

**WORKING PAPER** · NO. 2019-102

# The Rise of Niche Consumption

*Brent Neiman and Joseph S. Vavra*

AUGUST 2019

## THE RISE OF NICHE CONSUMPTION

Brent Neiman  
Joseph S. Vavra

We thank David Argente, Anhua Chen, Levi Crews, and Agustin Gutierrez for providing exceptional research assistance and Rodrigo Adao, Jonathan Dingel, J.P. Dubé, Austan Goolsbee, Pete Klenow, Thomas Mertens, Esteban Rossi-Hansberg, and Tom Wollmann for helpful comments and suggestions. Our analyses are calculated or derived based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from these Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. This paper supplants a previous draft circulated as “The Rise of Household Spending Concentration.”

© 2019 by Brent Neiman and Joseph S. Vavra. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Rise of Niche Consumption  
Brent Neiman and Joseph S. Vavra  
August 2019  
JEL No. D12,D4,E21,E31

### **ABSTRACT**

We show that over the last 15 years, the typical household has increasingly concentrated its spending on a few preferred products. However, this is not driven by “superstar” products capturing larger market shares. Instead, households increasingly focus spending on different products from each other. As a result, aggregate spending concentration has in fact decreased over this same period. We use a novel heterogeneous agent model to conclude that increasing product variety is a key driver of these divergent trends. When more products are available, households can select a subset better matched to their particular tastes, and this generates welfare gains not reflected in government statistics. Our model features heterogeneous markups because producers of popular products care more about maximizing profits from existing customers, while producers of less popular niche products care more about expanding their customer base. Surprisingly, however, our model can match the observed trends in household and aggregate concentration without any resulting change in aggregate market power.

Brent Neiman  
University of Chicago  
Booth School of Business  
5807 South Woodlawn Avenue  
Chicago, IL 60637  
and NBER  
[brent.neiman@chicagobooth.edu](mailto:brent.neiman@chicagobooth.edu)

Joseph S. Vavra  
Booth School of Business  
University of Chicago  
5807 South Woodlawn Avenue  
Chicago, IL 60637  
and NBER  
[joseph.vavra@chicagobooth.edu](mailto:joseph.vavra@chicagobooth.edu)

# 1 Introduction

Recent decades have played host to important changes in how goods are designed, produced, marketed, and delivered. In this paper, we explore the ramifications of the rapidly changing retail environment using household-level data on product spending. We first document striking trends in the concentration of household spending across products. We then develop a novel model that matches these trends and delivers important implications for welfare, market power and other aggregates that could not be gleaned using producer-, retailer-, or industry-level data alone.

We start by documenting that from 2004-2016, the typical household's spending has become significantly more concentrated on their set of preferred products. Further, the number of products consumed by the typical household within a given category has declined. These facts might on their own point toward an increasing importance of "superstar" products in retail bundles, but a similar analysis of aggregate spending paints quite a different picture. Pooling households together, we see that total spending has become more evenly distributed across the product space over the same period. How can this simultaneous increase in household concentration and decline in aggregate concentration be reconciled? Each household increasingly concentrates spending on its own preferred products, but households also increasingly differ on which products they consume. We refer to this greater fragmentation of the product space as a rise in "niche" consumption.

For example, the typical household in 2004 purchased an average of 16.6 UPCs in detailed product categories such as "Carbonated Beverages", "Laundry Supplies", and "Butter and Margarine", but by 2016 this had declined to 15.6 UPCs. Various measures of household spending concentration increased over this same time period. For example, the average household-category Herfindahl increased by roughly 8 percent (2.2 percentage points) from 2004-2016. However, we next add up product spending across all households in a category to compute aggregate market shares and concentration. We find that the aggregate Herfindahl in the typical category declined by more than 20 percent from 2004-2016. These divergent trends occur fairly steadily over time, are highly statistically significant, and are robust to a variety of specification and measurement choices.

The rise in niche consumption is also highly robust to the inclusion of a variety of observable controls. The divergence between household and aggregate concentration is not driven by a widening gap between the goods purchased by rich and poor households, between consumers in one region and another, or by other obvious demographic characteristics. Instead, we find that household consumption bundles are becoming more differentiated even when measuring within geographies, within store chains, and within demographic groups defined by income, race, education, and age. While there is variation in the magnitude, niche consumption grows in almost all product categories.

We next show that the rise of niche consumption in the data is strongly associated with product entry. The categories and the retail chains that have the largest aggregate variety growth also see

the largest increases in niche consumption. In addition, both household and aggregate concentration trends significantly attenuate when we eliminate product churn and apply our measure to a constant panel of goods. That is, the products accounting for rising spending shares for each household within narrow categories are predominantly products which are new to that household.

Exploring the dynamics of product market shares within particular categories further reinforces the role of product entry and differentiation for these trends. For example, Yoplait and Dannon lost substantial market share within the yogurt category over our sample period to products from new competitors like Chobani, Fage, and Wallaby. Similar increases in differentiation have also occurred within existing brands. Tostitos chips, for example, now complements its original line of chips with products such as “Scoops” and “Cantina Traditional”, but most households tend to concentrate purchases in one of these varieties rather than spreading evenly over the three. Overall this evidence suggests that the simultaneous rise in household concentration and decline in aggregate concentration is driven in part by the introduction of new products which cater to heterogeneous preferences.

In the second half of the paper, we develop a new model of product demand to better understand the underlying forces driving these trends, and to interpret their magnitudes and implications for welfare and market power. Many standard models cannot be used to confront these new empirical facts. For instance, the most common symmetric love-of-variety models imply that spending is evenly distributed across all goods. Standard discrete choice models imply that household spending within categories is completely concentrated on a single product. Any representative household model will, by construction, exhibit identical aggregate and household concentration. Instead, our model follows [Li \(2018\)](#) in featuring constant elasticity of substitution (CES) preferences but with product-specific taste shocks and a utility cost borne per variety consumed, which implies households consume only an endogenously determined subset of the available products. Assuming these tastes for products, adjusted for prices, are distributed Pareto, we obtain a closed-form expression relating the household’s Herfindahl to structural parameters in the model.

To move from an analysis of individual households to an analysis of aggregate spending, we introduce a continuum of households. We assume that tastes for all households decline identically from their favorite product to their second favorite, and so on (i.e. they have identical taste *distributions*). However, the actual identity of these first- and second-favorite products can differ from one household to the next. To capture this, we introduce a “rank” function, which maps each product to a relative position in each household’s tastes. The rank function is a weighted average of a common aggregate component and a random household-specific component. If the aggregate component receives all the weight, the environment collapses to a representative household economy with all households consuming the same products and with equal household and aggregate spending concentration. Conversely, if the household-specific component receives all the weight, there will be uniform aggregate spending across products and low aggregate concentration, even if individual household spending

is highly concentrated. We analyze an empirically-disciplined intermediate case and obtain another closed-form expression relating the aggregate Herfindahl to structural parameters in the model.

We start our analysis of the model by confirming that its basic predictions for spending patterns are broadly in line with relevant empirical moments. For example, the relationship between household Herfindahls and the number of consumed varieties that is implied by the model aligns well with the data. Further, the shape of the aggregate market share distribution in the model is consistent with that in the data for most product categories. Next, we confront the model's analytical expressions with the empirical trends in household and aggregate concentration to back out the key driving forces. The model is not fully identified, and multiple combinations of parameter values can allow it to fit the trends, but we demonstrate that we can limit or eliminate the salience of a number of otherwise plausible shocks, such as changes in the distribution of each household's tastes or the elasticity of substitution across products in demand. Instead, we show that quantitatively matching the relative movements in the Herfindahls requires an increase in the number of available varieties of about 70 percent, or a bit less than 5 percent per year, regardless of the other shocks involved in matching other moments.<sup>1</sup>

Why does this increase in product availability matter? In our model, an increase in the total number of varieties available leads to an increase in the number of varieties consumed by an individual household. As is standard in CES environments, this increase in the number of varieties consumed is welfare enhancing. However, this is not the primary effect of changing aggregate variety availability in our model. Rather, an increase in the total number of varieties available allows households to select a subset of products better matched to their particular tastes, resulting in welfare increases even with no increase in the number of varieties consumed by individual households. In a counterfactual exercise where we hold all else equal and raise the count of varieties available by the 70 percent implied by our model, we find that households experience consumption-equivalent welfare gains of 9.5 percent. This welfare gain from selection shows up in the ideal price index, but need not be captured by typical matched-model price indices used by national statistical agencies.

We note that while the model can generate the divergence in household and aggregate Herfindahls with expanded product availability, additional shocks are required to match the increase in household concentration. The model suggests that the two parameters in the per-consumed-variety utility cost, together with expanded product varieties, are the most likely candidates to allow the model to deliver both Herfindahl trends, rather than just their difference. Adding these shocks to the counterfactual does not dramatically alter our conclusions about overall welfare or the importance of selection gains.

Finally, given that changes in concentration are often referenced as proxies for changes in market power, we turn our focus to the implied distribution of markups in the model. Different products have different elasticities of demand because expansions in overall sales come in part from selling

---

<sup>1</sup>As discussed in our empirical results, there are a number of data challenges with directly measuring aggregate variety availability, so the strength of this channel is difficult to measure without our model structure.

more to existing customers (the intensive margin) and in part from selling to new customers (the extensive margin), and the relative importance of each margin varies across the product distribution. For example, the product with the largest amount of sales in the economy has many existing customers. Its elasticity of demand will mostly reflect the price-sensitivity of those customers, leading its producer to charge some non-zero markup. By contrast, the marginal product in the economy, with zero sales, effectively faces an infinite elasticity of demand as lowering its price may cause sales to jump from zero to non-zero. Its producer will set price at marginal cost.

We can characterize each product's markup analytically as a decreasing function of its percentile in the ranking of purchased products, with larger market-share products ranked lower and charging higher markups. Increases in the number of available products improves selection, which makes the set of incumbent products less preferable (i.e. shifts them to higher percentiles) and leaves them charging lower markups. By contrast, the newly consumed products, on average, are more preferable and have higher markups than those products that are dropped and are no longer consumed. In the aggregate, these two opposing forces exactly offset each other. Despite significant changes in both household and aggregate concentration measures, the cost-share weighted average of product markups – equal to the ratio of aggregate sales to aggregate costs – does not change when the product space expands.

We proceed as follows. Section 2 discusses the related literature, Section 3 demonstrates the empirical divergence between household and aggregate spending concentration, Section 4 develops a theoretical heterogeneous household model to interpret this empirical evidence, and Section 5 concludes.

## 2 Related Literature

Our work touches on and draws connections between a number of important themes in recent research. Several important papers document changes in top sales shares and industrial structure, including [Autor, Dorn, Katz, Patterson, and Reenen \(2017\)](#) and [Furman and Orszag \(2015\)](#). Our finding that household and aggregate concentration trends move in opposite directions is reminiscent of [Rossi-Hansberg, Sarte, and Trachter \(2018\)](#), who demonstrate that concentration trends also diverge when comparing measures done at the zip-code and national levels. Our work is also related to [Kaplan and Menzio \(2015\)](#), [Argente and Lee \(2017\)](#), and [Jaravel \(2019\)](#), who focus on inflation rather than concentration, but similarly use micro data to demonstrate that patterns at the household level may differ significantly from aggregate trends. Like [Jaravel \(2019\)](#), we conclude that innovation and product entry plays a large role in generating this mismatch.

The model with CES preferences, Pareto-distributed (price-adjusted) tastes, and endogenous choice of consumed varieties follows those in [Li \(2018\)](#) and [Arkolakis, Demidova, Klenow, and Rodriguez-Clare \(2008\)](#). Given heterogeneity across households, our structure generates heterogeneous markups because some producers adjust sales by selling more to existing customers while others adjust by sell-

ing non-zero amounts to more customers. To our knowledge, [Levin and Yun \(2008\)](#) is the only other paper in the recent literature that emphasizes this mechanism, though it also relates to [Hottman, Redding, and Weinstein \(2016\)](#), who emphasize heterogeneity in the degree to which price declines for one product cannibalize sales for others in multiproduct firms. Our emphasis on differences across firms in the importance of the intensive versus extensive margin contrasts with the more commonly used frameworks for generating variable markups, such as the linear demand environment in [Melitz and Ottaviano \(2008\)](#), nested-CES setup in [Atkeson and Burstein \(2008\)](#), translog preferences in [Feenstra and Weinstein \(2017\)](#), and [Kimball \(1995\)](#) kinked-demand curves as incorporated in [Gopinath and Itskhoki \(2010\)](#).

Our framework delivers analytical expressions for the full distribution of markups, a topic of increasing focus, such as in the work of [De Loecker and Eeckhout \(2017\)](#), [Edmond, Midrigan, and Xu \(2018\)](#), [Stroebel and Vavra \(2019\)](#) and [Anderson, Rebelo, and Wong \(2018\)](#). We note, however, that our model can easily deliver large trends in aggregate and household concentration without requiring any change in aggregate market power. Our work is therefore consistent with the skepticism expressed in [Syverson \(2018\)](#) and [Berry, Gaynor, and Morton \(2019\)](#) of the simple linkage often made between concentration trends and market power.

Our framework suggests that the divergent concentration trends reflect an increase in the number of products, which benefits consumers through a selection effect above-and-beyond the standard love-of-variety effect emphasized in [Feenstra \(1994\)](#) or [Broda and Weinstein \(2004, 2006\)](#).<sup>2</sup> In concurrent work, [Michelacci, Paciello, and Pozzi \(2019\)](#) document cyclical fluctuations in household variety adoption and model this phenomenon using a discrete choice model. Their empirical focus is on higher frequency business cycle effects, and their theoretical framework is very different from ours, but they reach similar conclusions about the important role of product selection for welfare. [Argente, Lee, and Moreira \(2018a,b\)](#) show that product introduction plays a key role in understanding patterns of firm growth. More broadly, our analysis relates to the large set of rich discrete choice models used in industrial organization to quantify welfare gains from particular new varieties or industries, such as those used in [Petrin \(2002\)](#) or [Berry, Eizenberg, and Waldfogel \(2016\)](#).<sup>3</sup> Our model allows for gains from selection which are often stressed in this literature but are absent from typical macro models.

