' and toddlers' clothing and accessories. Report footwear on line 4.)	**0240**				
4. Footwear (include accessories)	**0260**				

FORM RT-5302

an establishment either reports product line data for 100 percent of its sales or does not report any data. For the small number of establishments that report product lines data summing to a number other than 100 percent, we rescale the percentages so that they sum to one.[21] After this procedure, we have sales by product category for all establishments that reported lines data. The resulting categories are listed in Table B.1.

[21] This procedure has a minimal effect on aggregate retail sales in each category.

Table B.1: List of Product Categories

Product Category	Main	Corresponding Industry	Example Firm
Automotive Goods	N	441	Ford Dealer
Clothing	Y	448	Old Navy
Electronics and Appliances	Y	443	Best Buy
Furniture	Y	442	Ikea
Services	N	N/A	
Other Retail Goods	N	N/A	
Groceries	Y	445	Trader Joe's
Health Products	Y	446	CVS
Fuel	N	447	Shell Gasoline
Sporting Goods	Y	451	Dick's Sporting Goods
Toys	Y	451	Toys "R" Us
Home & Garden	Y	444	Home Depot
Paper Products	N	453210	
Jewelry	N	423940	Jared
Luggage	N	448320	Samsonite
Optical Goods	N	446130	Lenscrafters
Non-Retail Goods	N	N/A	
Books	N	451211	Borders

Notes: Authors' created list of product categories. The Main column indicates that a product category is included in concentration calculations. Firm names were created for illustrative purposes based on industries reported to the Securities and Exchange Commission and do not imply that the firm is in the analytical sample.

B.2 Imputing Missing Data

For the remaining establishments, we impute data using the NAICS code of the establishment, reported sales of other establishments of the same firm in the same industry, and reported activity of the same establishment in other census years.[22] Most establishments are part of single-unit firms, and many do not appear in multiple census years; thus their sales are imputed using only industry information.

Using this aggregation method, almost all establishments have significant sales in only two product categories, which increases confidence in the imputation. Additionally, we

[22]Reported product line sales are very similar across establishments of the same firm and the same establishment over time.

have compared the aggregate sales in our data to the Consumer Expenditure Survey (an independent Bureau of Labor Statistics program), and they are in line with the numbers from that source.[23]

Where relevant, all sales are deflated using consumer price indexes. We use the food deflator for Groceries, clothing and apparel deflator for Clothing, and the deflator for all goods excluding food and fuel for all other categories.

We find that this procedure predicts sales accurately for most establishments, but a small number of stores in each industry report selling very different products than all other stores in that industry. In these cases, the prediction can produce substantial error.

[23] Retail sales include some sales to companies, so it is expected that retail sales in a product category exceed consumer spending on that category.

C Additional Tables and Figures

C.1 The Role of Multi-Product Retailers

Table C.1 shows how sales for each main product category are distributed across sets of industries. This informs us of which type of establishment accounts for the sales of each product. The main subsector column refers to the NAICS subsector that most closely corresponds to the product category. The NAICS code of the subsector is indicated next to each product category. The main subsector accounts for just over half of sales on average, but this figure varies depending on the product. A larger fraction of sales of Furniture, Home Goods, and Groceries comes from establishments in their respective NAICS subsectors, while Electronics and Toys are more commonly sold by establishments in other subsectors. Over time, the share of sales accounted by the product's own subsector has decreased for most products, with the difference captured by establishments outside of the general merchandise subsector.

Table C.1: Share of Product Category Sales by Establishment Subsector

	Main Subsector			GM			Other		
	1992	2002	2012	1992	2002	2012	1992	2002	2012
Furniture (442)	76.3	73.1	64.4	16.9	13.3	11.2	6.8	13.6	24.4
Clothing (448)	50.9	51.8	51.1	41.4	37.7	27.4	7.7	10.5	21.5
Sporting Goods (451)	55.4	52.3	54.2	30.7	29.1	21.2	14.0	18.7	24.6
Electronic & Appliances (443)	30.3	31.0	29.5	34.1	27.1	24.9	35.6	41.9	45.6
Health Goods (446)	49.0	50.0	46.8	19.0	21.3	20.5	32.0	28.7	32.6
Toys (451)	40.7	27.6	22.0	45.2	47.7	46.9	14.1	24.7	31.1
Home Goods (444)	63.9	72.8	72.4	17.2	11.6	10.9	18.9	15.6	16.6
Groceries (445)	79.8	67.2	59.7	6.6	16.2	22.8	13.6	16.6	17.5
Average	55.8	53.2	50.0	26.4	25.5	23.2	17.8	21.3	26.8

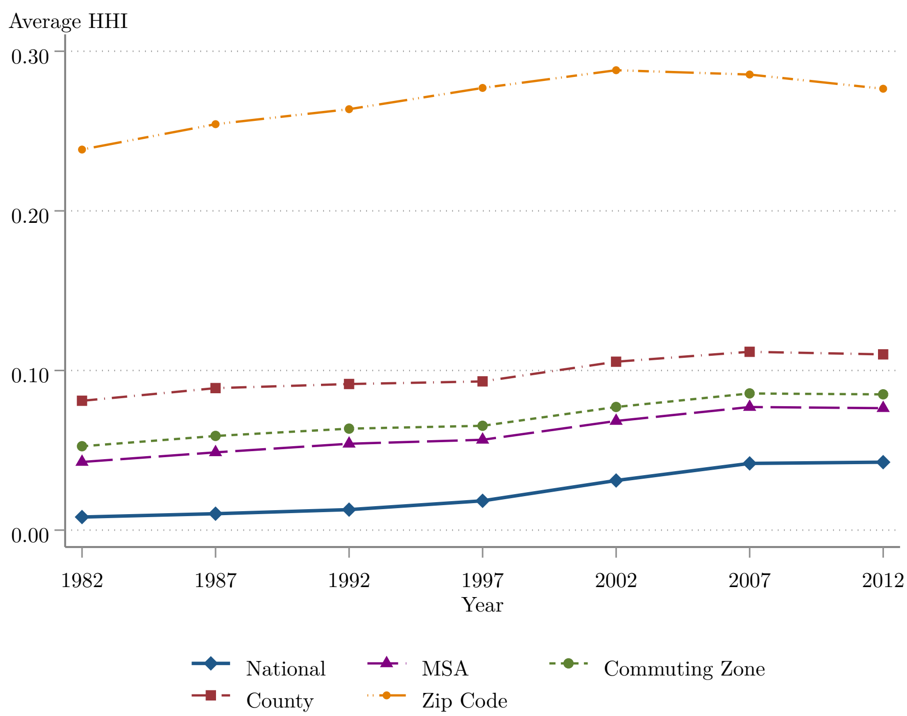
Notes: The numbers come from the Census of Retail Trade data. GM includes stores in subsector 452. Other includes sales outside of the main subsector (indicated in parenthesis) and GM. Average is the arithmetic mean of the numbers in the column.

C.2 Extended Sample

We now present results with an extended sample that covers the period 1982 to 2012. The 1982 and 1987 Censuses of Retail Trade do not include product-level sales for all the categories we consider in our main sample (1992-2012). The affected product categories, Toys and Sporting Goods, account for a relatively small share of total retail sales. Therefore, we focus on results for the retail sector as a whole which we believe are reliable for this time period.

Figure C.1 presents measured concentration indexes for different definitions of local markets and the retail sector as a whole going back to 1982. We use the store-level NAICS codes imputed by Fort and Klimek (2018) to identify retail establishments prior to 1992.

Figure C.1: National and Local Concentration



Notes: The data are from the Census of Retail Trade. The Herfindahl-Hirschman Index (HHI) for four different geographic definitions of local markets and national concentration are plotted. The local HHI is aggregated using each location's share of national sales within a product category. The numbers are sales weighted averages of the corresponding HHI in the product categories.

Relative to Figure 3 we also include a measure of local concentration where markets are defined by Metropolitan Statistical Areas (MSA). There are more MSAs than commuting zones (about 900 vs 722) and MSAs do not partition the U.S., omitting rural areas. In practice, the measured concentration level for MSAs is similar to that of commuting zones and only slightly lower.

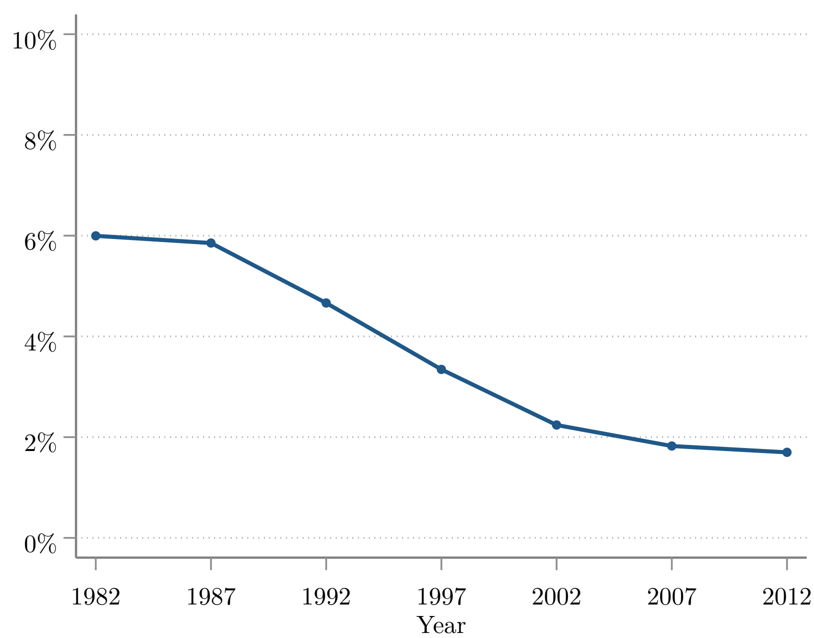
Extending our sample to 1982 does not change the main result of increasing national and local concentration. All measures show sustained increases between 1982 and 2002. Looking at the full sample highlights the change in the rate of increase of national concentration after 1997 which contrasts with the slow increase during the 1980s.

Finally, we extend the decomposition exercise of Figure 4 to 1982. The results, shown in Figure C.2, show a stark decrease in the contribution of local concentration to national concentration. Even though the role of local concentration was never large (always below 6 percent), the share of national concentration attributed to local concentration fell sharply during the 1990s, ending at roughly 2 percent in 2002.