Finally, although the underlying causes are potentially different, the rise in niche consumption of retail goods parallels the increasing segmentation or polarization witnessed in culture and digital content ([Aguado, Feijoo, and Martinez \(2015\)](#); [Alwin and Tufis \(2015\)](#)), in political ideology ([Pew Research Center \(2014\)](#); [Gentzkow, Shapiro, and Taddy \(2017\)](#)), in jobs and income ([Autor, Katz, and](#)

---

<sup>2</sup>[Handbury and Weinstein \(2014\)](#) emphasize the need to account for differences in variety availability when comparing the price level across U.S. cities. [Redding and Weinstein \(2016\)](#) demonstrate that welfare measures in CES environments are biased unless they account for heterogeneity in consumer tastes across products. [Atkin, Faber, and Gonzalez-Navarro \(2018\)](#) use similar scanner data on grocery purchases to calculate the welfare gains associated with entry of global retail chains into the Mexican market.

<sup>3</sup>See [Kroft, Laliberté, Leal-Vizcaíno, and Notowidigdo \(2017\)](#) for a sufficient statistic approach to the gains from new varieties.

Kearney (2006); Piketty, Saez, and Zucman (2016)), and in the geography of where households consume (Davis, Dingel, Monras, and Morales (2017)).<sup>4</sup> Our findings indicate that, along with these other manifestations of fragmentation in modern life, even our grocery purchases increasingly differ from the national average.<sup>5</sup>

### 3 Diverging Household and Aggregate Concentration

We start this section with a discussion of the aspects of the data that are particularly salient for our analysis, relegating a more detailed description to Appendix A. We then present our key finding that the concentration of household spending across products increased while, at the same time, aggregate concentration among the same goods decreased. Finally, we provide evidence that these trends are associated with product churning.

#### 3.1 AC Nielsen Homescan Data

We use Homescan data from AC Nielsen to measure household-level shopping behavior.<sup>6</sup> The data set contains a weekly household-level panel for the period 2004-2016. The panel has large coverage, with roughly 170,000 households in over 22,000 zip codes recording prices for almost 700 million unique transactions covering a large fraction of non-service retail spending.

Products are allocated by Nielsen into three levels of category aggregation: roughly 1304 product modules, 118 product groups, and 11 department codes. For example, "vegetables - peas - frozen" are a typical product module within the "vegetables - frozen" product group within the "frozen foods" department. Our baseline analysis focuses on annual spending by all households in the Nielsen sample and computes household spending shares across products within product groups, but all results are qualitatively robust to instead calculating household product spending shares within the more disaggregated product modules or within the more aggregated department codes.<sup>7</sup> We focus on the full sample of households for a number of reasons discussed in Appendix C, but this is relatively conservative since the magnitudes of our trends increase when we restrict to a balanced panel of households.<sup>8</sup>

---

<sup>4</sup>The fact that our results are not driven by a widening gap between the goods purchased by rich and poor households or between consumers in one region and another is also consistent with the finding in Bertrand and Kamenica (2018) that cultural distance between rich and poor has not grown over time.

<sup>5</sup>Our findings are also broadly consistent with forecasts of growing importance of "long-tail" consumption (Anderson (2006)) and shows that this phenomenon extends beyond e-commerce to broader retail spending.

<sup>6</sup>These data are available for academic research through a partnership with the Kilts Center at the University of Chicago, Booth School of Business. See <http://research.chicagobooth.edu/nielsen> for more details on the data. They have also been analyzed recently in Stroebel and Vavra (2019), Jaravel (2019), and Allcott, Diamond, Dube, Handbury, Rahkovsky, and Schnell (2017).

<sup>7</sup>All results weight each household using sampling weights provided by Nielsen, which are designed to make the Nielsen panel demographically representative of the broader U.S. population. Appendix Figure A2 shows that aggregate spending growth in our sample tracks government data on aggregate spending growth in comparable categories.

<sup>8</sup>In the appendix we also discuss the relevance for our results of additional measurement-related issues, such as the (unimportant) role of online shopping.

In our baseline analysis, we define a product as a Universal Product Code (UPC). Appendix B demonstrates, however, that our results are robust to instead defining a product as a "brand". Nielsen assigns UPCs to brands, which are more aggregated than UPCs but are still fairly disaggregated. "Pepsi", for example, is a brand and includes many different flavors and package sizes of the Pepsi drinks. "Caffeine Free - Pepsi", however, is considered a distinct brand. The UPC is our preferred notion of a product in part because UPCs are directly assigned by the manufacturer, whereas the brand variable is constructed by Kilts/Nielsen in a way that involves judgment and may differ across categories and over time. Further, although each generic has a unique UPC, all generics are assigned the same brand in order to preserve the anonymity of the stores in the Nielsen sample. Sales of generics are large and growing, so their inclusion, by construction, distorts concentration measures that define products as brands.<sup>9</sup> Finally, some of our analyses decompose expenditure changes into price and quantity effects, which is straightforward for the case of UPCs but not for brands.

We restrict our analysis to the set of product modules in the data for all years during 2004-2016. We exclude modules that enter or exit since this reflects changes in Nielsen's measurement – not actual household behavior – and could therefore lead to spurious changes in measured concentration. We also exclude fresh produce and other items without barcodes (these are labeled as "magnet" items in the data).

### 3.2 Household Spending Concentration

We begin our analysis by exploring how the concentration of household spending across products has changed over time. For each household  $i$ , UPC  $j$ , and product group  $c$  we calculate total expenditure  $E_{i,j,c,t}$  in year  $t$  and associated expenditure share:

$$s_{i,j,c,t} = \left( \frac{E_{i,j,c,t}}{\sum_j E_{i,j,c,t}} \right). \quad (1)$$

Our primary measure of household product concentration for a product category  $c$  at time  $t$  is the Herfindahl and equals the sum of the square of these expenditure shares:

$$\mathcal{H}_{i,c,t}^{\text{HH}} = \sum_j (s_{i,c,j,t})^2. \quad (2)$$

Next, we take the weighted average across households to generate the Household Herfindahl for product category  $c$ :

$$\mathcal{H}_{c,t}^{\text{HH}} = \sum_i \text{share}_{i,c,t} \mathcal{H}_{i,c,t}^{\text{HH}}, \quad (3)$$

---

<sup>9</sup>See, for example, Dube, Hitsch, and Rossi (2018). Our robustness checks using the brand definition of product exclude generics. Results using UPCs also remain qualitatively robust if we exclude generics.

where we use weights capturing household  $i$ 's share of aggregate spending in category  $c$ :

$$share_{i,c,t} = \frac{\sum_j (\omega_{i,t} E_{i,j,c,t})}{\sum_i \sum_j (\omega_{i,t} E_{i,j,c,t})}, \quad (4)$$

and where  $\omega_{i,t}$  is a household's sampling weight provided to make the Nielsen sample representative of aggregate consumption. Finally, we calculate the overall Household Herfindahl by averaging the category-specific Household Herfindahl in equation (3) across all categories:

$$\mathcal{H}_i^{\text{HH}} = \sum_c share_c \mathcal{H}_{c,t}^{\text{HH}}, \quad (5)$$

where  $share_c$  is the average share of category  $c$  in total spending across our entire sample.

Unlike the weights used in equation (3), we use fixed category spending shares over time in equation (5) to focus on concentration changes occurring within categories, rather than emerging from shifts in spending across categories with different levels of concentration. We do this to better interact with recent interest in changing market power and technological disruption, typically perceived to be occurring within sectors. Our results are qualitatively robust, though a bit noisier, when we allow compositional shifts across categories to influence our concentration measures.

Figure 1a plots  $\mathcal{H}_i^{\text{HH}}$  and reveals a nearly monotonic increase in Household concentration from 2004-2016.<sup>10</sup> Figure 1b shows that this increase in concentration is also associated with a decline in the average number of products consumed per household within a product category.<sup>11</sup> We delay interpreting the quantitative magnitude of these changes until we develop our model in Section 4 but note now that fitting these data with linear trends yields precise and highly significant estimates.

### 3.3 Aggregate Spending Concentration

What underlies this increase in the concentration of household expenditures? One possible explanation is that there has been an increase in the importance of "super-star products", along the lines of the rise of "super-star firms" documented in [Autor, Dorn, Katz, Patterson, and Reenen \(2017\)](#). This explanation, natural though it may be, finds no support in our data: we demonstrate in this subsection that at the same time the typical household's expenditures have grown more concentrated across products, aggregate spending has in fact become more evenly distributed across these same products.

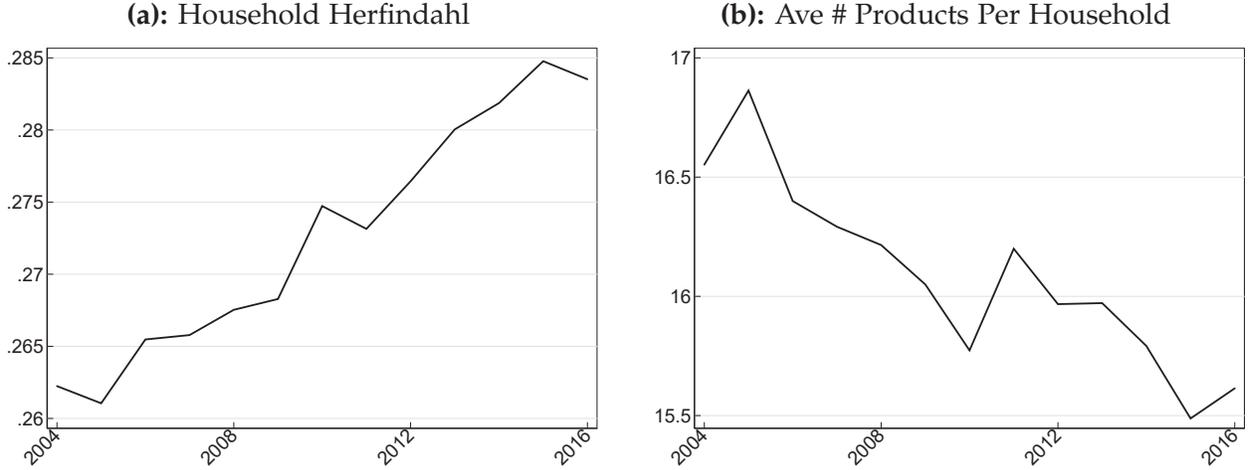
We sum spending on product  $j$  in category  $c$  across all households in our data and define the

---

<sup>10</sup>All results in the paper hold for alternative concentration measures such as the share of spending accounted for by the top 1 or the top 2 products. We use the Herfindahl as our primary concentration measure as it can be more easily interpreted through the lens of the structural model described in Section 4.

<sup>11</sup>Average product number counts are constructed using the same weights as were used in equations (3) and (5).

**Figure 1: Household Product Concentration**



aggregate market share of  $j$  in  $c$  as:

$$s_{j,c,t} = \frac{\sum_i (\omega_{i,t} E_{i,j,c,t})}{\sum_i \sum_j (\omega_{i,t} E_{i,j,c,t})}, \quad (6)$$

and the Aggregate Herfindahl in category  $c$  as:

$$\mathcal{H}_{c,t}^{\text{Agg}} = \sum_j (s_{j,c,t})^2. \quad (7)$$

Just as with the Household Herfindahl, we average these category Herfindahls using fixed category expenditure weights over time to generate the Aggregate Herfindahl of overall spending. Figure 2a plots this Aggregate Herfindahl and shows that the trend in product spending at the aggregate level is the reverse of what we see at the household level: aggregate spending concentration is declining, not rising.

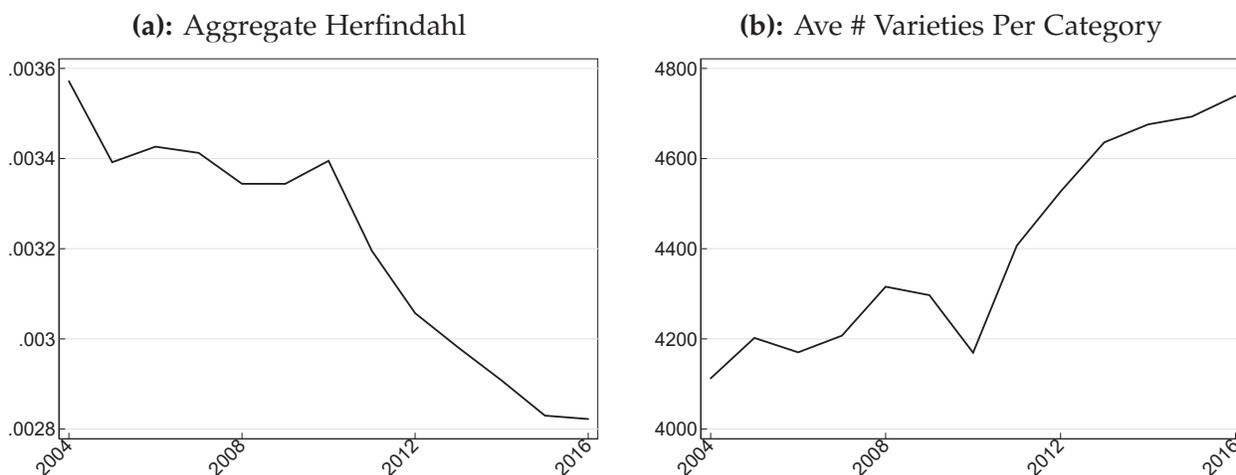
A decline in the number of products consumed by the typical household contributed to the rising Household Herfindahl measure, but we cannot easily observe the equivalent notion for the aggregate economy. The existence of thousands of products with tiny amounts of overall sales and incomplete coverage of households and stores in the data render a simple product count highly volatile, dependent on assumptions, and sensitive to measurement error. We therefore treat the total number of products as unobservable and, in Section 4, will use our model to infer it. To get a basic idea, however, we keep a constant number of households over time and count only products purchased by at least two households and with at least \$50 in aggregate annual spending, which generates a relatively stable measure, plotted in Figure 2b. This measure of the number of aggregate products is rising, a trend

that is also the reverse of its household-level equivalent.<sup>12</sup>

This result may, at first, seem at odds with the rise in sales concentration measured in Census data by papers including [Autor, Dorn, Katz, Patterson, and Reenen \(2017\)](#). Our aggregate concentration measure, however, captures expenditures at the product level whereas Census-based estimates aggregate products to the producer level.<sup>13</sup> The resulting trends may therefore differ significantly, particularly in the face of changes in the number of goods each manufacturer produces.

In the Appendix, we first show that production concentration measures from the Census for the relevant NAICS categories – “Food Manufacturing” (code 311) and “Beverage and Tobacco Product Manufacturing” (312) – are in fact flat or declining during the years covered in our sample. Next, we use a mapping of UPCs to manufacturers to generate a comparable producer-level concentration measure based on the sales in our Nielsen data. We offer a number of important caveats, including that the UPC-to-manufacturer mapping is highly imperfect for this purpose, but nonetheless find similar trends in manufacturer concentration in Nielsen and Census data.<sup>14</sup> We therefore conclude that our results are broadly consistent with the Census-based literature. Whether producer or product concentration is of greater interest depends, of course, on the question at hand. Our theory below will treat each good as produced and marketed independently such that it maps most naturally to our product-based concentration measure.

**Figure 2: Aggregate Product Concentration**



How can it be that aggregate concentration is declining if households are individually concentrating their spending on a smaller number of products? These divergent trends imply that households are

<sup>12</sup>Results are qualitatively similar with different spending thresholds and are also similar when using retail scanner data, although the latter is not a census of all stores.

<sup>13</sup>The categories within which we calculate concentration are also far less aggregated and cover a smaller set of economic activity than what is done in most Census-based studies. Further, our data begin in 2002, far later than the 1970s or 1980s start date commonly found in that literature.

<sup>14</sup>We further measure in our data an increase in the share of the top retailers or stores, a finding unrelated to our focus but one that closely resembles the results for the Retail industry in [Autor, Dorn, Katz, Patterson, and Reenen \(2017\)](#).

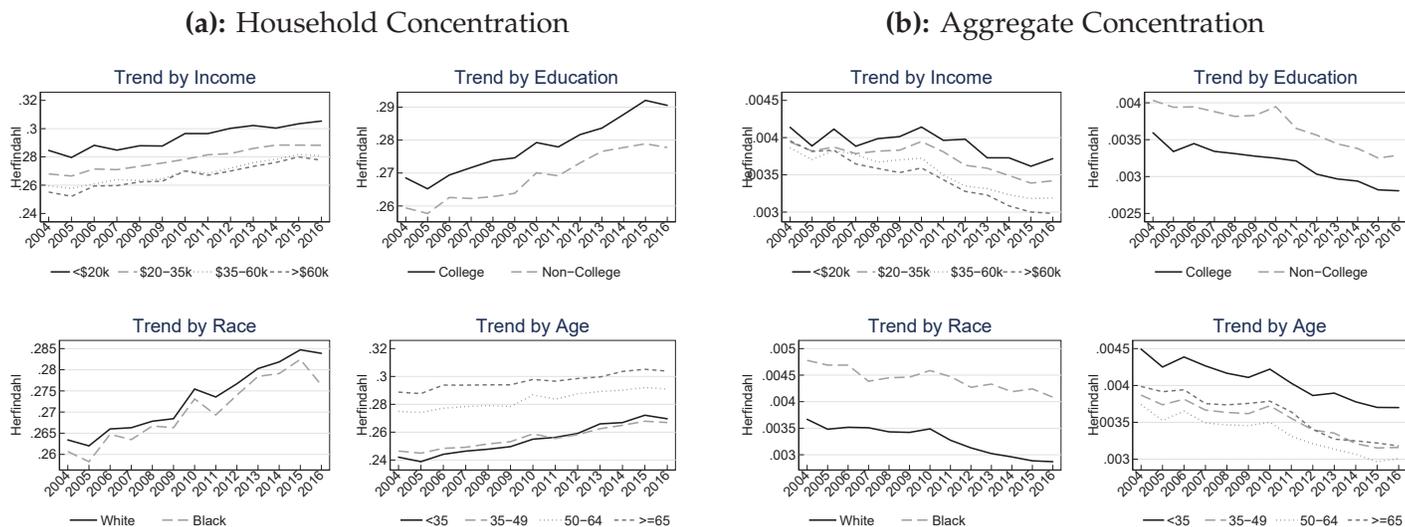
concentrating more and more spending on their top products over time, but that these top products increasingly differ across households. We view these divergent trends and resulting fragmentation of the product space as characterizing a rise in "niche" consumption.

### 3.4 The Pervasive Rise of Niche Consumption

This rise in niche consumption – the increase in the Household Herfindahl and decrease in the Aggregate Herfindahl – is highly robust to various measurement related choices. For example, Appendix Figures A3-A8 show that these divergent trends continue to hold if we exclude generics, compute concentration using more disaggregated categories (modules instead of groups), define products as brands instead of UPCs, use time-varying category weights, use alternative concentration measures instead of the Herfindahl, or focus on a balanced panel of households over time.

Is the rise of niche consumption driven by shifts in the importance of different groups, such as old and young or rich and poor? While there are differences in the level of concentration across different groups, the trends are primarily driven by *within* group variation.<sup>15</sup> To show this, we re-calculate annual Household and Aggregate Herfindahls using only expenditures by households with particular demographic characteristics such as income bracket, race, education, and age. Figures 3a and 3b show that rising household and declining aggregate concentration occurs within demographic groups.<sup>16</sup> The rising Household and falling Aggregate Herfindahls do not simply owe to changes in composition across groups with different levels of concentration.

**Figure 3: Trends within demographic group**



<sup>15</sup>While we primarily emphasize trends, the level differences in household concentration are similar to those documented in Hansen and Singh (2015). See also Bornstein (2018) for an analysis of age-specific results.

<sup>16</sup>See also Appendix Figure A9 showing divergent trends within households of different size. We also find diverging trends also obtain when computing concentration in more and less dense cities and cities with higher and lower Republican vote shares as a proxy for political polarization, which is unsurprising in light of Appendix Figure A10 discussed below.

While Figure 3 focuses on one set of demographics at a time, Appendix Table A1 reports estimated trends in household concentration when we control simultaneously for multiple demographic indicators, including household size. The positive trend in the Household Herfindahl is largely unchanged and remains highly statistically significant.<sup>17</sup>

As a simple summary statistic for the prominence of niche consumption, we consider the ratio of the Household Herfindahl to the Aggregate Herfindahl. A higher value for this “niche ratio” means that household consumption is more segmented into different niches. Figure 4 shows that the rise of niche consumption is pervasive across product categories, with three-quarters of product categories exhibiting increases in  $\mathcal{H}_{c,t}^{HH}$ , eighty percent of product categories exhibiting decreases in  $\mathcal{H}_{c,t}^{Agg}$ , and growth in the niche ratio in 92 percent of the categories.<sup>18</sup>

In Appendix Figure A10, we also show that the rise of niche consumption is occurring in the vast majority of locations, implying that shifts in the relative economic importance of cities and regions are not behind our findings. The niche ratio is highest in cities like Chicago, Washington DC, Tampa, Los Angeles, and Boston and lowest in Des Moines, Little Rock, Las Vegas, and “West Texas”, but it is increasing in most locations. Finally, Appendix Figure A11 shows that the rise in niche consumption is found within roughly two-thirds of the individual retailers in our data, so the aggregate patterns we observe are not simply driven by shifts in where households shop.<sup>19</sup> Together these results all imply that whatever forces are driving the rise in niche consumption, they are pervasive across demographics, geographies, retail chains, and product categories.

The level of the niche ratio is highest in “Cosmetics” and “Fragrances-Women” and is lowest for “Charcoal” and “Dough Products”. The rise in niche consumption is pervasive, but it is also clear from Figure 4 that there is substantial heterogeneity across categories in the extent of its ascent. The niche ratio has grown most rapidly for “Coffee”, “Hardware, Tools”, “Fresheners and Deodorizers”, and “Disposable Diapers”. It has declined by most for “Cottage Cheese”, “Eggs”, “Milk”, and “Bread and Baked Goods”.

### 3.5 The Role of Product Churn

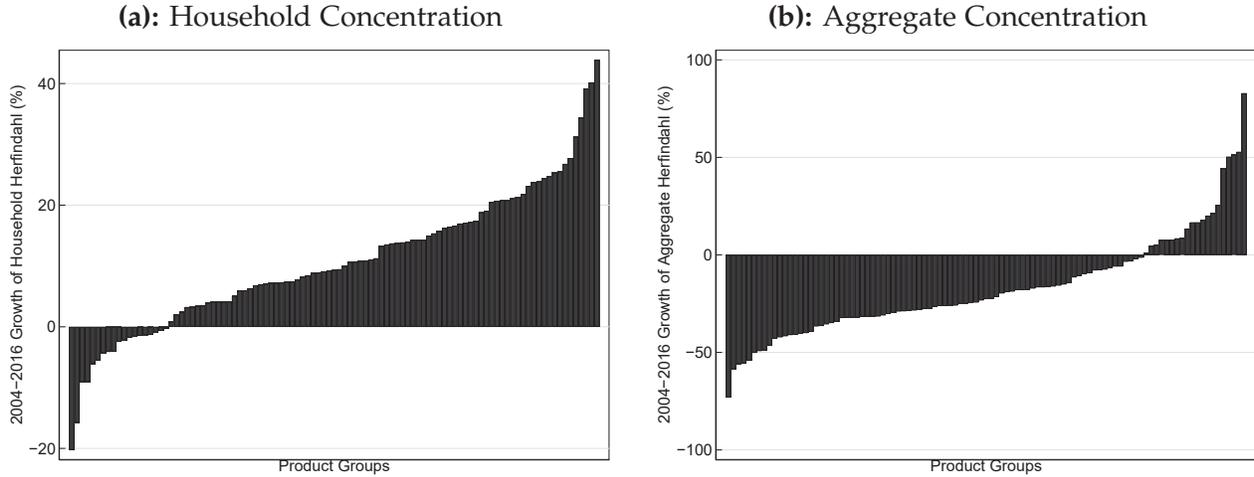
Interestingly, there is a common observable linking together the categories with the most rapid increases in niche ratios: they are also the categories with the fastest growth in the number of aggregate products, measured using the same assumptions as were used for Figure 2b above. We emphasize this

<sup>17</sup>Since  $\mathcal{H}_i^{HH}$  is measured for each household  $i$ , it is straightforward to regress household concentration measures on a variety of simultaneous demographics which vary across households. A similar exercise for aggregate concentration requires recalculating aggregate market shares and  $\mathcal{H}^{Agg}$  separately for each demographic group. This makes these calculations substantially more computationally burdensome. Even more importantly, measurement error in aggregate market shares increases rapidly for more narrow demographic groups.

<sup>18</sup>To improve visual exposition, Figures 4 and 5a drop 5 outlier categories whose variety counts more than double or decrease by more than 50 percent from 2004-2016: “Frozen Juices”, “Yeast”, “Canning Supplies”, “Greeting Cards” and “Photographic Supplies”. This does not affect any conclusions.

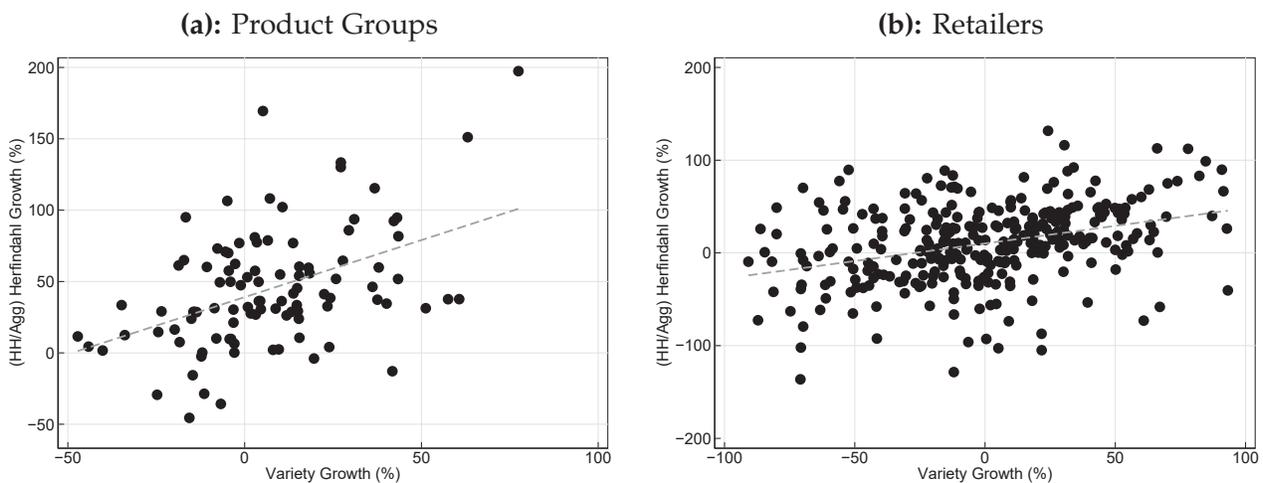
<sup>19</sup>We also show that our results are not driven by trends in the frequency of shopping and bulk purchasing (Coibion, Gorodnichenko, and Koustas (2017)).