C.3 Non-Store Retailer Market Shares

The penetration of non-store retailers varies widely across products. As Figure C.3 shows, the sales share of non-store retailers is highest in Electronics and Appliances, with an initial share of 7.5 percent in 1992 and a share of 20.9 percent in 2012. The initial differences were large, with only two categories (Electronics and Sporting Goods) having a share of more than 5 percent. By 2012, non-store retailers accounted for more than 15 percent of sales in five of the eight major categories. Despite this widespread increase, not all products are sold online. By 2012, only 0.7 percent of Groceries sales and 3 percent of Home Goods sales were accounted for by non-store retailers. These two categories account for almost half of all retail sales, which explains the overall low sales share of non-store retailers.

Figure C.2: Share of Local Concentration Term in National Concentration

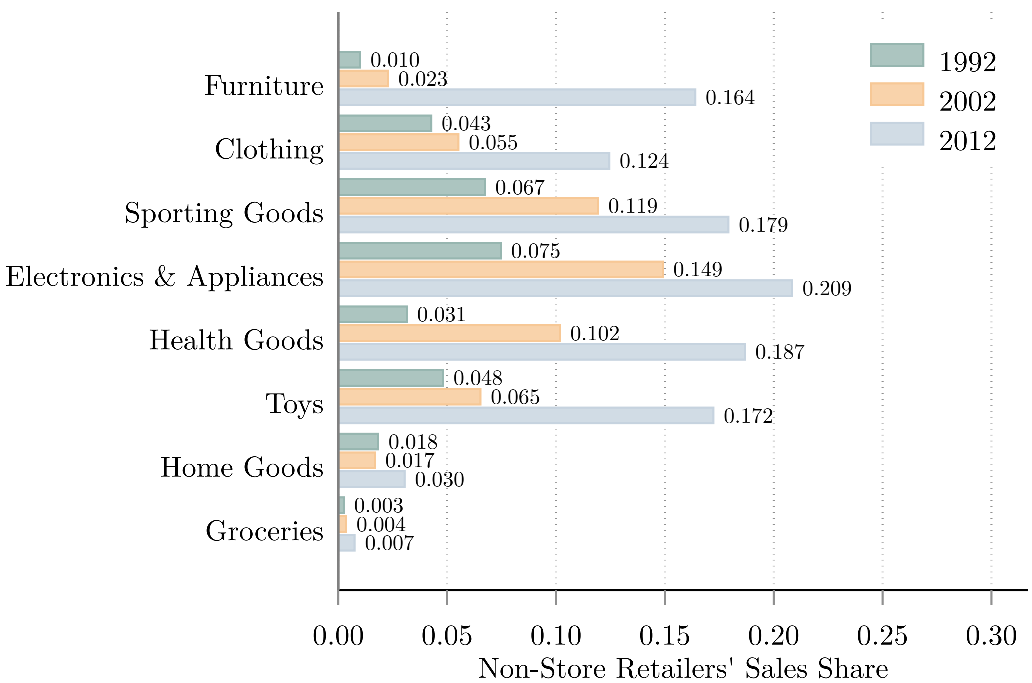


Notes: The numbers are from the Census of Retail Trade. The share of local concentration is measured as the ratio of the local concentration term in equation (3) to the national Herfindahl-Hirschman Index (HHI). We aggregate the local concentration terms across the product categories using their sales shares.

C.4 The Collocation Term and Local Concentration

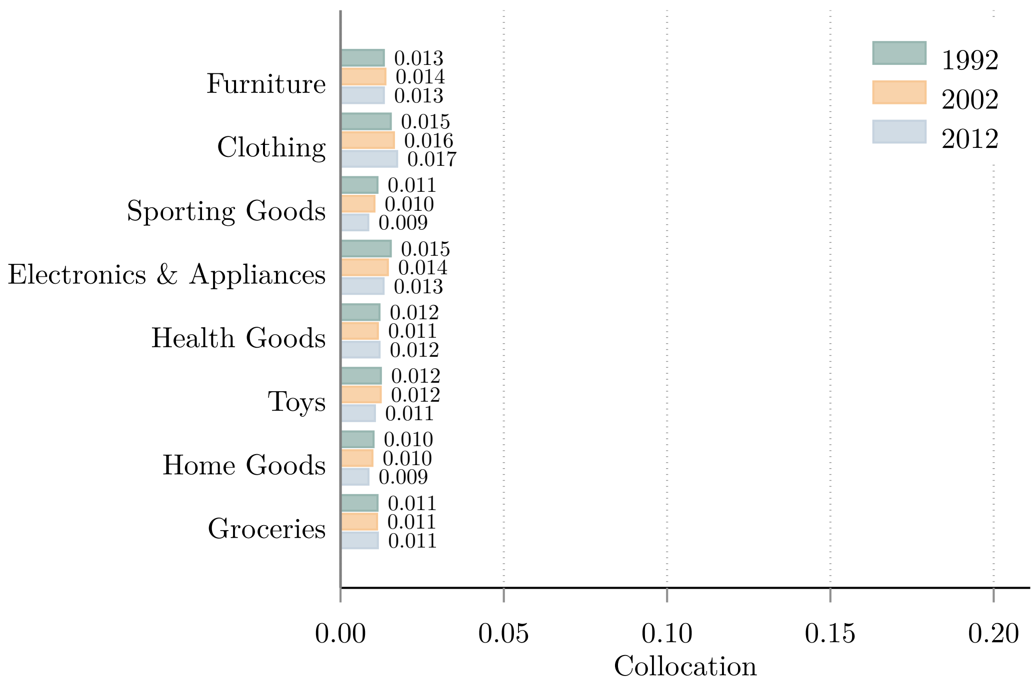
Figure C.4 shows the collocation term by product category. The numbers are the probability that two random dollars are spent in the same commuting zone for each year. These numbers are small, less than 2 percent, and stable over time. There is also little variation across product categories because spending on product categories is approximately proportional to each market's size. These numbers form the weights for the local HHI in the decomposition of national concentration. Their small magnitude explains the limited role of local concentration in explaining national changes. The contribution of local concentration varies across products but it is always low. By the early 1990s, only furniture and groceries have contributions of over 10 percent, with the local contribution in all other products being no higher than 5.5 percent, and as low as 2 percent.

Figure C.3: Non-Store Retailers Share across Product Categories



Notes: The numbers are the national sales shares of non-store retailers by product category from the Census of Retail Trade microdata.

Figure C.4: Collocation Term Product Categories



Notes: The data are from the CRT microdata. Numbers are the collocation term for commuting zones which forms the weight for the local HHI in the decomposition of national concentration.

D Industry-Based Results

A central contribution of this paper is the creation of store-level sales by product category for all U.S. retail stores. This allows us to define competition based on products rather than industry-based measures. Industries, either NAICS or standard industrial classification (SIC) codes, are regularly used to define markets. This approach is often necessitated by data availability and in many sectors is likely to be a good approximation (e.g. manufacturing).

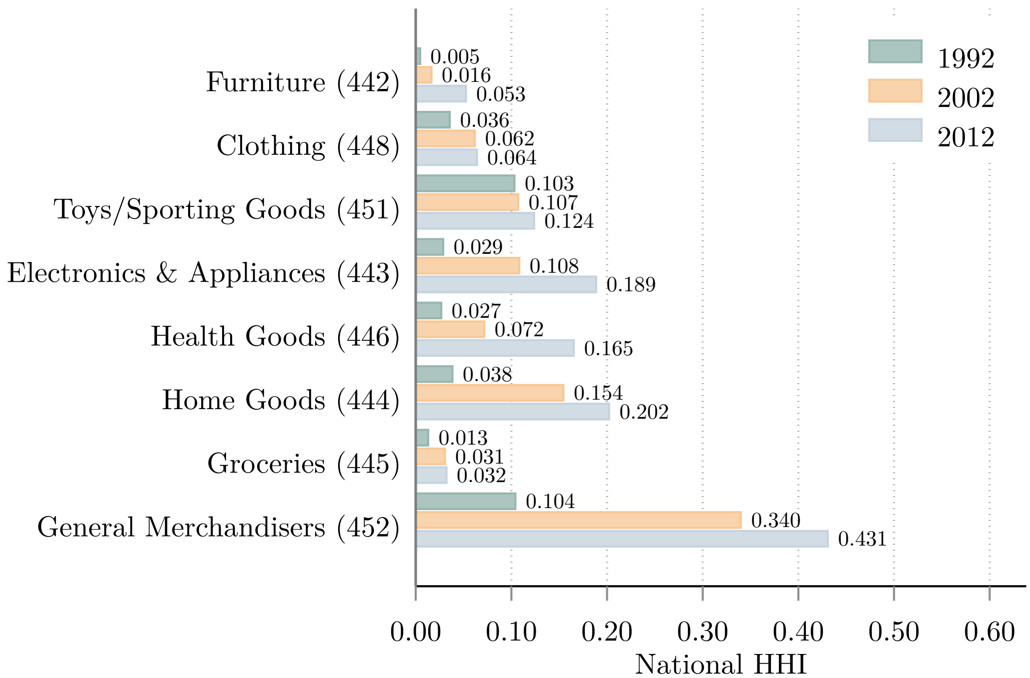
This is not the case in the retail sector. The retail sector has one set of industries, general merchandise stores (NAICS 452), that compete with stores in many industries. By construction these industries are composed by establishment that sell many types of products. Thus, industry-based measures ignore the competition faced by stores selling a given product, coming from general merchandise stores. The measures we developed in

Section 4 overcome this shortcoming.

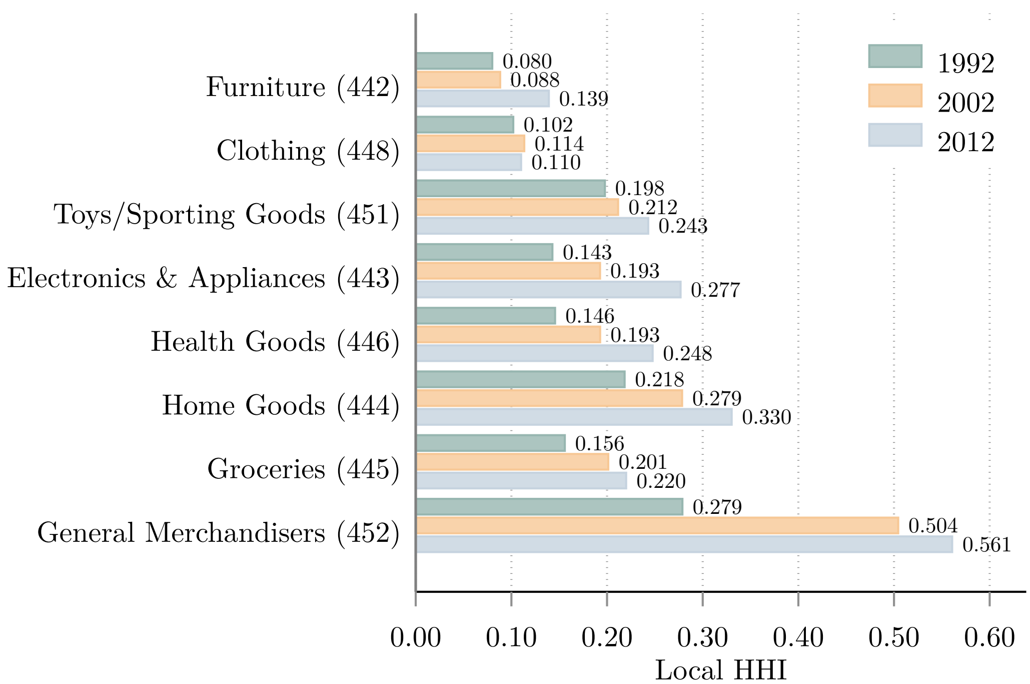
Figure D.1 shows national and local concentration for the main subsector (3-digit NAICS) of each of our product categories. Local concentration is defined at the commuting zone level. The increasing trends we documented for national concentration in the retail sector are present in all subsectors, but the increase is particularly strong for general merchandisers (NAICS 452) both at the national and at the local level. Industry based measures of concentration miss the impact of increasing concentration among general merchandisers across product markets. Similar patterns arise in local concentration. Figure D.1b shows local concentration for the major subsectors, calculated as a weighted average of the industries comprising each subsector. Local industry concentration also increases.

Figure D.1: National and Local Concentration Across Industries

(a) National Concentration

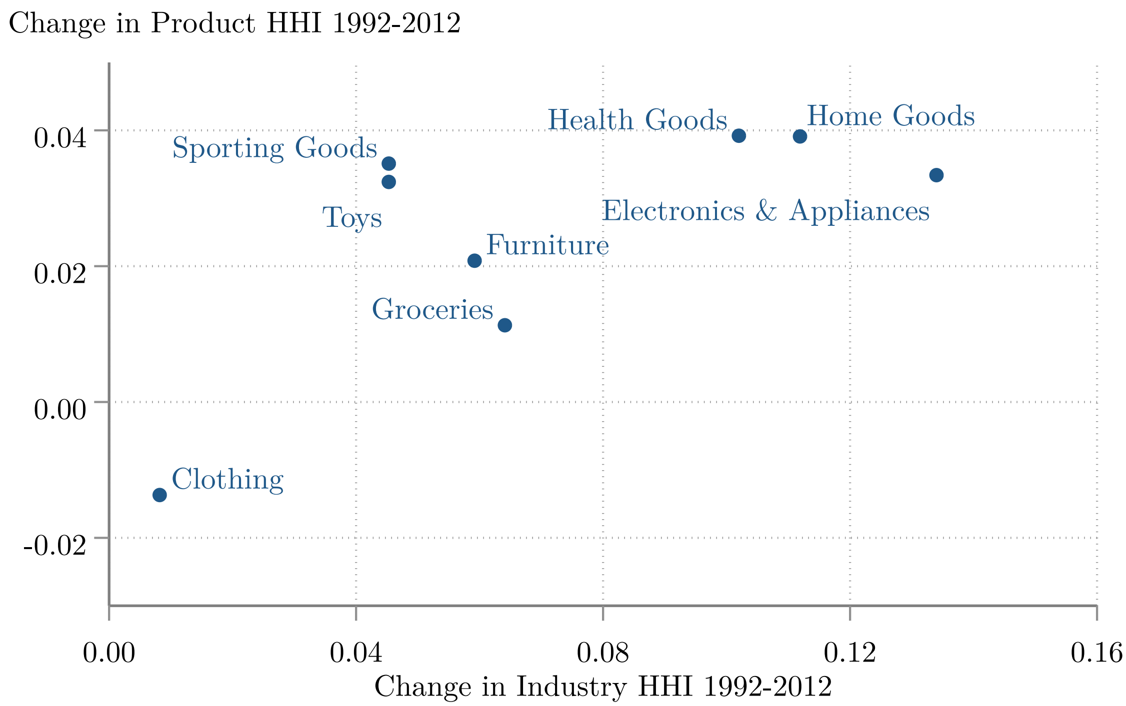


(b) Local Concentration



Notes: The data are from the Census of Retail Trade. Numbers are the national and Herfindahl-Hirschman Index (HHI) for various industries weighted by market size.

Figure D.2: National and Local Concentration



Notes: The data are from the Census of Retail Trade. Each point marks the change in local Herfindahl-Hirschman Index (HHI) of a product category and its main subsector between 1992 and 2012. Markets are defined at the commuting zone level.

Figure D.2 shows changes in product and industry based measures of local concentration. Despite changes in the level of the HHI both measures imply increasing concentration across most products/industries. Figure D.2 also makes clear that empirically there appears to be a correlation between product and industrial concentration, but this correlation is not perfect. Moreover, the significant sales share of non-store and general merchandise firms in most product categories means that this positive correlation could go away at any time.

E Comparison to Rossi-Hansberg, Sarte, and Trachter (2020)

This section compares our results to those in Rossi-Hansberg, Sarte, and Trachter (2020) (hereafter RST) for the retail sector and explains the factors contributing to the differences between our papers. Unlike us, they find a reduction in the local HHI for the retail sector between 1990 and 2014.[24]

There are three key differences between our paper and RST's that each partially explains the opposite results regarding local concentration. First, we use different data sources: while RST use the National Establishment Time Series (NETS), this paper uses confidential data from the CRT and the LBD. Second, we have different definitions of markets: this paper defines markets by product based on NAICS-6 classification of establishments, while RST define markets by industry based on SIC-8 or SIC-4 classification of establishments. Third, we differ in the methodology used to aggregate markets. This paper aggregates market-level concentration using contemporaneous weights, and we report the change in this (aggregate) index of local concentration. In contrast, RST aggregate the change in market-level concentration using end-of-period weights and report this (aggregate) change.

We argue that the CRT is likely to provide better data for the study of concentration in local markets, and we show that changing from NETS to CRT data alone explains a third of the discrepancy in the change of local concentration (while controlling for market definition and aggregation methodology). Another third of the difference in estimates is explained by the definition of product markets (by changing detailed SIC-8 industries to more aggregated SIC-4 industries). The proper definition of a product market (SIC-8, SIC-4, NAICS-6, product category) can depend on the question being asked. We argue in Section 3.3 that product categories are the proper way to study retail markets. The final third of the difference in estimates is explained by the aggregation methodology. We argue

[24]RST present results for many sectors of the economy. In what follows we discuss only their results in the retail sector. However, our discussion of aggregation methods is relevant for all sectors.

that the method used by RST is biased toward finding decreasing local concentration, and we show that their method could find evidence of decreasing concentration in a time series, even when concentration is not changing in the cross-section. This occurs when markets become less concentrated as they grow. Below we expand upon these differences and their implications for the measurement of local concentration.

Data sources The baseline results in RST are based on the NETS, a data product from Walls and Associates that contains information on industry, employment, and sales by establishments. These data have been shown to match county-level employment counts relatively closely (Barnatchez, Crane, and Decker, 2017), but the data do not match the dynamics of businesses Crane and Decker (2020). The results in this paper are based on the CRT, a data set assembled and maintained by the U.S. Census Bureau covering the universe of retail establishments.

Both the NETS data and the CRT use the establishment's reported industry and sales when available and both have some degree of imputation for establishments that do not report. However, the CRT can often impute using administrative records from the IRS.[25] Beyond this, the two data sets differ in other two relevant aspects. First, the CRT contains sales by product category for the majority of sales, while the NETS contains only industry, allowing us to define markets by product categories and account for cross-industry competition by general merchandisers (see Section 3.3). Second, the NETS includes non-employer establishments, while the CRT does not. According to official estimates, non-employer establishments account for about 2 percent of retail sales in 2012 (Economy-Wide Key Statistics: 2012 Economic Census of the United States).[26] On the whole, the CRT provides a more accurate picture of activity in the retail sector.

[25] Response to the CRT is required by law. Single-unit establishments are randomly sampled for sales in the CRT, while the non-sampled units have their sales imputed. See http://dominic-smith.com/data/CRT/crt_sample.html for more details.

[26] https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ECN_2012_US_00A1&prodType=table

Definition of product markets We adopt a different definition than RST for what constitutes a product market. Each definition of product market has its own pros and cons, and researchers may choose one over the other depending on the specific context. We define markets by a combination of a geographical location and a product category that we construct using the detailed data on sales provided by the CRT, along with the (NAICS-6) industry classification of establishments (see Section 3.3). As we mentioned above, doing this treats multi-product retailers as separate firms, ignoring economies of scope, in favor of putting all sales in a product category in the same market.

In contrast, RST define markets by the establishment's industry, using both SIC-8 and SIC-4 codes. Some examples of SIC-8 codes are department stores, discount (53119901); eggs and poultry (54999902); and Thai restaurants (58120115).[27] SIC-8 codes may be overly detailed for retail product markets, to the point that many retailers will sell multiple types of goods. For example, calculating concentration in eggs and poultry (54999902) would miss the fact that many eggs and poultry are sold by chain grocery stores (54119904) and discount department stores (53119901). This suggests that aggregating to less detailed codes may provide a better definition of product markets. To that end, RST present results for SIC-4 codes. When concentration is calculated using SIC-4 codes, the decrease in local concentration is much smaller, a 8 percentage point decrease instead of a 17 percentage point decrease.[28]

Incidentally, the SIC-4 codes are quite similar to the NAICS-6 codes available in the CRT, except restaurants are included in the SIC definition of retail but not in NAICS.[29] This makes the concentration measures based on each classification more closely comparable.

[27]NETS allows for 914 retail SIC-8 codes. A full list is available at https://www.dnb.com/content/dam/english/dnb-solutions/sales-and-marketing/sic_8_digit_codes.xls. RST indicate that many SIC-8 codes are rarely used (data appendix), but without access to the NETS data, we cannot assess the relative significance of each code for economic activity.

[28]The change from SIC-8 to SIC-4 has little effect on concentration outside of retail (RST Data Appendix). The numbers are read off graphs for the change in retail sector concentration for zip codes between 1990 and 2012.