**Figure 4: 2004-2016 Concentration Growth Within Category**



relationship as it will be central to the mechanism in our model in Section 4 and its implications for welfare. In particular, Figure 5a shows that categories with 50 percentage points more growth in the total number of products sold had, on average, 40 percentage points more growth in their niche ratios, with the relationship statistically significant at the 1 percent level. Figure 5b shows that a similar relationship also holds when comparing across retailers: retailers with 50 percentage points more growth in the number of products sold exhibited roughly 20 percentage points more growth in their niche ratios, with the relationship again significant at the 1 percent level.<sup>20</sup> The relationship between variety growth and the niche ratio becomes even steeper if we weight retailers by size.

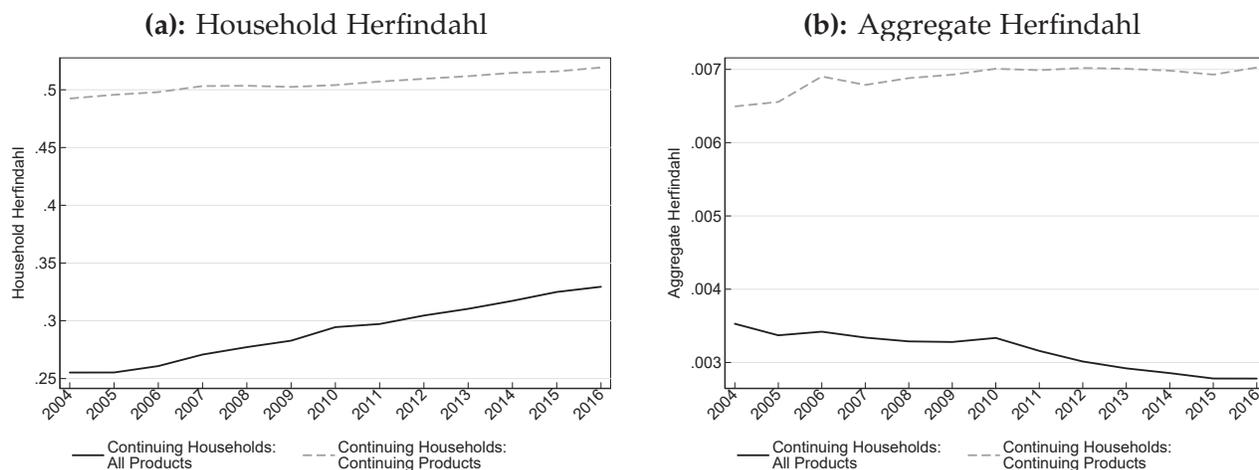
**Figure 5: Growth in Number of Products vs. Growth in Niche Ratio**



<sup>20</sup>To reduce the influence of outliers, we exclude retailers with absolute log variety changes above 2, which drops 6 out of 334 retailers. Results are similar for alternative thresholds. The panel of retailers is unbalanced, and growth rates for the remaining 328 retailers are calculated from their first to their last observation in the sample. Results are very similar if we instead calculate growth rates from 2004-2016 for the 179 retailers which are in the sample continuously.

We now provide additional evidence that product churn plays a key role in the rise of niche consumption by comparing concentration trends measured only among “continuing” products that are purchased by a household in two consecutive years with those measured using all spending by that household.<sup>21</sup> For each household  $i$  that is observed in both  $t$  and  $t + 1$ , we measure concentration of “continuing products” by using only those that are purchased by that household in both  $t$  and  $t + 1$ . These continuing products account for about 30 percent of transactions and 40 percent of spending. We also calculate Herfindahls for those same households using all their spending. We form an index by chaining together changes in these Herfindahls from  $t$  and  $t + 1$  and pin down the level using the values in the initial period.<sup>22</sup> Figure 6a shows the upward trend in household concentration is much stronger when using all UPCs than when restricting to continuing products, growing by 29 percent compared to 5 percent.<sup>23</sup> This implies a large role for product entry and exit in generating household concentration increases. Figure 6b shows that when focusing only on continuing products, aggregate concentration actually rises instead of declines.<sup>24</sup>

**Figure 6: 2004-2016 Concentration growth for continuing vs. all products**



## 4 Modeling the Rise of Niche Consumption

In this section, we develop a model able to match the rise of niche consumption documented in Section 3 and we use it to identify key driving forces and understand the resulting implications for welfare and

<sup>21</sup>For these results, therefore, we move to a panel specification since we require at least two time periods for a household to measure which products for that household are new and which are continuing.

<sup>22</sup>We report cumulative changes rather than levels since this procedure results in two levels of the Herfindahl for each year (one initial and one continuation year), so the levels are not meaningful.

<sup>23</sup>The trend for “All Products” is larger than that in Figure 1a as Figure 6a is calculated using within-household variation. See Appendix C for details.

<sup>24</sup>While this analysis shows that the diverging concentration trends are largely driven by extensive rather than intensive margin effects, for continuing products we can nevertheless decompose intensive margin concentration trends into P vs. Q effects. Appendix Figure A12 shows that changes in quantity rather than price are more important.

markups. Many standard models are not useful for studying a simultaneous increase in household concentration and decline in aggregate concentration because they either assume that all households consume a single product, assume that household tastes are symmetric across products, or assume that all households are identical. In our model, households choose how many products to consume, spend different amounts on each good, and differ from other households in their choice of which products to buy.

Following Li (2018), we assume that households must pay a fixed cost per consumed product, which implies they only consume a subset of available products despite their CES preferences that embed a “love of variety”. As in Arkolakis, Demidova, Klenow, and Rodriguez-Clare (2008), we assume each household’s tastes for products, adjusted for price, are distributed Pareto, which allows us to write the Household Herfindahl analytically. Further, we introduce a rank function that implies the preference ordering of products will differ across households, which allows us to write the Aggregate Herfindahl analytically.

Using the analytical expressions for the Household and Aggregate Herfindahl, we confront the model with empirical trends from 2004-2016 and back out the implied driving forces. We find that an increase in the number of available products is required to quantitatively match the rise of niche consumption. In the model, this increase leads to significant welfare gains as it implies that consumers can choose a consumption bundle better tailored to their tastes without raising their fixed cost expenditures.

Finally, the model features heterogeneous markups across products. Growth in the sales of products with larger aggregate market shares primarily reflect growth in spending by existing customers, or intensive margin adjustment. By contrast, growth in the sales of small products are more likely to come from the addition of new customers, or extensive margin adjustment. Intensive and extensive margin adjustments are characterized by different elasticities of demand, and this results in heterogeneous markups. While changes in the number of products shift markups across products, we demonstrate that they do not impact the aggregate degree of market power in the economy.

## 4.1 Household Problem

We assume that a continuum of households  $i \in [0, 1]$  spend  $E$  on a continuum of varieties  $k \in [0, N]$  to maximize:<sup>25</sup>

$$U_i = \left( \int_{k \in \Omega_i} (\gamma_{i,k} C_{i,k})^{\frac{\sigma-1}{\sigma}} dk \right)^{\frac{\sigma}{\sigma-1}} - F \times (|\Omega_i|)^{\epsilon}, \quad (8)$$

where  $\Omega_i$  is the set of products consumed by  $i$  (with  $|\Omega_i| \leq N$ ),  $\gamma_{i,k}$  is a household-specific taste for product  $k$ , and the term multiplied by  $F$  captures a fixed cost that increases exponentially in the

---

<sup>25</sup>To ease notation, we do not index  $k$  by  $i$ , but importantly note that the same  $k$  may represent a different actual product for each different households. This is unimportant for the analysis of individual households, but will be crucial when we move to the aggregate analysis.

measure of varieties consumed.

We write the price of product  $k$  as  $p_k$ , so  $\tilde{\gamma}_{i,k} = \gamma_{i,k}/p_k$  captures the price-adjusted taste of household  $i$  for  $k$ . We assume price-adjusted tastes are distributed Pareto:

$$Pr(\tilde{\gamma}_{i,k} < y) = G(y) = 1 - (y/b)^{-\theta},$$

where  $y \geq b > 0$  and where we assume  $\theta > 2(\sigma - 1)$ . Since larger  $\theta$  means a flatter distribution of tastes, the latter condition simply ensures that tastes are not "too concentrated" relative to  $\sigma$  and that the model delivers a finite Household Herfindahl. We also assume  $\epsilon > 1/(1 - \sigma) - 1/\theta$ , which implies that higher fixed costs  $F$  lead to less purchased products  $|\Omega|$ . Household  $i$  will consume the set of goods with  $\tilde{\gamma}_{i,k} \in [\tilde{\gamma}^*, \infty)$  for some  $\tilde{\gamma}^* \geq b$ .

The ideal price index in this environment will be equal for all households and is defined as:

$$\begin{aligned} P_i = P &= \left( \int_{k \in \Omega_i} (\tilde{\gamma}_{i,k})^{\sigma-1} dk \right)^{\frac{1}{1-\sigma}} \\ &= \underbrace{\left( 1 + \frac{1-\sigma}{\theta} \right)^{\frac{1}{\sigma-1}} b^{-1}}_{\text{Average Price}} \times \underbrace{(|\Omega_i|)^{\frac{1}{1-\sigma}}}_{\text{Variety Gains}} \times \underbrace{\left( \frac{|\Omega_i|}{N} \right)^{\frac{1}{\theta}}}_{\text{Selection Effects}}. \end{aligned} \quad (9)$$

The price index has three terms, each with an intuitive interpretation. We refer to the first term as the average price since it summarizes the full distribution of price-adjusted tastes for available products as if there were a single purchase price for one unit of the full bundle. It varies with the shape  $\theta$  and scale  $b$  of the Pareto distribution as well as with the elasticity of substitution  $\sigma$ . The second term is the standard love-for-varieties term in CES models, which decreases with the measure of consumed products and with the elasticity of substitution (given  $|\Omega_i| > 1$ ). Finally, the third term represents a selection effect from the fact that when households only consume a subset  $\Omega_i$  of the full measure  $N$  of products, they choose the subset they like best. This term decreases in the share of products that are consumed and in the extent to which households prefer some products to others.

The index reduces to more standard expressions in special cases. For example, consider  $\theta \rightarrow \infty$ , which implies that households value all products identically at  $b$ , i.e.  $\tilde{\gamma}_{i,k} = b$  for all  $i$  and  $k$ . In such a case, the expression reduces to  $b^{-1}|\Omega_i|^{1/(1-\sigma)}$ , which is the standard price index for symmetric CES preferences. Alternatively, imagine some products are preferred to others,  $\theta < \infty$ , but all products are nonetheless purchased,  $\Omega_i = N$ . In this case, the last term reduces to 1 as there are no selection effects and the average price term fully captures impact of heterogeneity in the desirability of the products.

The properties of the CES price index imply we can re-write equation (8) as:

$$U_i = \frac{E}{P} - F \times (|\Omega_i|)^\epsilon.$$

Consumers choose  $|\Omega_i|$  to maximize utility. The first order condition implies that the optimal number of products is:

$$|\Omega_i| = |\Omega| = \left( \frac{\left(\frac{1}{1-\sigma} - \frac{1}{\theta}\right) \left(1 + \frac{1-\sigma}{\theta}\right)^{\frac{1}{1-\sigma}} N^{\frac{1}{\theta}}}{\tilde{F}\epsilon} \right)^{\left(\epsilon - \frac{1}{1-\sigma} + \frac{1}{\theta}\right)^{-1}}, \quad (10)$$

where  $\tilde{F} = F/(bE)$  is a parameter which shifts spending, aggregate prices, and variety costs.<sup>26</sup> Importantly, the optimal choice of varieties yields a "cutoff" taste  $\tilde{\gamma}^*$  that satisfies:  $G(\tilde{\gamma}^*) = 1 - \frac{|\Omega|}{N}$ , and the share of household  $i$ 's expenditure on variety  $k$  is then given by

$$s_{i,k} = \begin{cases} (P\tilde{\gamma}_{i,k})^{\sigma-1}, & \tilde{\gamma}_{i,k} > \tilde{\gamma}^* \\ 0, & \tilde{\gamma}_{i,k} \leq \tilde{\gamma}^* \end{cases} \quad (11)$$

with  $\int_k s_{i,k} dk = 1$ .