[29]In the results in the main text, we exclude automotive dealers, gas stations, and non-store retailers because of concerns related to ownership data and defining which markets they serve (see Section 3 for further discussion). This has little impact on the estimates for local concentration.

Yet, even in this setting (NETS SIC-4 versus CRT NAICS-6) there are still significant differences between our studies. We will go back to this comparison when we discuss Figure E.2 and Table E.1 below.

Aggregation methodology The final difference comes from how we aggregate the market-level changes in concentration into an aggregate index of local concentration. We compute the local HHI index by first computing the HHI for each pair of product category (j) and location (ℓ). Then we aggregate across locations, weighting each market (location-product) HHI by the market's share of the product's national sales. Doing this provides a measure of the average local HHI for each product. Finally, we aggregate across products, weighting by the product's share of national retail sales, to obtain an average local HHI. We do this for each period (t) and report the time series for this index. The average local HHI is then given by

$$HHI_t = \underbrace{\sum_j s_j^t}_{Products} \underbrace{\sum_\ell s_\ell^{jt}}_{Locations} \cdot HHI_{j\ell t}, \quad \text{where } HHI_{j\ell t} = \sum_i \left(s_i^{j\ell t}\right)^2. \tag{E.1}$$

RST use a different methodology. Instead of computing concentration in the cross-section, they calculate the change in concentration between t and some initial period and then aggregate these changes weighting by the period t share of employment of each industry (j) in total retail employment. Their index for the change in concentration is given by[30]

$$\Delta HHI_t^{RST} = \sum_{j\ell} s_{j\ell}^t \Delta HHI_{j\ell t}, \tag{E.2}$$

where $s_{j\ell}^t$ is the sales share of industry j and location ℓ in the country at time t[31] and $\Delta HHI_{j\ell t}$ is the change in the revenue-based HHI in industry j and location ℓ between the base period and time t.

[30]Equation E.2 is taken from RST, with notation adjusted to match the notation in this paper.

[31]Note that RST weight markets by their employment share $\left(e_{j\ell}^t\right)$ instead of their sales share $\left(s_i^{j\ell t}\right)$. However, their data appendix shows this has no effect on the results.

The key difference between the methodologies is that RST do not account for the size of a market in the initial period. This is shown in equation E.3, which subtracts the two measures of concentration from each other. After canceling terms, the difference between the two measures is

$$\Delta HHI - \Delta HHI^{RST} = \sum_{j\ell} \underbrace{(s_{j\ell}^{t} - s_{j\ell}^{0})}_{\Delta s_{j\ell}^{t}} \cdot HHI_{m\ell 0}. \tag{E.3}$$

RST will weight markets that increase in size over time by more in the initial period, while those that decrease will be weighted less relative to our measure. As markets grow, they typically become less concentrated resulting in RST weighting markets with decreasing concentration more than markets with increasing concentration.[32]

Figure E.1 shows that this methodology can find decreasing concentration in a time series, even when concentration is not changing in the cross-section. Consider three firms (A, B, and C) that operate in two markets and have the same size. In the first period ($t-1$), firms A and B operate in market 1 and firm C operates in market 2. Consequently, the HHI is 0.5 and 1 for each market, respectively, and the aggregate (cross-sectional) HHI is $^2/_3$. In period t, market 1 shrinks and market 2 grows, with firm B changing markets. This change does not affect the cross-sectional distribution of local (market-specific) concentration, but it does imply an increase in concentration in market 1 and a decrease in market 2. Despite there being no changes in the cross-sectional HHI, RST's methodology would report a decrease in local concentration ($\Delta HHI = {}^{-1}/_{6}$), driven by the decrease in market 2's HHI (which happens to be the largest market in period t).

Quantifying differences Figure E.2 quantifies the role of each of the differences highlighted above for the change in local concentration between 1992 and 2012.[33] To

[32] A similar point is made in Appendix E of Ganapati (2020) using LBD data.

[33] RST use 1990 as the base year instead of 1992. This is unlikely to matter as RST find small changes in concentration between 1990 and 1992.

make the comparison clear, we define markets by industry throughout the exercise.[34] Overall, Figure E.2 shows that the difference in the estimated change of local HHI is explained in roughly equal parts by the three differences highlighted above: data source (CRT versus NETS), industry definition (NAICS-6 versus SIC-8), and aggregation methodology. We discuss each step in more detail below.

The lowest estimate for the change in local concentration (a decrease of 0.17 points in local HHI) corresponds to RST's baseline estimate using NETS data and SIC-8 for industry classification. Once industries are aggregated to the SIC-4 level (to improve comparability across establishments), the estimate increases by 9 percentage points, still implying a reduction of 8 percentage points in the local HHI. The next estimate reproduces RST's methodology using microdata from the CRT. Changing from NETS to CRT data implies a further increase in the estimate of 6.5 percentage points, with the overall change suggesting a minor decease of local HHI of 1.5 percentage points.[35] Next we change the weighting methodology to ours (as explained above). Doing so increases the estimated change of local concentration again (by 9.5 percentage points), implying an overall increase of local HHI of 8 percentage points.

Table E.1 provides a more detailed account of the estimates presented in Figure E.2 and also includes estimates of changes in local concentration for intermediate census years (1997, 2002, and 2007). In the first panel, national concentration, we compare the numbers in RST (Figure 1b) to numbers calculated for NAICS-based measures (including all 6-digit industries in NAICS) and product-based measures. In all three cases, national concentration is increasing significantly. Despite differences in the initial levels of concentration (column 1), the national HHI increases by two to three times.[36]

[34]To be precise, we define a market either by an SIC-8, an SIC-4, or a NAICS-6 industry in a given location. Our preferred definition of markets by product categories implies a change in the level of the HHI that makes the comparison with the results in RST less transparent.

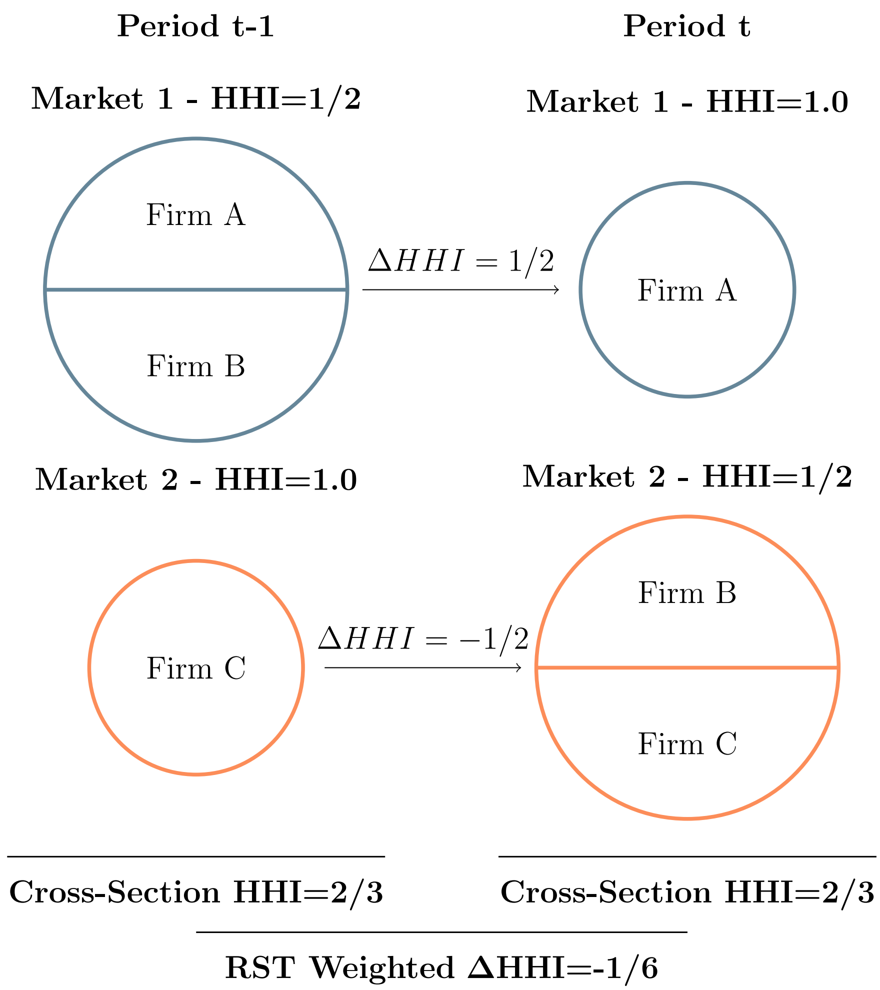
[35]Part of this difference could be explained in theory by the inclusion of restaurants in SIC-4; however, the industry by industry results in RST's Figure 7 suggest that this is not the case because they find diverging trends in most retail industries.

[36]The level of concentration is not provided in RST.

The second panel of Table E.1 compares concentration measured at the zip code level using RST's weighting methodology as described above. We also provide results for the set of establishments that are included in the product-based results in the paper. Using their methodology, we find evidence for slight decreases in local concentration of 1 to 2 percentage points whether markets are aggregated using sales or employment weights. These decreases are much less severe than the 17 percentage point decrease in RST.

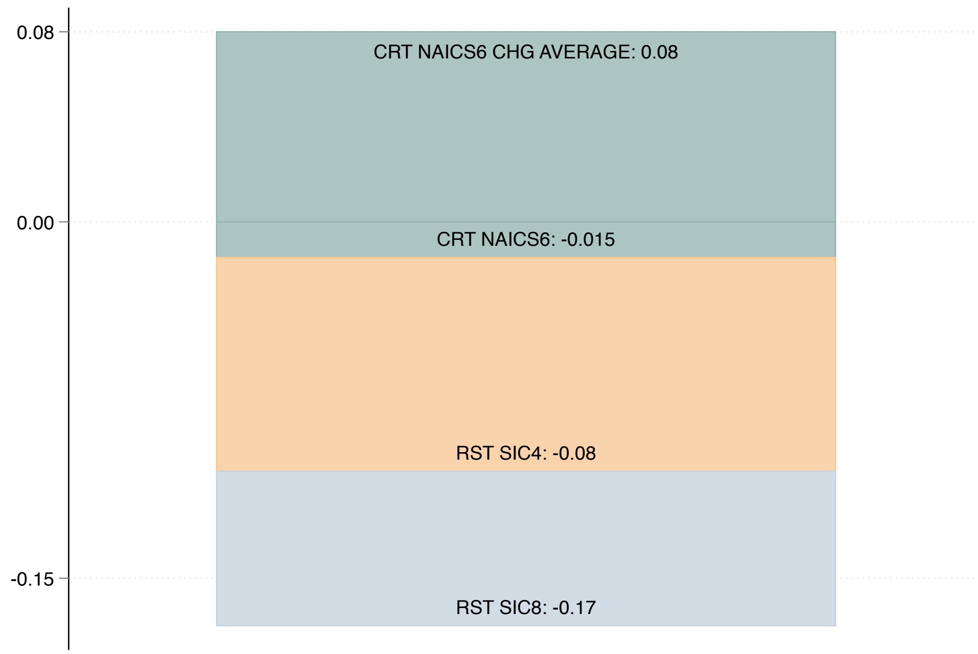
The final panel of Table E.1 compares concentration measured at the zip code level using our aggregation method. This method finds significant increases in local concentration across both NAICS samples. Local HHI increased between 7.1 and 8.5 percentage points; that is, the average dollar in 2012 is spent in a more concentrated market than the average dollar in 1992.

Figure E.1: Example of RST Methodology



Notes: The figure shows how market and cross-sectional concentration indices are computed under our methodology (difference in cross-section Herfindahl-Hirschman Index (HHI)) and that of Rossi-Hansberg et al. (2020). The economy has two markets and three firms. Firms are of the same size. Markets change size from period $t-1$ to period t, but the cross-sectional distribution of markets and concentration does not change. The weighting methodology used by Rossi-Hansberg et al. (2020) puts more weight on market 2, which increases size between $t-1$ and t and has a reduction in concentration. The result is a decrease in aggregate concentration when changes are measured according to this methodology, while cross-section HHI does not change.

Figure E.2: RST Comparison



Notes: The figure shows various estimates for the change in local HHI between 1992 and 2012. The estimates vary according to the data source, industry definition, and aggregation methodology. The lowest estimate corresponds to Rossi-Hansberg et al. (2020)'s estimate using SIC-8 industries, and the second lowest estimate corresponds to using SIC-4 industries. The second highest estimate corresponds to using Census of Retail Trade microdata and NAICS-6 industries (which are similar to SIC-4 industries), and the highest estimate computes indices under our aggregation methodology instead of that of Rossi-Hansberg et al. (2020).

Table E.1: Comparison of Concentration to RST

	Weight	Level 1992	1997	2002	2007	2012
National Concentration						
	Weight	Level	Change from 1992			
		1992	1997	2002	2007	2012
RST	Emp.	N/A	0.020	0.030	0.050	0.055
NAICS Based	Sales	0.029	0.017	0.056	0.076	0.087
Product Based	Sales	0.013	0.006	0.018	0.029	0.030
Zip Code Concentration: End-of-Period Weights						
	Weight	Level	Change from 1992			
RST	Emp.	N/A	-0.070	-0.100	-0.140	-0.170
NAICS Based	Emp.		-0.022	-0.016	-0.019	-0.015
	Sales		-0.023	-0.015	-0.017	-0.011
Paper Sample	Emp.		-0.002	-0.013	-0.018	-0.017
	Sales		-0.024	-0.009	-0.013	-0.011
Product based	N/A		N/A	N/A	N/A	N/A
Zip Code Concentration: Current Period Weights						
	Weight	Level	Change from 1992			
RST	N/A	N/A	N/A	N/A	N/A	N/A
NAICS Based	Emp.	0.507	0.025	0.060	0.068	0.080
	Sales	0.498	0.018	0.052	0.062	0.071
Paper Sample	Emp.	0.524	0.029	0.069	0.075	0.083
	Sales	0.530	0.022	0.073	0.081	0.085
Product Based	Sales	0.2637	0.013	0.024	0.022	0.013

Notes: The numbers come from the Census of Retail Trade and Rossi-Hansberg et al. (2020) (RST). Numbers from RST are taken from retail series in Figure 2. The level column contains the 1992 level of concentration. The formula for changes in concentration using end-of-period weights does not depend on the initial 1992 level as shown in RST, and consequently the level column does not apply to these calculations. NAICS-based measures concentration calculated including all NAICS industries. Paper sample uses only establishments included in the sample for the product-based results. Retail in RST is defined using SIC codes that include restaurants.

F Model of Firm's Markups

We now provide more detail on the model described in section 5. We follow Grassi (2017) who builds on Atkeson and Burstein (2008). The model's objective is to provide a link between local retail concentration and markups faced by consumers. We focus on how heterogeneous firms compete in an oligopolistic setup. Firms have market power in the local product markets in which they operate. To ensure tractability, we keep the modeling of demand as simple as possible. Demand for goods comes from a representative consumer, who supplies labor inelastically in each market and demands a national consumption good—a composite of all goods in the economy. The model closes with a perfectly competitive sector that aggregates individual goods from each market into the national consumption good.

F.1 The Model Economy

The model economy contains L locations, in each of them there are J products being transacted in local markets. Each location has N_ℓ retail firms that compete with one another in each good. Competition takes place at the location-product level. A perfectly competitive sector aggregates goods across firms for each product and location, aggregates products by location into location-specific retail goods, and aggregates each location's retail output into a final consumption good. A single representative consumer demands the final consumption good and supplies labor in each location.

F.1.1 Technology

A retailer i selling product j in location ℓ produces using only labor through a linear technology. $z_i^{j\ell}$ represents the productivity of the retailer:

$$y_i^{j\ell} = z_i^{j\ell} n_i^{j\ell}. \tag{F.1}$$

Labor is immobile across locations, but not products, so each location has a specific wage, w_ℓ. Firms maximize profits for each market they operate in:

$$\pi_i^{j\ell} = p_i^{j\ell} y_i^{j\ell} - \lambda_i^{j\ell} y_i^{j\ell}, \tag{F.2}$$

where $\lambda_i^{j\ell} = {}^{w_\ell}/_{z_i^{j\ell}}$ is the marginal cost of production.

The demand faced by the individual retailer comes from the aggregation sector that serves the consumer. Aggregation takes place in three levels. First, a local aggregator firm that combines the output of the N_ℓ retail firms selling product j in location ℓ. The firm operates competitively using the following technology:

$$y_j^\ell = \left(\sum_{i=1}^{N_\ell} \left(y_i^{j\ell} \right)^{\frac{\epsilon_j - 1}{\epsilon_j}} \right)^{\frac{\epsilon_j}{\epsilon_j - 1}} ; \qquad \epsilon_j > 1. \tag{F.3}$$

Then, the combined product bundles, y_j^ℓ, are themselves aggregated into local retail output, y_ℓ, through the following technology:

$$y_\ell = \prod_{j=1}^{J} \left(y_j^\ell \right)^{\gamma_j^\ell} ; \qquad \sum_{j=1}^{J} \gamma_j^\ell = 1, \tag{F.4}$$

where γ_j^ℓ is the share of product j in retail sales in location ℓ

Finally, the national retail output is created by combining local output, y_ℓ, from the L locations in the country:

$$y = \prod_{\ell=1}^{L} \left(y_\ell \right)^{\beta_\ell} ; \qquad \sum_{\ell=1}^{L} \beta_\ell = 1, \tag{F.5}$$

where β_ℓ corresponds to the share of location ℓ in national retail sales.

The aggregation process implies the following demand and prices:

$$y_\ell = \beta_\ell \frac{P}{p_\ell} \cdot y \qquad\qquad P = \prod_{\ell=1}^{L} \left(\frac{p_\ell}{\beta_\ell} \right)^{\beta_\ell} \tag{F.6}$$

$$y_j^\ell = \gamma_j^\ell \frac{p_\ell}{p_j^\ell} y_\ell \qquad\qquad p_\ell = \prod_{j=1}^{J} \left(\frac{p_j^\ell}{\gamma_j^\ell} \right)^{\gamma_j^\ell} \tag{F.7}$$

$$y_i^{j\ell} = \left(\frac{p_i^{j\ell}}{p_j^\ell} \right)^{-\epsilon_j} y_j^\ell \qquad\qquad p_j^\ell = \left(\sum_{i=1}^{N} \left(p_i^{j\ell} \right)^{1-\epsilon_j} \right)^{\frac{1}{1-\epsilon_j}} \tag{F.8}$$

F.1.2 Pricing to market

Firms compete directly in the sales of each product in a given location. We assume that firms are aware of the effect of their choices $\left(p_i^{j\ell}, y_i^{j\ell}\right)$ on the price and quantity of the product in the market they operate in $\left(p_j^{\ell}, y_j^{\ell}\right)$, but take as given the prices and quantities of other products in the same market, and of all products in other markets.

Firms choose either the price of their good $\left(p_i^{j\ell}\right)$ of the quantity $\left(y_i^{j\ell}\right)$ in a noncooperative fashion, taking as given the choices of other firms. We solve the pricing problem for Bertrand and Cournot competition (choosing prices or quantities respectively), as well as for the Dixit-Stiglitz monopolistic competition case, which serves as a useful framework.

The solution to the pricing problem is summarized in the following proposition taken from Grassi (2017):

Proposition 1. *The optimal price of a firm takes the form:* $p_i^{j\ell} = \mu_i^{j\ell}\lambda_i^{j\ell}$, *where* $\mu_i^{j\ell}$ *is a firm-product-market specific markup that depends on the form of competition:*

$$\mu_i^{j\ell} = \begin{cases} \frac{\epsilon_j}{\epsilon_j - 1} & \text{if Dixit-Stiglitz monopolistic competition} \\ \frac{\epsilon_j - (\epsilon_j - 1)s_i^{j\ell}}{\epsilon_j - 1 - (\epsilon_j - 1)s_i^{j\ell}} & \text{if Bertrand competition} \\ \frac{\epsilon_j}{\epsilon_j - 1 - (\epsilon_j - 1)s_i^{j\ell}} & \text{if Cournot competition} \end{cases} \tag{F.9}$$

and $s_i^{j\ell}$ *is the sales share of the firm in the given product market:*

$$s_i^{j\ell} = \frac{p_i^{j\ell} y_i^{j\ell}}{p_j^{\ell} y_j^{\ell}} = \left(\frac{p_i^{j\ell}}{p_j^{\ell}}\right)^{1-\epsilon_j} = \left(\frac{y_i^{j\ell}}{y_j^{\ell}}\right)^{\frac{\epsilon_j - 1}{\epsilon_j}} \tag{F.10}$$

We show details for the derivation in what follows.