## 4.2 Household Herfindahls

Given equation (11), it follows that the Household Herfindahl  $\mathcal{H}^{\text{HH}}$  will be equal for all  $i$  and can be written as:

$$\begin{aligned} \mathcal{H}^{\text{HH}} &= \mathcal{H}_i^{\text{HH}} = \int_{k \in \Omega_i} (s_{i,k})^2 dk = N \int_{\tilde{\gamma}_i^*}^{\infty} (P_i \tilde{\gamma}_{i,k})^{2(\sigma-1)} dG(y) \\ &= \frac{(\eta + 1)^2}{4\eta} \frac{1}{|\Omega|}, \end{aligned} \quad (12)$$

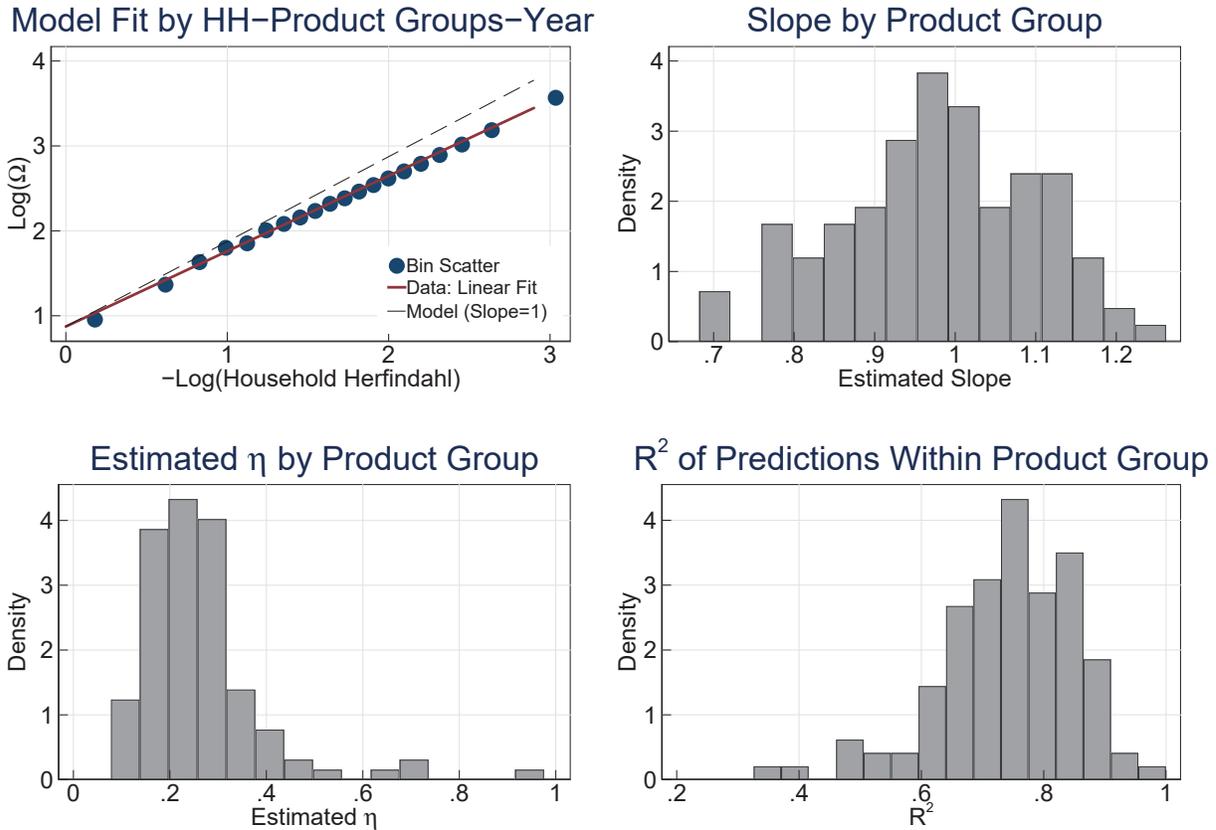
where we introduce the variable  $\eta = 1 - 2(\sigma - 1)/\theta$ . The above parameter restrictions imply  $\eta \in (0, 1)$ . For fixed  $\theta$  and  $\sigma$ , which implies fixed  $\eta$ , household concentration declines monotonically with the number of consumed varieties. And for fixed  $|\Omega|$ , concentration declines monotonically with  $\eta$ . All else equal, flatter taste distributions (higher  $\theta$ ) or less substitutability across products in preferences (lower  $\sigma$ ) reduce Household Herfindahls.

How well does this model fit household spending data? Interpreting our model as applying to each household's spending decisions for a given product category  $c$  in a given year, we have the testable prediction that  $\mathcal{H}_{i,c}^{\text{HH}}$  is proportional to  $1/|\Omega_{i,c}|$ . Indeed, when we pool categories, years, and households and regress  $\ln |\Omega_{i,c}|$  on  $-\ln \mathcal{H}_{i,c}^{\text{HH}}$  (with category-year fixed effects), we get a coefficient of 0.89 – close to the model-consistent value of 1 – and a large  $R^2$  of 0.82. The upper left panel of Figure 7 shows a binscatter (with category-year fixed effects) of the 54 million observations underlying this regression to demonstrate that linearity with a coefficient of 1 is a close approximation to the raw data.<sup>27</sup> In the upper right panel, we estimate these regressions separately for each category in 2016 and

<sup>26</sup>When  $N = 1$ , this expression is the same as that in Li (2018) after substituting in his special case for  $b$ .

<sup>27</sup>This specification has large explanatory power even though it only allows  $\eta$  to vary across category-years and not across households. With arbitrary heterogeneity in  $\eta$  across households within category-years, there would be as many parameters as observations so it would be trivial to perfectly fit the data.

Figure 7: Model Fit on Household-Category Data



plot a histogram of the estimated slopes. The values are largely clustered around the model-consistent value of 1.

Next, rather than estimating the slope, we constrain it to equal 1 and back out the  $\eta$  values implied for each category. The model imposes the restriction that  $0 < \eta < 1$  and the bottom left panel of Figure 7 shows that this restriction is satisfied in every category. The values of  $\eta$  range from lows of 0.08 (Baby Food) and 0.10 (Carbonated Beverages) to highs of 0.69 (Greeting Cards) and 0.97 (Yeast).<sup>28</sup> Finally the lower right panel shows that the  $R^2$ 's from these restricted regressions are generally high. Overall, we conclude that the empirical relationship between household-level concentration measures and the number of consumed products is consistent with the relationships implied in our model.

### 4.3 Aggregation

In order to account for divergent trends in household and aggregate concentration measures, we must specify how tastes for particular products differ across households. We index all products in the

<sup>28</sup>The value of 0.08 for baby food implies that the typical household in this category has spending which is almost 4 times more concentrated than if that household spent evenly on all the baby food products they consumed, while the value of 0.97 for yeast implies that household spending in that category is essentially evenly divided across products. With homogeneous tastes across products (i.e.  $\theta \rightarrow \infty$ ) – the setup in many standard models – we cannot capture this large extent of sectoral heterogeneity in concentration as  $\eta \rightarrow 1$ .

economy with  $j \in [0, N]$ , and assume each household assigns each product a “rank”, where lower ranks indicate higher price-adjusted tastes. Households will consume all goods which they rank less than or equal to  $|\Omega|$ .

We introduce the following rank function for each household  $i$ :

$$r_{i,j} = (1 - \alpha)j + \alpha x_{i,j}, \quad (13)$$

where  $j$  identifies a common aggregate rank for a product,  $x_{i,j}$  is an i.i.d. draw from the uniform distribution with support  $[0, N]$  representing a household-specific taste component, and  $\alpha \in (0, 1)$ . If  $\alpha$  is close to zero, the model approximates a representative agent model where all households rank products in the same order. If  $\alpha$  approaches 1, tastes are purely idiosyncratic and resulting aggregate spending will be evenly distributed over all consumed products even if individual household tastes are very concentrated. Thus, even though all households have identical *distributions* of taste-adjusted prices, this rank function allows for different households to have different ranks for the exact same product  $j$ .

To compute the aggregate spending share on product  $j$ , we need to know the cumulative distribution function (CDF) of product ranks  $R(r)$ , integrating over all household and products. Without loss of generality, we assume  $\alpha < 1/2$  and write:<sup>29</sup>

$$R(r) = \begin{cases} \frac{1}{2} \left( \frac{r}{N} \right)^2 \frac{1}{\alpha(1-\alpha)}, & 0 \leq r < \alpha N \\ \frac{r}{N} \frac{1}{1-\alpha} - \frac{1}{2} \frac{\alpha}{1-\alpha}, & \alpha N \leq r < (1-\alpha)N \\ -\frac{1}{2} \left( \frac{r}{N} \right)^2 \frac{1}{\alpha(1-\alpha)} + \frac{r}{N} \frac{1}{\alpha(1-\alpha)} - \frac{1}{2} \left( \frac{\alpha}{1-\alpha} + \frac{1-\alpha}{\alpha} \right), & (1-\alpha)N \leq r \leq N. \end{cases} \quad (14)$$

Note that this CDF satisfies the properties that  $R(0) = 0$ ,  $R(N) = 1$ , is continuous at  $r = \alpha N$  and  $r = (1 - \alpha)N$ , and is monotonically increasing.

There are three distinct regions in  $R(r)$  with different functional forms. If households only consume goods with ranks in the first region, this implies that there is no single product in the economy that is purchased by all households. If households consume so many varieties that some have ranks in the second region, this implies that at least one product is purchased by everyone. Finally, if even the worst possible product in the economy is purchased by at least one household, then the ranks of some consumed goods will fall into the third region.<sup>30</sup> As long as  $0 \leq \frac{|\Omega|}{N} < \frac{\alpha}{2(1-\alpha)} < \frac{1}{2}$ , it can be shown that

<sup>29</sup>Replacing  $\alpha$  with  $1 - \alpha$  in all instances in equation (14) yields the corresponding  $R(r)$  for the alternative case of  $\alpha > 1/2$ . Furthermore, this leaves the rank function unchanged for the first of the three regions of  $R(r)$ , which will be the focus of our analysis.

<sup>30</sup>More specifically, the product with the best aggregate taste shock is  $j = 0$ . The worst possible idiosyncratic rank for this product occurs when  $x_{i,j} = N$ , in which case  $r = \alpha N$ , so if we are in the first region of the parameter space, even the best product is not purchased by some households. Conversely, the product with the worst aggregate taste shock is  $j = N$ . The best possible idiosyncratic rank for this product occurs when  $x_{i,j} = 0$ , in which case  $r = (1 - \alpha)N$ . This means that if we are in the third region of the CDF, this worst product will still be consumed by some household.

all consumed products in the economy will have an  $r$  value confined to the first region of  $R(r)$ . This is the empirically relevant region of the parameter space, since the number of varieties purchased by an individual household is orders of magnitude less than the aggregate number of varieties, and there are no varieties in the data that are consumed by all households. To simplify analytical solutions, we thus impose this parameter restriction for the remainder of the analysis.

Noting that  $\tilde{\gamma}_{i,j} = G^{-1}(1 - R(r_{i,j}))$ , the spending share of household  $i$  that is dedicated to product  $j$  can be written as a function of  $j$ 's rank:

$$s_{i,j} = P^{\sigma-1} \tilde{\gamma}_{i,j}^{\sigma-1} = (Pb)^{\sigma-1} (R(r_{i,j}))^{-\frac{\sigma-1}{\theta}} = \frac{\eta+1}{2} N^{\frac{\eta-1}{2}} |\Omega|^{-\frac{\eta+1}{2}} (R(r_{i,j}))^{\frac{\eta-1}{2}}, \quad (15)$$

if  $R(r_{i,j}) \leq |\Omega|/N$  and zero otherwise. To determine the products for which the share  $s_{i,j}$  in equation (15) jump from positive to zero, we solve for the rank of the marginal, or least-preferred, variety that is consumed in positive quantities by household  $i$ . Note that this good's identity will differ across households, but its rank  $r^*$  will be the same and satisfies  $R(r^*) = |\Omega|/N$ . Substituting into equation (14) under the assumption that  $0 \leq \frac{|\Omega|}{N} < \frac{\alpha}{2(1-\alpha)}$ , we get:

$$r^* = (2\alpha(1-\alpha)|\Omega|N)^{\frac{1}{2}}. \quad (16)$$

Under our parameter restrictions, individual households each consume only a fraction of the total products available  $N$ , but the exact products consumed will differ across households. However, even when aggregating across all households, there are some products which are consumed by no households. This means that for the economy as a whole, there is a difference between the measure of available goods  $N$  and the measure of goods that are actually consumed, which we denote  $j^*$ . This marginal consumed good for the economy as a whole,  $j^*$ , is that  $j$  for which the best possible idiosyncratic taste draw – a draw of zero – yields rank  $r^*$  for the household with that zero draw. Solving for this cutoff,  $j^* = r^*/(1-\alpha)$ , we get:

$$j^* = \left( \frac{2\alpha|\Omega|N}{1-\alpha} \right)^{\frac{1}{2}}. \quad (17)$$

Importantly, since  $r_{i,j}$  is strictly increasing in  $j$ , all goods with  $j < j^*$  will have positive aggregate sales and all goods with  $j \geq j^*$  will have zero aggregate sales. Finally, substituting in the definition of the rank function from equation (13) into the expression (16), and using the definition of  $j^*$  in equation (17), we can write the highest value or “cutoff” random draw  $x_j^*$  that yields positive consumption of  $j$  as:

$$x_j^* = \frac{1-\alpha}{\alpha} (j^* - j). \quad (18)$$

#### 4.4 The Aggregate Herfindahl

We now integrate spending shares across households  $i$  to get the aggregate spending share on good  $j$ :

$$\begin{aligned}
s_j &= \frac{1}{\int_i Edi} \int_i Es_{i,j} di \\
&= \frac{\eta+1}{2} N^{\frac{\eta-1}{2}} |\Omega|^{-\frac{\eta+1}{2}} \int_0^{\frac{1-\alpha}{\alpha}(j^*-j)} 2^{\frac{1-\eta}{2}} N^{1-\eta} \left( \alpha (1-\alpha)^{\frac{1-\eta}{2}} \right) ((1-\alpha)j + \alpha x_{i,j})^{\eta-1} \frac{dx}{N} \\
&= (\eta+1) (2N|\Omega|)^{-\frac{\eta+1}{2}} (\alpha(1-\alpha))^{\frac{1-\eta}{2}} \int_0^{\frac{1-\alpha}{\alpha}(j^*-j)} ((1-\alpha)j + \alpha x_{z,i})^{\eta-1} dx \\
&= \frac{\eta+1}{\eta} \left( \frac{2\alpha N|\Omega|}{1-\alpha} \right)^{-\frac{\eta+1}{2}} ((j^*)^\eta - j^\eta) \\
&= \frac{\eta+1}{\eta j^*} \left( 1 - \left( \frac{j}{j^*} \right)^\eta \right). \tag{19}
\end{aligned}$$