Dixit-Stiglitz monopolistic competition The problem takes as given the product's price $\left(p_j^{\ell}\right)$ and aggregate demand $\left(y_j^{\ell}\right)$. The objective is to maximize profits by choosing the firm's price $\left(p_i^{j\ell}\right)$:

$$\max_{p_i^{j\ell}} p_i^{j\ell} y_i^{j\ell} - \lambda_i^{j\ell} y_i^{j\ell} \qquad \text{s.t. } y_i^{j\ell} = \left(\frac{p_i^{j\ell}}{p_j^{\ell}}\right)^{-\epsilon_j} y_j^{\ell}$$

Replacing the constraint:

$$\max_{p_i^{j\ell}} \left[\left(p_i^{j\ell}\right)^{1-\epsilon_j} - \lambda_i^{j\ell}\left(p_i^{j\ell}\right)^{-\epsilon_j}\right] \left(p_j^{\ell}\right)^{\epsilon_j} y_j^{\ell}$$

The first order condition is:

$$0 = (1-\epsilon_j)\left(p_i^{j\ell}\right)^{-\epsilon_j} + \epsilon_j \lambda_i^{j\ell}\left(p_i^{j\ell}\right)^{-\epsilon_j - 1}$$

$$0 = (1-\epsilon_j)\, p_i^{j\ell} + \epsilon_j \lambda_i^{j\ell}$$

Rearranging gives the result:

$$p_i^{j\ell} = \mu_i^{j\ell} \lambda_i^{j\ell} \qquad \mu_i^{j\ell} = \frac{\epsilon_j}{\epsilon_j - 1}$$

Bertrand competition The problem takes into account the effect of changes in the firm's own price on the product's price $\left(p_j^{\ell}\right)$ and aggregate demand $\left(y_j^{\ell}\right)$. The objective is to maximize profits by choosing the firm's price $\left(p_i^{j\ell}\right)$:

$$\max_{p_i^{j\ell}} p_i^{j\ell} y_i^{j\ell} - \lambda_i^{j\ell} y_i^{j\ell}$$

$$\text{s.t. } y_i^{j\ell} = \left(\frac{p_i^{j\ell}}{p_j^{\ell}}\right)^{-\epsilon_j} y_j^{\ell} \qquad y_j^{\ell} = \gamma_j^{\ell} \frac{p_m y_\ell}{p_j^{\ell}} \qquad p_j^{\ell} = \left(\sum_{i=1}^{N} \left(p_i^{j\ell}\right)^{1-\epsilon_j}\right)^{\frac{1}{1-\epsilon_j}}$$

Replacing the constraints:

$$\max_{p_i^{j\ell}} \left[\left(p_i^{j\ell}\right)^{1-\epsilon_j} - \lambda_i^{j\ell}\left(p_i^{j\ell}\right)^{-\epsilon_j}\right] \left(\sum_{i=1}^{N} \left(p_i^{j\ell}\right)^{1-\epsilon_j}\right)^{-1} \gamma_j^{\ell} p_\ell y_\ell$$

The first order condition is:

$$0=\left[\left(1-\epsilon_j\right)\left(p_i^{j\ell}\right)^{-\epsilon_j}+\epsilon_j\lambda_i^{j\ell}\left(p_i^{j\ell}\right)^{-\epsilon_j-1}\right]\left(\sum_{i=1}^N\left(p_i^{j\ell}\right)^{1-\epsilon_j}\right)^{-1}$$
$$-\left(1-\epsilon_j\right)\left[\left(p_i^{j\ell}\right)^{1-\epsilon_j}-\lambda_i^{j\ell}\left(p_i^{j\ell}\right)^{-\epsilon_j}\right]\left(\sum_{i=1}^N\left(p_i^{j\ell}\right)^{1-\epsilon_j}\right)^{-2}\left(p_i^{j\ell}\right)^{-\epsilon_j}$$
$$0=\left[\left(1-\epsilon_j\right)p_i^{j\ell}+\epsilon_j\lambda_i^{j\ell}\right]-\left(1-\epsilon_j\right)\left[p_i^{j\ell}-\lambda_i^{j\ell}\right]\left(p_i^{j\ell}\right)^{1-\epsilon_j}\left(p_j^{\ell}\right)^{\epsilon_j-1}$$
$$0=\left[\left(1-\epsilon_j\right)p_i^{j\ell}+\epsilon_j\lambda_i^{j\ell}\right]-\left(1-\epsilon_j\right)\left[p_i^{j\ell}-\lambda_i^{j\ell}\right]s_i^{j\ell}$$

Rearranging gives the result:

$$p_i^{j\ell}=\mu_i^{j\ell}\lambda_i^{j\ell}\qquad \mu_i^{j\ell}=\frac{\epsilon_j-\left(\epsilon_j-1\right)s_i^{j\ell}}{\epsilon_j-1-\left(\epsilon_j-1\right)s_i^{j\ell}}$$

Cournot competition The problem takes into account the effect of changes in the firm's own price on the product's price $\left(p_j^{\ell}\right)$ and aggregate demand $\left(y_j^{\ell}\right)$. The objective is to maximize profits by choosing the firm's quantity $\left(y_i^{j\ell}\right)$:

$$\max_{y_i^{j\ell}} p_i^{j\ell}y_i^{j\ell}-\lambda_i^{j\ell}y_i^{j\ell}$$

$$\text{s.t. } p_i^{j\ell}=\left(\frac{y_i^{j\ell}}{y_j^{\ell}}\right)^{\frac{-1}{\epsilon_j}}p_j^{\ell}\qquad p_j^{\ell}=\gamma_j^{\ell}\frac{p_\ell y_\ell}{y_j^{\ell}}\qquad y_j^{\ell}=\left(\sum_{i=1}^N\left(y_i^{j\ell}\right)^{\frac{\epsilon_j-1}{\epsilon_j}}\right)^{\frac{\epsilon_j}{\epsilon_j-1}}$$

Replacing the constraints:

$$\max_{y_i^{j\ell}}\left(y_i^{j\ell}\right)^{\frac{\epsilon_j-1}{\epsilon_j}}\left(\sum_{i=1}^N\left(y_i^{j\ell}\right)^{\frac{\epsilon_j-1}{\epsilon_j}}\right)^{-1}\gamma_j^{\ell}p_\ell y_\ell-\lambda_i^{j\ell}y_i^{j\ell}$$

The first order condition is:

$$0=\frac{\epsilon_j-1}{\epsilon_j}\left[\left(y_i^{j\ell}\right)^{\frac{-1}{\epsilon_j}}\left(y_j^{\ell}\right)^{\frac{1-\epsilon_j}{\epsilon_j}}-\left(y_i^{j\ell}\right)^{2\frac{\epsilon_j-1}{\epsilon_j}-1}\left(y_j^{\ell}\right)^{2\frac{1-\epsilon_j}{\epsilon_j}}\right]\gamma_j^{\ell}p_\ell y_\ell-\lambda_i^{j\ell}$$
$$0=\left(\epsilon_j-1\right)\left[1-\left(\frac{y_i^{j\ell}}{y_j^{\ell}}\right)^{\frac{\epsilon_j-1}{\epsilon_j}}\right]\left(\frac{y_i^{j\ell}}{y_j^{\ell}}\right)^{\frac{-1}{\epsilon_j}}\frac{\gamma_j^{\ell}p_\ell y_\ell}{y_j^{\ell}}-\epsilon_j\lambda_i^{j\ell}$$
$$0=\left(\epsilon_j-1\right)\left[1-s_i^{j\ell}\right]p_i^{j\ell}-\epsilon_j\lambda_i^{j\ell}$$

Rearranging gives the result:

$$p_i^{j\ell} = \mu_i^{j\ell} \lambda_i^{j\ell} \qquad \mu_i^{j\ell} = \frac{\epsilon_j}{\epsilon_j - 1 - (\epsilon_j - 1)\, s_i^{j\ell}}$$

F.1.3 Consumers

There is a representative consumer who has preferences over consumption of a national retail good, c. The consumer supplies labor inelastically in each location, with the local labor supply given by $\{n_\ell\}$.[37] The consumer receives income from profits and wages. The consumer's problem is:

$$\max_{\{c\}} u(c) = \frac{c^{1-\sigma}}{1-\sigma} \quad \text{ s.t. } p \cdot c \leq \sum_{\ell} n_\ell w_\ell + \Pi. \tag{F.11}$$

We normalize total labor supply, $n^S \equiv \sum_{\ell=1}^{L} n_\ell$, to one.

F.1.4 Equilibrium

The equilibrium of the model is standard and consists of a set of prices $\left\{P, \{p_\ell\}, \{p_j^\ell\}, \{p_i^{j\ell}\}\right\}$, wages $\{w_\ell\}$, outputs $\left\{y, \{y_\ell\}, \{y_j^\ell\}, \{y_i^{j\ell}\}\right\}$ and aggregate consumption demand, c, such that:

1. Aggregate prices and quantities satisfy F.6, F.7, and F.8.
2. Firm prices satisfy F.9, with the market share of each firm satisfying F.10.
3. Firm i's labor demand is given by $n_i^{j\ell} = y_i^{j\ell}/y_i^{j\ell}$.
4. Wages are such that local labor markets clear, that is, for each ℓ:

$$n_\ell = \sum_{j=1}^{J} \sum_{i=1}^{N_\ell} n_i^{j\ell},$$

where $n_i^{j\ell} = y_i^{j\ell}/y_i^{j\ell}$ corresponds to firm i's labor demand.

[37] In appendix F.4.2 we extend the model to include elastic labor supply.

F.2 Aggregating Markups

We now aggregate markups and productivity at the three levels of the economy (product-location, location, national).

F.2.1 Product-Location Level

The objective is to define an average markup for product j in location ℓ $\left(\mu_j^\ell\right)$, as well as the average productivity of firms producing product j in location ℓ $\left(z_j^\ell\right)$.

Average Markup The average markup is given by the ratio between the price p_j^ℓ and product-market marginal cost λ_j^ℓ. Because of constant returns to scale λ_j^ℓ is also the average cost:

$$\lambda_j^\ell = \frac{\sum_i \lambda_i^{j\ell} y_i^{j\ell}}{y_j^\ell} = \sum_{i=1}^{N} \lambda_i^{j\ell} \frac{y_i^{j\ell}}{y_j^\ell}$$

then the average markup is:

$$\mu_j^\ell = \frac{p_j^\ell}{\lambda_j^\ell} = \left[\sum_{i=1}^{N} \lambda_i^{j\ell} \frac{y_i^{j\ell}}{p_j^\ell y_j^\ell}\right]^{-1} = \left[\sum_{i=1}^{N} \left(\frac{\lambda_i^{j\ell}}{p_i^{j\ell}}\right)\left(\frac{p_i^{j\ell} y_i^{j\ell}}{p_j^\ell y_j^\ell}\right)\right]^{-1} = \left[\sum_{i=1}^{N} \left(\mu_i^{j\ell}\right)^{-1} s_i^{j\ell}\right]^{-1},$$

that is, a harmonic mean of individual markups, weighted by sales shares.

It is possible to further solve for the markup using the solution to the pricing problem above. The result is taken from Proposition 4 in Grassi (2017):

Proposition 2. *The average markup for product j in market m is:*

$$\mu_j^\ell = \begin{cases} \frac{\epsilon_j}{\epsilon_j - 1} & \textit{if Dixit-Stiglitz monopolistic competition} \\ \frac{\epsilon_j}{\epsilon_j - 1}\left[\frac{1}{\epsilon_j - 1}\sum_{k=2}^{\infty}\left(\frac{\epsilon_j - 1}{\epsilon_j}\right)^{k-1}\left(HK_j^\ell(k)\right)^k\right]^{-1} & \textit{if Bertrand competition} \\ \frac{\epsilon_j}{\epsilon_j - 1}\left[1 - HHI_j^\ell\right]^{-1} & \textit{if Cournot competition} \end{cases}$$

where $HK_j^\ell(k)$ is the Hanna & Kay (1977) concentration index of order k:

$$HK_j^\ell(k) = \left[\sum_{i=1}^{N}\left(s_i^{j\ell}\right)^k\right]^{\frac{1}{k}},$$

and $HHI_j^\ell = HK_j^\ell(2)^2 = \sum_i \left(s_i^{j\ell}\right)^2$ *is the Herfindahl-Hirschman Index.*

Average Productivity The average product is also obtained from the marginal (average) cost:

$$\lambda_j^\ell = \sum_{i=1}^{N} \lambda_i^{j\ell} \frac{y_i^{j\ell}}{y_j^\ell} = \left[\sum_{i=1}^{N} \left(z_i^{j\ell}\right)^{-1} \frac{y_i^{j\ell}}{y_j^\ell}\right] w_\ell$$

which implies:

$$z_j^\ell = \left[\sum_{i=1}^{N} \left(z_i^{j\ell}\right)^{-1} \frac{y_i^{j\ell}}{y_j^\ell}\right]^{-1},$$

an output-weighted harmonic mean of productivities.

F.2.2 Local market and national level

Markups and productivities can be aggregated again at the market level (aggregating across products) by defining first the market's marginal (average) cost:

$$\lambda_\ell = \frac{\sum \lambda_j^\ell y_j^\ell}{y_\ell}$$

For markups this implies:

$$\mu_\ell = \frac{p_\ell}{\lambda_\ell} = \left[\sum_{j=1}^{J} \left(\mu_j^\ell\right)^{-1} s_j^\ell\right]^{-1} = \left[\sum_{j=1}^{J} \left(\mu_j^\ell\right)^{-1} \gamma_j^\ell\right]^{-1}$$

For productivity:

$$z_\ell = \frac{w_\ell}{\lambda_\ell} = \left[\sum_{j=1}^{J} \left(z_j^\ell\right)^{-1} \frac{y_j^\ell}{y_\ell}\right]^{-1}$$

The same procedure gives the markup for the national level:

$$\mu = \left[\sum_{\ell=1}^{L} \left(\mu_\ell\right)^{-1} \beta_\ell\right]^{-1}$$

We define the productivity at the national level as the harmonic mean of local productivities weighted by output shares:

$$z \equiv \left[\sum_{\ell=1}^{L} \left(z_\ell\right)^{-1} \frac{y_\ell}{y}\right]^{-1}$$

This expression does not follow as the others because the cost of production (w_ℓ) differs across markets.

Multi-product/Multi-market firm Note that the equations above also apply to firms that sell various products and operate in various markets, modifying the sums to account only for the firm's products and markets.

F.2.3 Product aggregation

We also compute the average markup of a product across markets. This measure is relevant because it can be obtained directly from the data. We define the average markup

$$\mu_j \equiv \frac{\sum_{\ell=1}^{L} p_j^\ell y_j^\ell}{\sum_{\ell=1}^{L} w_\ell l_j^\ell}$$

as the ratio between product $j's$ total sales and total labor costs of the product across markets ($\ell = 1, \ldots, L$). The average markup is given in the model by:

$$\mu_j \equiv \frac{\sum_{\ell=1}^{L} p_j^\ell y_j^\ell}{\sum_{\ell=1}^{L} w_\ell l_j^\ell} = \frac{\sum_{\ell=1}^{L} p_j^\ell y_j^\ell}{\sum_{\ell=1}^{L} \frac{\lambda_j^\ell}{p_j^\ell} p_j^\ell y_j^\ell} = \left[\sum_{\ell=1}^{L} \left(\mu_j^\ell\right)^{-1} \theta_j^\ell\right]^{-1},$$

a harmonic mean of market level markups for product j, weighted by the share of product j sales in market ℓ captured by $\theta_j^\ell \equiv \frac{p_j^\ell y_j^\ell}{\sum_{\ell=1}^{L} p_j^\ell y_j^\ell} = \frac{\gamma_j^\ell \beta_\ell}{\sum_{\ell=1}^{L} \gamma_j^\ell \beta_\ell}$.

Using the result in Proposition 2 it is possible to express the product markup in terms of market concentration. For the case of Cournot competition it gives:

$$\mu_j = \left[\sum_{\ell=1}^{L} \left(\frac{\epsilon_j^\ell}{\epsilon_j^\ell - 1}\right)^{-1} \left[1 - \mathrm{HHI}_j^\ell\right] \theta_j^\ell\right]^{-1}$$

If the elasticity of substitution across varieties of good j is common across markets the expression simplifies to:

$$\mu_j = \frac{\epsilon_j}{\epsilon_j - 1} \left[1 - \mathrm{HHI}_j\right]^{-1},$$

where $\mathrm{HHI}_j \equiv \sum_{\ell=1}^{L} \mathrm{HHI}_j^\ell \theta_j^\ell$ is the sales weighted Herfindahl-Hirschman Index of product j across market.

F.3 Estimation Steps

We estimate the model using product level data from the Census of Retail Trade and the Annual Retail Trade Survey. This allows us to discuss how conditions in the average U.S. market has changed. To accomplish this we use the estimates of local concentration from section 4.2 and data on markups, prices, output, and labor supply. As in the empirical analysis of sections 3 and 4, we define markets in the model as pairs of a commuting zone and one of the product categories described in Table B.1.

The Cobb-Douglas parameters, β_ℓ and γ_j^ℓ, are obtained from the Census of Retail Trade as the share of spending on each product in a commuting zone. The estimation of the elasticity of substitution parameters consists on matching the product level markup from the ARTS given the product's average local concentration. From equation (9) we get:

$$\hat{\epsilon}_j = \frac{\hat{\mu}_j \left[1 - \sum_\ell s_\ell^j HHI_j^\ell\right]}{\hat{\mu}_j \left[1 - \sum_\ell s_\ell^j HHI_j^\ell\right] - 1} \tag{F.12}$$

where $\hat{\mu}_j = {}^{\text{Sales}_j}/_{\text{Cost of Goods Sold}_j}$ is the gross markup for product j. We use 2007 ARTS data for the estimation of the elasticity of substitution, matching all products' markups in that year by construction. Using our estimate of the elasticity of substitution parameters and the measured series for the product-level HHI we construct the series of markups implied by the model through equation (9).

We also define implicit price and quantity indexes for each product such that they are consistent with total sales of the product across markets:

$$P_j Y_j = \sum_\ell p_j^\ell y_j^\ell \tag{F.13}$$

Given the quantity index we define the average (marginal) cost of goods for a product, λ_j, as the output-weighted average of the individual market costs:

$$\lambda_j \equiv \sum_\ell \lambda_j^\ell \frac{y_j^\ell}{Y_j}. \tag{F.14}$$

Note that the average cost satisfies the following pricing equation at the product level:

$$P_j = \mu_j \lambda_j. \tag{F.15}$$

Finally, we can aggregate our product-level results to obtain a measure of the average retail cost and markup. The average cost is defined, as before, as the output-weighted average of the individual product costs:

$$\lambda \equiv \sum_j \lambda_j \frac{y_j}{Y}, \tag{F.16}$$

where Y is a quantity index for the retail sector. The average markup is defined as the ratio of total sales to cost:

$$\mu \equiv \frac{\sum_j P_j Y_j}{\sum_j \lambda_j Y_j} = \frac{\sum_j P_j Y_j}{\sum_j \frac{\lambda_j}{P_j} P_j Y_j} = \left[\sum_j (\mu_j)^{-1} s_j \right]^{-1}, \tag{F.17}$$

where s_j is the expenditure share of product j. As before this measure of markup satisfies the pricing equation at the national level:

$$P = \mu \lambda, \tag{F.18}$$

where P is a retail price index satisfying:

$$PY = \sum_j P_j Y_j. \tag{F.19}$$

F.3.1 Comparing Results Across Time

To compare our model's cross-sectional results across time we choose normalizations for prices that make aggregate numbers consistent with published statistics.[38] We use data on the change of retail good prices from the Price Indexes for Personal Consumption, from the U.S. Bureau of Economic Analysis (2020). These data provides us with series for the price index of each good category.[39] Each price index defines the inflation of prices in its

[38]The level of the aggregate price does not affect relative prices, output, or markups in the model.

[39]The price index for some product categories is not directly provided by the BEA data. In these cases we construct the category's index from individual product's series in the same way as we construct the aggregate retail index from the product category indexes.

respective category. We normalize the index so that $P_j^{1987} = 1$ for all product categories $j = 1, \ldots, J$. The level of the price index in year t reflects the cumulative (gross) inflation of prices in the product category.

We aggregate the individual category price indexes following the same procedure as the BEA. This procedure defines the aggregate index as an expenditure share weighted geometric average of the categories' indexes, the same definition as in our model (see equation F.6). Since the level of the individual indexes is arbitrary and only allows for direct comparisons across time and not products, we construct the aggregate index indirectly by computing its change over time:

$$\frac{P_t}{P_{t-1}} = \prod_{j=1}^{J} \left(\frac{P_j^t}{P_j^{t-1}} \right)^{s_j^t}. \tag{F.20}$$

We normalize the aggregate index so that $P_{1992} = 1$, and obtain the level in subsequent periods by concatenating the changes obtained in equation (F.20). As before the index provides the cumulative (gross) inflation in retail prices since 1992.