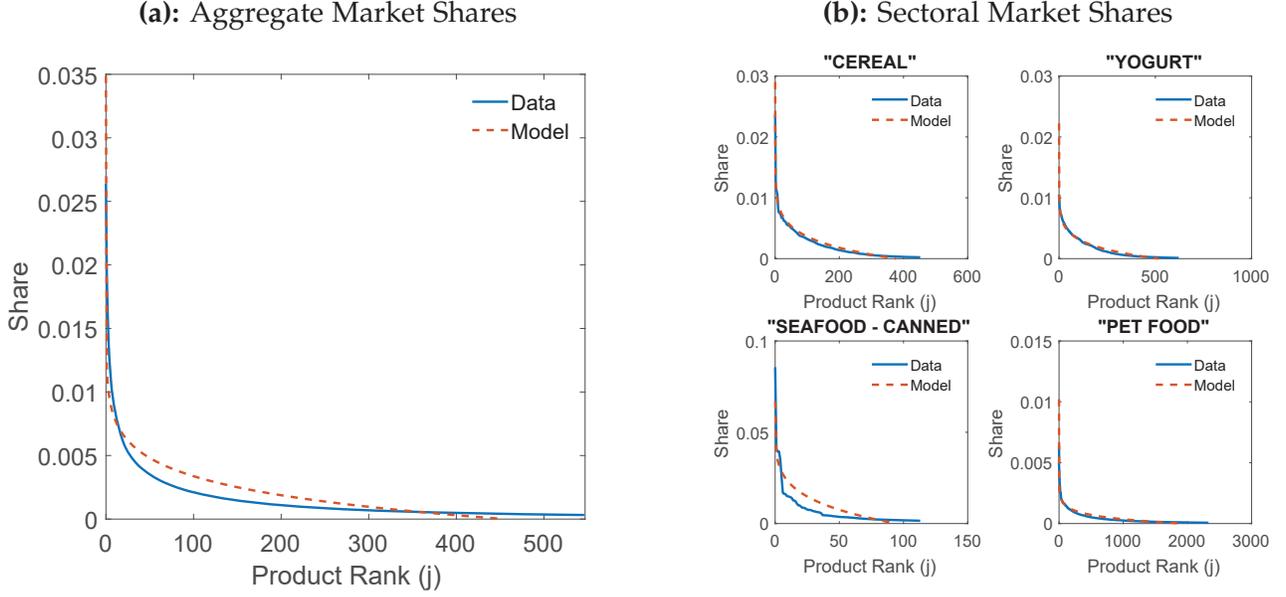
Using equation (19), we immediately obtain the Aggregate Herfindahl:

$$\begin{aligned}
\mathcal{H}^{\text{Agg}} &= \int_{j=0}^{j^*} s_j^2 dj = \left( \frac{\eta+1}{\eta j^*} \right)^2 \int_{j=0}^{j^*} \left( 1 - \left( \frac{j}{j^*} \right)^\eta \right)^2 dj \\
&= \left( \frac{\eta+1}{\eta j^*} \right)^2 j^* \left( 1 - \frac{2}{\eta+1} + \frac{1}{2\eta+1} \right) \\
&= \frac{2(\eta+1)}{(2\eta+1)} \frac{1}{j^*} \\
&= \frac{2(\eta+1)}{(2\eta+1)} \left( \frac{1}{2\tilde{N}|\Omega|} \right)^{\frac{1}{2}}, \tag{20}
\end{aligned}$$

where we define  $\tilde{N} = N\alpha/(1-\alpha)$ . Aggregate concentration declines monotonically with  $\tilde{N}$ . For fixed  $\theta$  and  $\sigma$ , aggregate concentration declines monotonically with the number of consumed products. And for fixed  $|\Omega|$ , concentration declines monotonically with  $\eta$ . Importantly, changes in  $|\Omega|$  and  $\eta$  move the Household Herfindahl and Aggregate Herfindahl in the same direction. As we discuss in the next subsection, this imposes strong restrictions on the set of possible forces which can explain the opposite empirical trends for  $\mathcal{H}^{\text{HH}}$  and  $\mathcal{H}^{\text{Agg}}$  and implies an important role for increases in  $\tilde{N}$ .

How well do these model-based relationships fit aggregate sales distributions in the data? To assess this, we start by measuring  $|\Omega|$  directly in the data and then solve for the two remaining free parameters,  $\eta$  and  $\tilde{N}$ , to match  $\mathcal{H}^{\text{HH}}$  and  $\mathcal{H}^{\text{Agg}}$  in equations (12) and (20). Figure 8 then plots the market share distribution across products implied by our model in equation (19) (the red dashed line) against the actual market share distribution in the data (the solid blue line). We do this for total spending in Figure 8a as well as separately for a number of product categories in Figure 8b. In several categories such as cereal and yogurt, the model fits extremely well, while it is notably less successful in others such as greeting cards or canned seafood. Overall, however, we consider these good fits as validating our use of the model, particularly given the distributions are fully determined by only

Figure 8: Model Fit on Aggregate Category Data



three parameters and reflect parametric assumptions and functional forms chosen largely for analytical convenience.

#### 4.5 Understanding the Empirical Trends

We now confront our model with concentration measures and other moments from the data to infer which structural forces led to the rise in niche consumption. Collecting previous results, our model implies that:

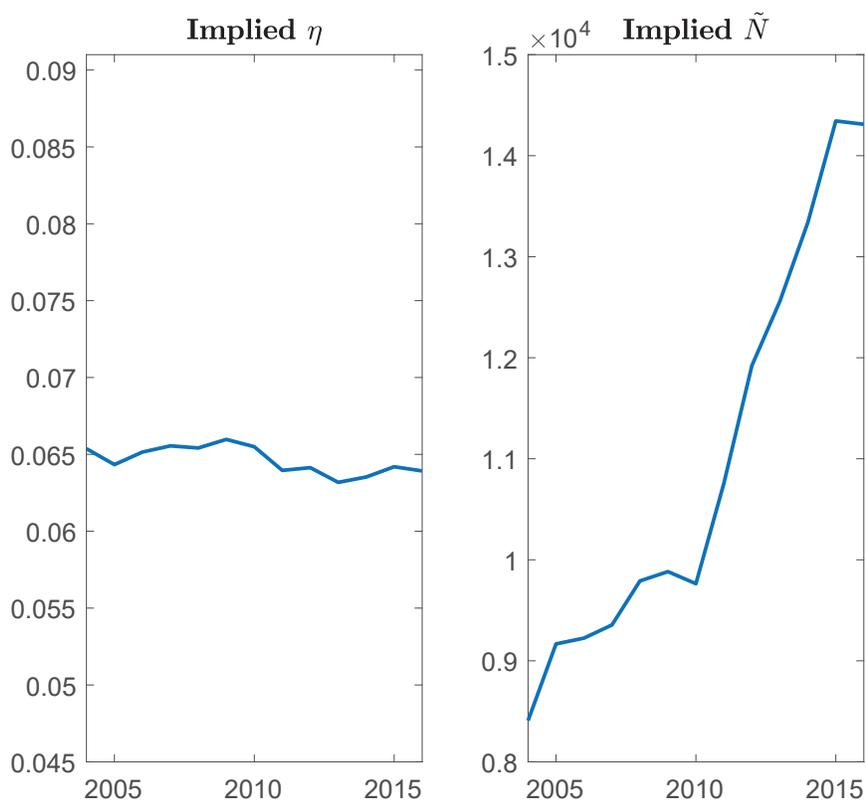
$$\mathcal{H}^{\text{HH}} = \frac{(\eta + 1)^2}{4\eta} \frac{1}{|\Omega|}$$

$$\mathcal{H}^{\text{Agg}} = \frac{2(\eta + 1)}{(2\eta + 1)} \left( \frac{1}{2\tilde{N}|\Omega|} \right)^{\frac{1}{2}}.$$

Since  $\mathcal{H}^{\text{HH}}$ ,  $\mathcal{H}^{\text{Agg}}$ , and  $|\Omega|$  are directly observable in the data, this produces a system of two equations that can be solved to determine  $\eta$  and  $\tilde{N}$  for each year. Figure 9 shows the time-series for  $\tilde{N}$  and  $\eta$  necessary to hit these observables in each year. Through the lens of the model, given the observed path for  $|\Omega|$ , matching the concentration trends requires nearly constant values for  $\eta$  and a strong upward trend in  $\tilde{N}$ . From 2004-2016,  $\eta$  falls by 2 percent while  $\tilde{N}$  rises by 70 percent.

Inspecting the equations for the two concentration measures, it is clear that increases in  $\tilde{N}$  push the Aggregate Herfindahl down relative to the Household Herfindahl. Is it obvious that  $\tilde{N}$  must *necessarily* rise to explain opposite trends in the two concentration measures? In Appendix D, we demonstrate that the answer is no. Theoretically, there are combinations of  $|\Omega|$  and  $\eta$  that deliver increases in  $\mathcal{H}^{\text{HH}}$  and decreases in  $\mathcal{H}^{\text{Agg}}$ , even without a change in  $\tilde{N}$ . These combinations, however, are grossly at odds

Figure 9: Modeling the Rise of Niche Consumption



with multiple moments of the data. Multiple factors can cause the observed changes in  $|\Omega|$ , but all explanations of the data require an increase in  $\tilde{N}$ , so we focus our analysis on this force.

Mechanically, increases in  $\tilde{N}$  can arise from increases in  $\alpha$  or  $N$ . Changes in  $\alpha$  are straightforward to interpret, since  $\alpha$  is simply an exogenous parameter governing preference heterogeneity. While increases in  $\alpha$  could easily rationalize the data, our empirical results show that the rise of niche consumption occurs pervasively across all of our narrowly-defined demographic groups. The within-group trends are far more important than across-group trends in generating our aggregate results. While this does not rule out increases in  $\alpha$  as a driving force, it seems unlikely that fundamental preferences within narrow groups have become dramatically more heterogeneous over a twelve-year period. Based on this logic, we hold  $\alpha$  fixed and explore the effects of increases in  $N$  in the model, holding constant all other parameters, including the price-adjusted taste distribution.

#### 4.6 Implications of An Increase in the Number of Products

We set all parameter values to match key empirical moments in 2004 and increase  $N$  by 70 percent to generate the increase in  $\tilde{N}$  backed out in Figure 9.<sup>31</sup> We then calculate the change in household

<sup>31</sup>The exact initial parameters are not important for our qualitative conclusions, but we set  $\alpha = 0.36$ ,  $\epsilon = 2$ ,  $E = 35$ ,  $\sigma = 3.76$ ,  $b = 1$ ,  $F = 0.14$  and  $\theta = 5.9$ .  $E$  is set to match average household category expenditures,  $\theta$  and  $\sigma$  are set to match

welfare, expressed as the percentage change in expenditures on the initial set of goods that would bring the same change in household utility as that delivered by the increase in  $N$ . We find that a 70 percent increase from  $N_{2004}$  to  $N_{2016}$  generates total welfare gains of 9.5 percent, or about 0.8 percent per year. That is:

$$U_{2016} = \frac{E}{P_{N_{2016}}} - F \times (|\Omega_{N_{2016}}|)^\epsilon = 1.095 \times \frac{E}{P_{N_{2004}}} - F \times (|\Omega_{N_{2004}}|)^\epsilon, \quad (21)$$

where we change  $N$  and calculate the endogenous change in  $P$  and  $\Omega$ , but hold fixed all other parameters.

These welfare gains of 9.5 percent arise from three sources. First, as seen clearly in the third term of the ideal price index in equation (9), an increase in  $N$  for a given  $|\Omega|$  leads to gains from selection. With more choices, households consume those products better suited to their particular tastes. Second, as seen clearly in equation (10), increases in  $N$  lead to increases in  $|\Omega|$ . As seen in the second term of the ideal price index, this leads to welfare gains from the love-for-variety effect on welfare, even if selection effects  $|\Omega|/N$  are held constant. Third, increases in  $|\Omega|$  lead to increases in the fixed costs paid by households, which reduces welfare. Selection effects account for the bulk of the gains (8.5 percent), with love-for-variety effects improving welfare by 1.8 percent and the increase in fixed costs causing welfare losses of approximately 1.0 percent (these do not sum perfectly to 9.5 due to non-linearity).

An increase in  $N$  (or  $\alpha$ ) is a necessary part of any explanation that quantitatively matches the simultaneous rise in  $\mathcal{H}^{\text{HH}}$  and decline in  $\mathcal{H}^{\text{Agg}}$ .<sup>32</sup> Other parameters must change as well, however, in order to fully fit the concentration trends. After all, as just discussed, an increase in  $N$  on its own causes  $|\Omega|$  to rise and generates a (small) decline rather than an increase in  $\mathcal{H}^{\text{HH}}$ .

Equation (10) shows that if  $N$  increases, declines in  $|\Omega|$  must reflect declines in measured real expenditures  $\left(E (1 + (1 - \sigma) / \theta)^{\frac{1}{1-\sigma}} b\right)$ , increases in effective costs per number of products consumed ( $F$  or  $\epsilon$ ), or declines in an "effective curvature" of utility term  $\left(\frac{1}{1-\sigma} - \frac{1}{\theta}\right)$ .<sup>33</sup> Expenditures, however, increase in the data, and while changes in either  $\sigma$  or  $\theta$  could change the curvature term, they would have to change in a very particular way so as to match the decline in  $|\Omega|$  without leading to changes in  $\eta$ . We therefore find it most plausible that the decline in  $|\Omega|$  reflected an increase in  $F$  or  $\epsilon$ .<sup>34</sup>

How does our inference about welfare change when we consider joint increases in  $N$  and in  $F$  (or  $\epsilon$ ) that are chosen to match both the rise of  $\mathcal{H}^{\text{HH}}$  and the declines in  $\mathcal{H}^{\text{Agg}}$  and  $|\Omega|$ ? Repeating

---

the initial  $\eta$  in the data given observed  $\Omega$  subject to generating a ratio of aggregate sales to aggregate costs of 1.2 percent. Given  $b$  and  $\epsilon$ , we choose  $F$  to match  $\Omega$ .