Finally, we deflate our retail price index by overall inflation. Without this adjustment the index reflects not only changes in retail prices, but also trends in overall inflation due to monetary or technological phenomena that are outside of the scope of the model. From these data we find retail prices decreased 35 percent relative to overall inflation. We use aggregate price index we obtain and the average retail markup (equation F.17) to compute the value of the average marginal cost λ, implied by equation (F.18).

F.4 Extensions

F.4.1 Marginal Costs Functional Form

In the baseline model production at a retail uses only labor as an input. In this section we evaluate how our setup maps to the case where the firms uses an arbitrary constant-returns-to-scale technology that uses labor and other materials.

Consider the problem of retail firms that use multiple inputs $\{x_k\}_{k=1}^{K}$ in addition to labor to produce:

$$y_i^{j\ell} = \tilde{z}_i^{j\ell} F\left(x_1, \ldots, x_K, n_i^{j\ell}\right), \tag{F.21}$$

where the function F is strictly concave, twice continuously differentiable, and has constant returns to scale. Letting the prices of inputs be $\{\tilde{p}_k\}_{k=1}^K$ and $\tilde{w}_\ell$ respectively, we know from the firm's optimality condition that:

$$\tilde{z}_i^{j\ell} F_k\left(\frac{x_1}{n_i^{j\ell}}, \ldots, \frac{x_K}{n_i^{j\ell}}, 1\right) = \tilde{p}_k \tag{F.22}$$

recalling that, because of Euler's theorem, F_k is homogeneous of degree zero for every k.

The equations defined by (F.22) define a square system in the ratio ratio of each input x_k to labor. The system has a solution that gives the ratios in terms of parameters:

$$\frac{x_k}{n_i^{j\ell}} = g_k\left(\tilde{z}_i^{j\ell}, \tilde{p}_1, \ldots, \tilde{p}_K\right) \tag{F.23}$$

The existence of a solution follows from the inverse function theorem applied to the function $\nabla_x F : \mathbb{R}_{++}^K \to \mathbb{R}_{++}^K$, where the operator ∇_x gives the first derivatives of F with respect to the variables $\{x_k\}_{k=1}^K$. Note that the Jacobian of ∇F is given by the first K rows and columns of the Hessian of F, which is negative definite for all interior points by the strict concavity of F. The negative definiteness of the Jacobian ensures the invertibility of $\nabla_x F$.

Given the system's solution we express the production function in terms of labor alone:

$$y_i^{j\ell} = \tilde{z}_i^{j\ell} F\left(\frac{x_1}{n_i^{j\ell}}, \ldots, \frac{x_K}{n_i^{j\ell}}, 1\right) n_i^{j\ell} = z_i^{j\ell} n_i^{j\ell}$$

where we define the effective productivity of labor as $z_i^{j\ell} \equiv \tilde{z}_i^{j\ell} F\left(x_1/n_i^{j\ell}, \ldots, x_K/n_i^{j\ell}, 1\right)$ with the ratios $x_k/n_i^{j\ell}$ given as in (F.23). Thus, z_i^j is a function of productivity $\tilde{z}_i^{j\ell}$ and the price of the other inputs. This is the production function we use in the main model.

Finally, the cost of labor must take into account that other inputs react to changes in labor according to (F.23). Then, the cost of the firm is given by:

$$\sum_{k=1}^K \tilde{p}_k x_k + \tilde{w}_\ell n_i^{j\ell} = \left(\sum_{k=1}^K \tilde{p}_k \frac{x_k}{n_i^{j\ell}} + \tilde{w}_\ell\right) n_i^{j\ell} = w_\ell n_i^{j\ell}$$

where $w_\ell n_i^{j\ell}$ represents the cost of goods sold, and w_ℓ is not directly the wage, but a measure of costs that takes into account the price of other inputs and the change in their demand in response to changes in the firm's labor demand.

F.4.2 Elastic Labor Supply

In this section we outline a version of the model where consumers have preferences over national consumption (c) and leisure/labor in each location: $u(c, n_1, \ldots, n_L)$. This setup does not affect any of the results in the paper as all results using the markup equation go through unchanged.

We consider a utility function that is separable in consumption and labor:

$$u(c, \{n_\ell\}) = \frac{c^{1-\sigma}}{1-\sigma} - \chi \sum_{\ell=1}^{L} \frac{(n_\ell)^{1+\frac{1}{\phi}}}{1+\frac{1}{\phi}} \qquad \text{and} \qquad \chi c^{\sigma} (n_\ell)^{\frac{1}{\phi}} = \frac{w_\ell}{P}.$$

ϕ corresponds to the Frisch elasticity of labor supply. σ is the curvature of utility in consumption.

The first order conditions of the consumer imply:

$$\frac{u_{n_\ell}(c, \{n_\ell\})}{u_c(c, \{n_\ell\})} = \frac{w_\ell}{P}.$$

This governs how total labor supply reacts to changes in prices and changes in markups. These results affect how productivity and output respond to changes in prices.

F.4.3 Uniform prices across locations

Consider now the problem of firm i that sales product j across various markets $\ell \in \mathcal{L}_i$. There are three options for pricing: pricing to market, ignoring linkages of demand across markets, pricing to market incorporating linkages of demand, uniform pricing. We deal with them in turn.

The first price option (pricing to market, ignoring effects on demand across markets) gives the same solution as above, and the aggregation is also the same. The second option would require the firm to take into account the effect on the demand for groceries in New York of a price change in groceries in Minneapolis. We consider this to be implausible, and the effect to be likely very small (even if firms are taking into account). Thus we think this case is well approximated by our baseline case above. The final option is uniform pricing, which we solve for below.

The problem of the firm is:

$$\max_{p_i^j} \sum_{\ell \in \mathcal{L}_i} \left[p_i^j y_i^{j\ell} - \lambda_i^{j\ell} y_i^{j\ell} \right]$$

$$\text{s.t. } y_i^{j\ell} = \left(\frac{p_i^j}{p_j^\ell} \right)^{-\epsilon_j} y_j^\ell \qquad y_j^\ell = \gamma_j^\ell \frac{p_\ell y_\ell}{p_j^\ell} \qquad p_j^\ell = \left(\sum_{i=1}^N \left(p_i^{j\ell} \right)^{1-\epsilon_j} \right)^{\frac{1}{1-\epsilon_j}}$$

Replacing the constraints:

$$\max_{p_i^j} \sum_{\ell \in \mathcal{L}_i} \left[\left(p_i^j \right)^{1-\epsilon_j} - \lambda_i^{j\ell} \left(p_i^j \right)^{-\epsilon_j} \right] \left(\sum_{i=1}^N \left(p_i^{j\ell} \right)^{1-\epsilon_j} \right)^{-1} \gamma_j^\ell p_\ell y_\ell$$

The first order condition is:

$$0 = \sum_{\ell \in \mathcal{L}_i} \left[\left[(1-\epsilon_j) p_i^j + \epsilon_j \lambda_i^{j\ell} \right] \frac{y_i^{j\ell}}{p_i^j} - (1-\epsilon_j) \left[p_i^j - \lambda_i^{j\ell} \right] s_i^{j\ell} \frac{y_i^{j\ell}}{p_i^j} \right]$$

$$0 = \sum_{\ell \in \mathcal{L}_i} \left[-(\epsilon_j - 1) \left(1 - s_i^{j\ell} \right) y_i^{j\ell} p_i^j + \left(\epsilon_j - (\epsilon_j - 1) s_i^{j\ell} \right) \lambda_i^{j\ell} y_i^{j\ell} \right]$$

Rearranging:

$$p_i^j = \frac{\sum_\ell \left(\epsilon_j - (\epsilon_j - 1) s_i^{j\ell} \right) \lambda_i^{j\ell} y_i^{j\ell}}{\sum_\ell (\epsilon_j - 1) \left(1 - s_i^{j\ell} \right) y_i^{j\ell}}$$

If marginal cost is constant across markets then we define the markup :

$$p_i^j = \mu_i^j \lambda_i^j \qquad \mu_i^j = \frac{\sum_\ell \left(\epsilon_j - (\epsilon_j - 1) s_i^{j\ell} \right) y_i^{j\ell}}{\sum_\ell (\epsilon_j - 1) \left(1 - s_i^{j\ell} \right) y_i^{j\ell}}$$

The firm's markup reflects its market power across different markets, captured by the firm's output-weighted average share, $\hat{s}_i^j$. Define $\hat{y}_i^{j\ell} = y_i^{j\ell} / \sum_\ell y_i^{j\ell}$, then:

$$\mu_i^j = \frac{\sum_\ell \left(\epsilon_j - (\epsilon_j - 1) s_i^{j\ell} \right) \hat{y}_i^{j\ell}}{\sum_\ell (\epsilon_j - 1) \left(1 - s_i^{j\ell} \right) \hat{y}_i^{j\ell}} = \frac{\epsilon_j - (\epsilon_j - 1) \sum_\ell s_i^{j\ell} \hat{y}_i^{j\ell}}{(\epsilon_j - 1) \left(1 - \sum_\ell s_i^{j\ell} \hat{y}_i^{j\ell} \right)} = \frac{\epsilon_j - (\epsilon_j - 1) \hat{s}_i^j}{(\epsilon_j - 1) \left(1 - \hat{s}_i^j \right)}$$

The firm's uniform markup is lower than the average markup if the firm chooses prices in each market separately. To see this, define the firm's average price in product j such that:

$$p_i^j y_i^j = \sum_{\ell} p_i^{j\ell} y_i^{j\ell},$$

where $y_i^j \equiv \sum_{\ell} y_i^{j\ell}$. It follows that $p_i^j = \sum_{\ell} p_i^{j\ell} \hat{y}_i^{j\ell}$. The average markup would then be:

$$\mu_i \equiv \frac{p_i^j}{\lambda_i^j} = \sum_{\ell} \frac{p_i^{j\ell}}{\lambda_i^j} \hat{y}_i^{j\ell} = \sum_{\ell} \mu_i^{j\ell} \hat{y}_i^{j\ell},$$

which is the output-weighted average of the individual market markups. This average is higher than the uniform markup. The result follows from Jensen's inequality as the Bertrand markup is convex in the firm's sales share.