<sup>32</sup> $N$  impacts the extent of selection effects while  $\alpha$  does not, so if one considers that growth in  $\tilde{N}$  is driven in part by growth in  $\alpha$ , the welfare gains will be smaller.

<sup>33</sup>Real measured expenditures equal  $E$  divided by the expression labeled "Average Price" in equation (9).

<sup>34</sup>While some technological advances such as the rise of the internet or better advertising technology might be expected to lower variety costs, it is also likely that increases in the number of available varieties  $N$  make it more costly to sort through and identify the particular products a household wants to purchase. An increase in  $F$  or  $\epsilon$  can be interpreted as a simple proxy for these latter forces when it is accompanied by the increase in  $N$ .

our comparative statics exercise when we vary both  $N$  and  $F$ , we find that welfare rises by about 7.5 percent instead of 9.5 percent. This decomposes into selection gains above 10 percent, love-for-variety losses of about 2 percent, and an additional loss of 1 percent from the higher fixed costs. Results are vary similar when we vary  $N$  and  $\epsilon$  instead of  $N$  and  $F$ .

Finally, we note that if one simply viewed our data through the lens of a representative household model with CES preferences, it might be natural to only consider this love-of-variety loss and misleadingly conclude that welfare *declined* from 2004-2016, since the typical household consumed fewer varieties in 2016 than in 2004. Heterogeneity in product consumption across households is crucial for capturing the divergent concentration trends in our data. Representative agent models abstract from this heterogeneity, and our results show that this can potentially lead to misleading conclusions about the welfare effects arising from changes in the number of products households consume.

#### 4.7 Elasticities of Demand, Markups, and Aggregate Profits

Concentration measures like the Herfindahl are often used as proxies for market power. What implications does the divergence in Aggregate and Household Herfindahls have for markups and aggregate profits in our model? In typical CES environments, the elasticity of demand and markups are fully determined by the exogenous elasticity of substitution  $\sigma$ . By contrast, we show in this section that the elasticity of demand in our model depends both on this standard “intensive margin” force as well as on an endogenous “extensive margin” force that arises from the possibility for products to gain new customers (or lose existing ones). Since these forces are of different importance for products with different aggregate market shares, the model generates heterogeneous markups.

How does an increase in  $N$  influence markups? On the one hand, competition from new products pushes existing consumed products closer to the margin where they are no longer consumed, increasing their elasticity of demand and decreasing their markups. On the other hand, selection effects mean those new products are on average more desirable and have higher markups than those in the bundle prior to the change in  $N$ . We show that in the aggregate, these opposing forces exactly cancel and the ratio of total revenues to total costs remains unchanged. Our model therefore shows how the economy can exhibit large changes in aggregate and household concentration without any change in aggregate market power.<sup>35</sup>

To solve for the price elasticity of aggregate demand for product  $j$ , we start by expressing its total sales as the integral of each household’s spending on  $j$ , taken over all households:

$$s_j = \frac{1}{N} \int_0^{x_j^*} s_{i,x,j} dx, \quad (22)$$

---

<sup>35</sup>We note that if each consumer’s taste for each good remains fixed, and markups change for any reason, this would result in a change in the price-adjusted taste distribution and would affect the expressions above that were derived assuming that price-adjusted tastes were distributed a la Pareto. We explore this in more detail in Appendix E, but here note that in order to preserve the Pareto distribution of price-adjusted tastes in the face of increases in  $N$  and endogenous markups, the changes in tastes that we additionally require are relatively minor.

where we use the notation  $s_{i_x,j}$  to denote the spending share of a household with taste draw on product  $j$  equal to  $x$ . Since  $j$  will only be purchased by those households with a sufficiently high idiosyncratic taste for it, we need only integrate from households drawing  $x_{i,j} = 0$  to the marginal household that draws  $x_{i,j} = x_j^*$ .

We take the partial derivative of  $s_j$  in equation (22) with respect to  $p_j$  to get:

$$\frac{\partial s_j}{\partial p_j} = \frac{1}{N} \left( \int_0^{x_j^*} \frac{\partial s_{i_x,j}}{\partial p_j} dx + s_{i_{x_j^*},j} \frac{\partial x_j^*}{\partial p_j} \right), \quad (23)$$

where the right hand side of equation (23) follows from Leibniz's rule. The first term can be solved using equation (15) as:

$$\frac{\partial s_{i_x,j}}{\partial p_j} = \frac{\partial P^{\sigma-1} p_j^{1-\sigma} \gamma_{i,j}^{\sigma-1}}{\partial p_j} = (1-\sigma) \frac{s_{i_x,j}}{p_j}, \quad (24)$$

where we take the aggregate price index  $P$  as fixed. Moving on to the second term, we can evaluate equation (15) at the marginal household with taste  $x_j^*$  to get:

$$s_{i_{x_j^*},j} \frac{\partial x_j^*}{\partial p_j} = \frac{\eta+1}{2} N^{\frac{\eta-1}{2}} |\Omega|^{-\frac{\eta+1}{2}} (R(r^*))^{\frac{\eta-1}{2}} \frac{\partial x_j^*}{\partial p_j}. \quad (25)$$

Substituting equations (24) and (25) back into equation (23), we get:

$$\begin{aligned} \frac{\partial s_j}{\partial p_j} &= (1-\sigma) \frac{1}{p_j} \frac{1}{N} \int_0^{x_j^*} s_{i_x,j} dx + \frac{1}{N} s_{i_{x_j^*},j} \frac{\partial x_j^*}{\partial p_j} \\ &= (1-\sigma) \frac{s_j}{p_j} + \frac{1}{N} \frac{\eta+1}{2} N^{\frac{\eta-1}{2}} |\Omega|^{-\frac{\eta+1}{2}} (R(r^*))^{\frac{\eta-1}{2}} \frac{\partial x_j^*}{\partial p_j} \\ &= (1-\sigma) \frac{s_j}{p_j} + \frac{1}{N} \frac{\eta+1}{2} N^{\frac{\eta-1}{2}} |\Omega|^{-\frac{\eta+1}{2}} \left( \frac{1}{2} N^{-2} \alpha^{-1} (1-\alpha) (j^*)^2 \right)^{\frac{\eta-1}{2}} \frac{\partial x_j^*}{\partial p_j} \\ &= (1-\sigma) \frac{s_j}{p_j} + \frac{\eta+1}{2N|\Omega|} \frac{\partial x_j^*}{\partial p_j}. \end{aligned} \quad (26)$$

We cannot directly consider a price change, required to exactly solve for  $\partial x_j^* / \partial p_j$ , because the full distribution of price-adjusted tastes in our model is set exogenously and must be maintained as Pareto. Instead, to *approximate* the change in mass of households consuming a product when that product's price increases, we start with the relationship:<sup>36</sup>

$$\frac{1}{2} \left( \frac{r_{i,j}}{N} \right)^2 \frac{1}{\alpha(1-\alpha)} = R((1-\alpha)j + \alpha x_{i,j}) = 1 - G\left(\frac{\gamma_{i,j}}{p_j}\right) = b^\theta \gamma_{i,j}^{-\theta} p_j^\theta, \quad (27)$$

<sup>36</sup>In Appendix E we use numerical simulations which do not require any distributional assumption on tastes to verify that the approximation that follows is highly accurate.

and differentiate to yield:

$$\frac{r_{i,j}}{1-\alpha} \frac{1}{N^2} \frac{\partial x_{i,j}}{\partial p_j} = \theta b^\theta \gamma_{i,j}^{-\theta} p_j^{\theta-1}, \quad (28)$$

where we've substituted  $\partial r_{i,j}/\partial x_{i,j} = \alpha$ . We then evaluate equation (28) at  $r_{i,j} = r^*$  and  $\tilde{\gamma}_{i,j} = \tilde{\gamma}^*$  using equation (16), and add a minus sign to reflect the fact that increase in the price of good  $j$  should reduce the set of households purchasing that good, to get:

$$\frac{dx_j^*}{dp_j} = -\frac{\theta}{j^*} |\Omega| N \frac{1}{p_j}. \quad (29)$$

Inserting this into equation (26), we have:

$$\begin{aligned} \frac{\partial s_j}{\partial p_j} &= (1-\sigma) \frac{s_j}{p_j} - \frac{\eta+1}{2N|\Omega|} \frac{\theta}{j^*} |\Omega| N \frac{1}{p_j} \\ &= (1-\sigma) \frac{s_j}{p_j} - \frac{\eta\theta}{2\left(1-\left(\frac{j}{j^*}\right)^\eta\right)} \frac{s_j}{p_j}. \end{aligned} \quad (30)$$

Equation (30) implies that product  $j$ 's price elasticity of demand  $\varepsilon_j$  can be written as:

$$\begin{aligned} \varepsilon_j &= 1 - \frac{\partial s_j}{\partial p_j} \frac{p_j}{s_j} \\ &= \underbrace{\sigma}_{\text{Intensive Margin}} + \underbrace{\left(1 - \left(\frac{j}{j^*}\right)^\eta\right)^{-1} [\theta/2 - (\sigma - 1)]}_{\text{Extensive Margin}} > \sigma. \end{aligned} \quad (31)$$

In addition to the standard intensive margin term  $\sigma$ , there is a strictly positive contribution from the extensive margin, since lowering the price of a product can induce new households to start consuming the product. Low  $j$  or "mass-market" products are consumed by many households, so the intensive margin is relatively more important for them. High  $j$  or "niche" products are consumed by few households, so the extensive margin is relatively more important. As a result, the elasticity of demand increases as market share falls.<sup>37</sup> As  $j \rightarrow j^*$  and a product approaches the point where it is dropped from the aggregate consumption bundle, the elasticity approaches infinity, i.e.  $\varepsilon \rightarrow \infty$ . The markup  $\mu_j$  can then be written (in gross terms) as:

$$\mu_j = \frac{\varepsilon_j}{\varepsilon_j - 1} = \frac{\sigma + \frac{\theta(\eta+1)}{2j^*s_j}}{\sigma + \frac{\theta(\eta+1)}{2j^*s_j} - 1}, \quad (32)$$

and ranges from a high of  $(1 + 2/\theta)$  for the largest good  $j = 0$  to a low of 1 for  $j = j^*$ .

<sup>37</sup>Interestingly, for good  $j = 0$ , which has the largest aggregate demand, the positive impact of  $\sigma$  on the elasticity coming through the intensive margin exactly cancels with the negative impact of  $\sigma$  coming from the extensive margin, leaving a total elasticity of  $(\theta/2 + 1)$ . This result echos a closely related point in [Chaney \(2008\)](#), where the impact of the equivalent parameter for the elasticity of trade flows to trade costs also fully cancels when combining the intensive and extensive margin effects.

What are the implications of the rise in niche consumption for the distribution of products and for market power? The “aggregate markup” is equal to the ratio of aggregate sales to aggregate costs. Using equations (19) and (31), it can be written as:

$$\mu^{Agg} = \frac{\int_0^{j^*} s_j dj}{\int_0^{j^*} s_j \frac{\epsilon_j - 1}{\epsilon_j} dj} = \left[ \frac{\theta + (\sigma - 1)^2}{\sigma^2} - \frac{1}{2} \frac{\eta \theta^2}{\sigma^2} \left( \frac{\eta + 1}{2 + \theta} \right) \times {}_2F_1 \left( 1, \frac{1}{\eta}; 1 + \frac{1}{\eta}; \frac{2\sigma}{2 + \theta} \right) \right]^{-1}, \quad (33)$$

where  ${}_2F_1(\cdot)$  is the hypergeometric function.<sup>38</sup> Importantly, while this aggregate profit share is a relatively complicated function of  $\sigma$  and  $\theta$ , it is *not* a function of  $\tilde{N}$ ,  $F$ , or  $\epsilon$ , the key model forces we argued drove the rise in niche consumption. Changes in  $\tilde{N}$ ,  $F$ , and  $\epsilon$  may have implications for the distribution of markups across products in the economy, since they impact  $j^*$ , but these changes leave the “aggregate markup” exactly constant, since  $j^*$  drops out of equation (33). Given the common association of changes in concentration with changes in market power, we find this to be a surprising and useful result.

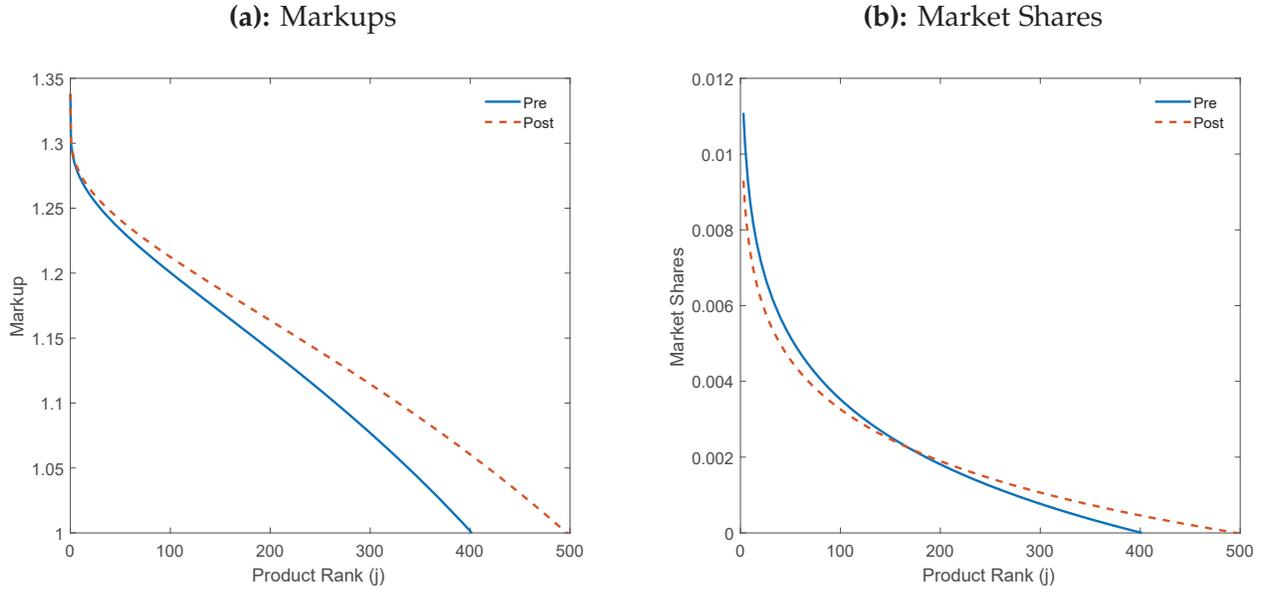
More intuitively, this result arises from two opposing forces which exactly cancel when  $j^*$  changes. On the one hand, the  $j$ th good in an economy with a low  $j^*$  is closer to being the marginal consumed good and will therefore have a lower markup than the  $j$ th good in an economy with a high  $j^*$ . This can be seen in equation (31), which shows that the elasticity of demand is strictly increasing in the ratio of  $j$  to  $j^*$ . The solid blue line in Figure 10a plots markups as a function of product rank  $j$  using the same calibrated values meant to represent 2004 in the exercise in Section 4.6, and the red dashed line plots this relationship after increasing both  $\tilde{N}$  and  $F$  to fit the rise of niche consumption seen by 2016. The entire distribution of markups on consumed products shifts upward, with the top product ( $j = 0$ ) as the only exception. On the other hand, in the new economy with more product options, the top high-markup products account for a smaller share of aggregate spending, as shown in Figure 10b. Both forces are individually quite powerful, but in our model they exactly offset such that there is no change in the aggregate profit share.<sup>39</sup>

While Figure 10 demonstrates the impact of this change in  $\tilde{N}$  and  $F$  on the joint distribution of markups and market shares, additional structure is required if one wishes to comment on the implications of entry for a given product. For example, in response to changes in  $\alpha$  or  $F$ , it is natural to assume that a particular product  $j$ 's aggregate rank does not change, i.e.  $j^{post} = j^{pre}$ . A comparison of the lines in Figure 10 evaluated at a particular product rank  $j$ , therefore, would capture the change impacting that particular product. By contrast, a given product's  $j$  would likely change in response to a change in  $N$ , depending on whether the new expanded set of products are drawn from a better,

<sup>38</sup>The hypergeometric function is defined as follows:  ${}_2F_1(a, b; c; z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!}$ .  $(x)_n$  is the Pochhammer symbol, which equals  $\frac{(x+n-1)!}{(x-1)!}$  for all  $n > 0$  and equals 1 for  $n = 0$ .

<sup>39</sup>Putting Figures 10a and 10b together, the distribution of actual profits flattens across  $j$  values, as shown in Appendix Figure A13.

Figure 10: Effects of  $\tilde{N} \uparrow$  and  $F \uparrow$  on Markup and Market Share Distributions



worse, or identical distribution as the original smaller set of products. For example, if the new and old products have the same quality distribution, it is most natural to think of a given product's  $j$  as evolving according to:  $j^{post} = j^{pre} \times (N^{post} / N^{pre})$ . Since we cannot pin down the change in markup or market share of any particular incumbent product in our model without making further assumptions on the nature of entry, and given we do not have data on product markups suitable to empirically measure such changes, we leave this to future work to explore.

## 5 Conclusions

This paper empirically documents a rise in what we call "niche" consumption. Households are increasingly concentrating their spending. This pattern, however, does not appear to be driven by the emergence of superstar products. Rather, households are increasingly buying different goods from one another. The increase in segmentation seen in many other walks of modern life also applies to consumption: our grocery baskets look less and less similar. As a result, aggregate spending has become less concentrated.

We develop a new model of product demand in order to explore the drivers and implications of the rise in niche consumption. In our model, households choose how many products to consume, spend different amounts on each good, and differ from other households in their choice of which products to buy. The model delivers simple analytical expressions for household and aggregate concentration indices, and these closed form solutions allow us to match the model to data and infer the drivers of our empirical findings. Increases in product availability played a critical role in the divergent concentration trends, and led to welfare gains from households being able to consume a subset of

products that better satisfied their tastes. This welfare effect would be difficult to find using standard statistics such as measured price indices. Finally, our model delivers endogenous and heterogeneous markups. Matching the trends in household and aggregate concentration carries implications for the distribution of markups, but does not imply changes in aggregate market power.

Our model highlights the importance of greater product choice but treats the set of available products as an exogenous parameter. We suspect the nature of product introduction and development, however, reflects recent progress in supply chain integration, big-data marketing research, targeted advertising, and the growing importance of online sales. Unpacking the product innovation process and relating it to these important trends is a fruitful avenue for future research on consumption behavior.

## References

- AGUADO, J. M., C. FEIJOO, AND I. MARTINEZ (2015): *Emerging Perspectives on the Mobile Content Evolution*. IGI Global.
- ALLCOTT, H., R. DIAMOND, J.-P. DUBE, J. HANDBURY, I. RAHKOVSKY, AND M. SCHNELL (2017): "Food Deserts and the Causes of Nutritional Inequality," Working Paper 24094, National Bureau of Economic Research.
- ALWIN, D., AND P. TUFIS (2015): "The Changing Dynamics of Class and Culture in American Politics: a Test of the Polarization Hypothesis," *The Annals of the American Academy of Political and Social Science*, 663.
- ANDERSON, C. (2006): *The Long Tail: Why the Future of Business Is Selling Less of More*. Hyperion.
- ANDERSON, E., S. REBELO, AND A. WONG (2018): "Markups Across Space and Time," Working Paper 24434, National Bureau of Economic Research.
- ARGENTE, D., AND M. LEE (2017): "Cost of Living Inequality During the Great Recession," mimeo.
- ARGENTE, D., M. LEE, AND S. MOREIRA (2018a): "How do firms grow? The life cycle of products matters," .
- (2018b): "Innovation and product reallocation in the great recession," *Journal of Monetary Economics*, 93, 1–20.
- ARKOLAKIS, C., S. DEMIDOVA, P. J. KLENOW, AND A. RODRIGUEZ-CLARE (2008): "Endogenous variety and the gains from trade," *American Economic Review*, 98(2), 444–50.
- ATKESON, A., AND A. BURSTEIN (2008): "Pricing-to-market, trade costs, and international relative prices," *American Economic Review*, 98(5), 1998–2031.
- ATKIN, D., B. FABER, AND M. GONZALEZ-NAVARRO (2018): "Retail globalization and household welfare: Evidence from Mexico," *Journal of Political Economy*, 126(1), 1–73.
- AUTOR, D., D. DORN, L. F. KATZ, C. PATTERSON, AND J. V. REENEN (2017): "The Fall of the Labor Share and the Rise of Superstar Firms," Working Paper 23396, National Bureau of Economic Research.
- AUTOR, D. H., L. F. KATZ, AND M. S. KEARNEY (2006): "The Polarization of the U.S. Labor Market," *The American Economic Review*, 96(2).
- BERGER, D., AND J. VAVRA (2015): "Consumption Dynamics During Recessions," *Econometrica*, 83(1).
- BERGER, D., AND J. VAVRA (2019): "Shocks vs. Responsiveness: What Drives Time-Varying Dispersion?," *Journal of Political Economy*.
- BERRY, S., A. EIZENBERG, AND J. WALDFOGEL (2016): "Optimal product variety in radio markets," *The RAND Journal of Economics*, 47(3), 463–497.
- BERRY, S. T., M. GAYNOR, AND F. S. MORTON (2019): "Do Increasing Markups Matter? Lessons from Empirical Industrial Organization," Working Paper 26007, National Bureau of Economic Research.
- BERTRAND, M., AND E. KAMENICA (2018): "Coming apart? Lives of the Rich and the Poor Over Time," Discussion paper.
- BORNSTEIN, G. (2018): "Entry and Profits in an Aging Economy: The Role of Consumer Inertia," .

- BRODA, C., AND D. E. WEINSTEIN (2006): "Globalization and the Gains from Variety," *The Quarterly Journal of Economics*, 121(2), 541–585.
- BRODA, C., AND D. W. WEINSTEIN (2004): "Variety growth and world welfare," *American Economic Review*, 94(2), 139–144.
- CHANEY, T. (2008): "Distorted gravity: the intensive and extensive margins of international trade," *American Economic Review*, 98(4), 1707–21.
- COIBION, O., Y. GORODNICHENKO, AND D. KOUSTAS (2017): "Consumption Inequality and the Frequency of Purchases," Discussion paper.
- DAVIS, D. R., J. I. DINGEL, J. MONRAS, AND E. MORALES (2017): "How Segregated is Urban Consumption?," .
- DE LOECKER, J., AND J. EECKHOUT (2017): "The Rise of Market Power and the Macroeconomic Implications," Working Paper 23687, National Bureau of Economic Research.
- DUBE, J.-P., G. HITSCH, AND P. ROSSI (2018): "Income and Wealth Effects on Private-Label Demand: Evidence from the Great Recession," *Marketing Science*.
- EDMOND, C., V. MIDRIGAN, AND D. Y. XU (2018): "How costly are markups?," Discussion paper, National Bureau of Economic Research.
- FEENSTRA, R. C. (1994): "New product varieties and the measurement of international prices," *The American Economic Review*, pp. 157–177.
- FEENSTRA, R. C., AND D. E. WEINSTEIN (2017): "Globalization, markups, and US welfare," *Journal of Political Economy*, 125(4), 1040–1074.
- FURMAN, J., AND P. ORSZAG (2015): "A firm-level perspective on the role of rents in the rise in inequality," *Presentation at AIAA Just Society Centennial Event in Honor of Joseph Stiglitz Columbia University*, 16.
- GENTZKOW, M., J. M. SHAPIRO, AND M. TADDY (2017): "Measuring Polarization in High-Dimensional Data: Method and Application to Congressional Speech," Working Paper 22423, National Bureau of Economic Research.
- GOPINATH, G., AND O. ITSKHOKI (2010): "Frequency of price adjustment and pass-through," *The Quarterly Journal of Economics*, 125(2), 675–727.
- HANDBURY, J., AND D. E. WEINSTEIN (2014): "Goods prices and availability in cities," *The Review of Economic Studies*, 82(1), 258–296.
- HANSEN, K., AND V. SINGH (2015): "Choice Concentration," Discussion paper.
- HOTTMAN, C. J., S. J. REDDING, AND D. E. WEINSTEIN (2016): "Quantifying the sources of firm heterogeneity," *The Quarterly Journal of Economics*, 131(3), 1291–1364.
- JARAVEL, X. (2019): "The Unequal Gains from Product Innovations: Evidence from the US Retail Sector," *Quarterly Journal of Economics*, 134(2).
- KAPLAN, G., AND G. MENZIO (2015): "The morphology of price dispersion," *International Economic Review*, 56(4), 1165–1206.
- KIMBALL, M. S. (1995): "The Quantitative Analytics of the Basic Neomonetarist Model," *Journal of Money, Credit and Banking*, 27(4), 1241–1277.

- KROFT, K., J.-W. P. LALIBERTÉ, R. LEAL-VIZCAÍNO, AND M. J. NOTOWIDIGDO (2017): “Quantifying the Welfare Gains of Variety: A Sufficient Statistics Approach,” Discussion paper, Mimeo, available at [http://korykroft.com/wordpress/KKLN variety july2017. pdf](http://korykroft.com/wordpress/KKLN%20variety%20july2017.pdf).
- LEVIN, A., AND T. YUN (2008): “Reconsidering the microeconomic foundations of price-setting behavior,” *Unpublished Manuscript, Board of Governors of the Federal Reserve System*.
- LI, N. (2018): “An Engel curve for variety,” *Working Paper*.
- MELITZ, M. J., AND G. I. OTTAVIANO (2008): “Market size, trade, and productivity,” *The review of economic studies*, 75(1), 295–316.
- MICHELACCI, C., L. PACIELLO, AND A. POZZI (2019): “The Extensive Margin of Aggregate Consumption Demand,” *Mimeo*.
- NEIMAN, B., AND J. VAVRA (2018): “The Rise of Household Spending Concentration,” *Mimeo*.
- PETRIN, A. (2002): “Quantifying the benefits of new products: The case of the minivan,” *Journal of political Economy*, 110(4), 705–729.
- PEW RESEARCH CENTER (2014): “Political Polarization in the American Public,” Discussion paper.
- PIKETTY, T., E. SAEZ, AND G. ZUCMAN (2016): “Distributional National Accounts: Methods and Estimates for the United States,” Working Paper 22945, National Bureau of Economic Research.
- REDDING, S. J., AND D. E. WEINSTEIN (2016): “Measuring aggregate price indexes with demand shocks: Theory and evidence for CES Preferences,” Discussion paper, National Bureau of Economic Research.
- ROSSI-HANSBERG, E., P.-D. SARTE, AND N. TRACHTER (2018): “Diverging trends in national and local concentration,” Discussion paper, National Bureau of Economic Research.
- STROEBEL, J., AND J. VAVRA (2019): “House Prices, Local Demand and Retail Prices,” *Journal of Political Economy*, 127(3).
- SYVERSON, C. (2018): “Remarks at ‘Changing Market Structure and Implications for Monetary Policy’,” *Remarks, 2018 Jackson Hole Symposium*.