# The Rise of Niche Consumption*

Brent Neiman
University of Chicago

Joseph Vavra
University of Chicago

June 2019

**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 differ in terms of which products they focus their spending on. 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 force underlying these divergent trends. When more products are available, households can select a subset better matched to their particular tastes. This results in substantial welfare gains that are not reflected in standard government statistics. The 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 set of customers. Surprisingly, however, our model can match the observed trends in household and aggregate concentration without any resulting change in aggregate market power.

*We thank David Argente and Anhua Chen for providing exceptional research assistance and Rodrigo Adao, Jonathan Dingel, J.P. Dubé, Austan Goolsbee, Pete Klenow, Thomas Mertens, 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."

# 1 Introduction

Recent decades have played host to important changes in how goods are designed, produced, marketed, and delivered. Studies exploring the ramifications of these changes for macroeconomic aggregates and welfare typically focus on producer-, retailer-, or industry-level data. In this paper, we use household-level data on retail purchases to provide a new perspective on the rapidly changing retail environment. We show that these changes manifest even in individual shopping carts, develop a novel model capable of matching this micro evidence, and demonstrate that it delivers implications that could not be gleaned from aggregate sales 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 and decline in aggregate concentration be reconciled? Each household increasingly concentrates spending on its own preferred products, but households also increasing 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 this 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.[1] Finally, while there is variation in the magnitude, niche consumption grows in almost all product categories.

[1] We also show that our results are not driven by trends in the share of generic products (Dube, Hitsch, and Rossi (2018)) or by trends in the frequency of shopping and bulk purchasing (Olivier Coibion (2017)).

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. We also show that both household and aggregate concentration trends are attenuated significantly 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 demand we assume that there is a continuum of households and that each household has an identical distribution of tastes for all products. Although all households have identical taste *distributions*, we allow for cross-household heterogeneity in product demand by letting the ranking of particular products within this distribution vary across households. For example, we assume that tastes for all households decline identically from their favorite product to their second favorite, and so on. However, the the actual identity of two different household's favorite products can potentially differ. To capture this, we introduce a "rank" function, which maps each product to a relative position in each households' tastes. The rank function is a weighted average of a common aggregate component and a random household-

specific component. When the rank function fully weights the aggregate component, the economy converges to a representative household economy with full agreement on which products to consume and identical household and aggregate spending concentration. Conversely, when the rank function fully weights the idiosyncratic household-specific component, there may be uniform aggregate spending across products and low aggregate concentration even if individual household spending is very 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 some of its basic predictions are broadly in line with relevant empirical moments. For example, the relationship of household Herfindahls and the number of consumed varieties in the model resembles that in the data. 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.[2]

Why does this increase in product availability matter? In our model, an increase in the total number of varieties available leads to a small 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. In our model, 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 individual households consume. In a counterfactual exercise where we hold all else equal and raise varieties by the 70% implied by our model, we estimate gains from selection that households value equivalently to the ability to increase consumption expenditures by 9.5 percent on a fixed set of products. This welfare gain from selection shows up in the ideal price index, but would 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

[2] 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.

both Herfindahl trends, rather than just their difference. We show how 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 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 will vary across the product distribution. For example, the product with the largest amount of sales in the economy has many existing customers. It's elasticity of demand will mostly reflect the price-sensitivity of those customers, leading the firm 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. It will set price close to marginal cost.

We can characterize each product's markup analytically as an increasing function of its percentile in the distribution of purchased products, with larger market-share products charging higher markups. Increases in the number of available products improves selection, which shifts the set of continuing products to higher percentiles in that distribution (noting that percentile 0 is the top-ranked aggregate product) and leaves them charging lower markups. By contrast, the newly consumed products, on average, occupy lower percentiles 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.

Our work touches on and draws connections between a number of important themes in recent research. Although we argue that 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 idiology (Pew Research Center (2014); Gentzkow, Shapiro, and Taddy (2017)), in jobs and income (Autor, Katz, and Kearney (2006); Piketty, Saez, and Zucman (2016)), and in the geography of where households consume (Davis, Dingel, Monras, and Morales (2017)).[3] Our findings indicate that, along with these other manifestations of fragmentation in modern life, even our grocery purchases increasingly differ from the national average.[4]

Several important papers document changes in top sales shares and industrial structure, including

[3]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.

[4]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.

David, Dorn, Katz, Patterson, and Van 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 selling 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. This contrasts with the more commonly used frameworks for generating variable markups, such as the linear demand environment in Melitz and Ottaviano (2008), the 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) of the simple linkage often made between concentration trends and market power.

Finally, 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). 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. 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).[5] Our model allows for gains from selection which are often stressed in this literature but are absent from typical macro models.

[5]See Kroft, Laliberté, Leal-Vizcaíno, and Notowidigdo (2017) for a sufficient statistic approach to the gains from new varieties.

## 2 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 B.[6] 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.

### 2.1 AC Nielsen Homescan Data

We use Homescan data from AC Nielsen to measure household-level shopping behavior.[7] 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.[8] Appendix Figure A12 shows that aggregate spending growth in our sample tracks government data on aggregate spending growth in comparable categories. We focus on the full sample of households for a number of reasons discussed in Appendix C.1, but this is relatively conservative since the magnitudes of our trends increase when we restrict to a balanced panel of households.[9]

In our baseline analysis, we define a product as a Universal Product Code (UPC). The Appendix demonstrates, however, that our results are robust to instead defining a product as a "brand". Nielsen assignes 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 they 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 cate-

[6] These data have also recently been analyzed in Stroebel and Vavra (2019), Jaravel (2019), and Allcott, Diamond, Dube, Handbury, Rahkovsky, and Schnell (2017).

[7] 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.

[8] 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.

[9] In the appendix we also discuss the relevance for our results of additional measurement-related issues, such as the (unimportant) role of online shopping.

gories and over time. Further, though 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.[10] 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 produce 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).

### 2.2 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 $E^t_{i,j,c}$ in year $t$ and associated expenditure share:

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

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:

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

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

$$\mathcal{H}^{\text{HH}}_{c,t} = \sum_i share_{i,c,t} \mathcal{H}^{\text{HH}}_{i,c,t}, \tag{3}$$

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

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

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

[10] See, for example, Dube, Hitsch, and Rossi (2018). Our robustness checks using the brand definition of product, therefore, exclude generics. Results using UPCs also remain qualitatively robust if we exclude generics.

category-specific Herfindahl in equation (3) across all categories:

$$\mathcal{H}_t^{\text{HH}} = \sum_c share_c \mathcal{H}_{c,t}^{\text{HH}}, \tag{5}$$

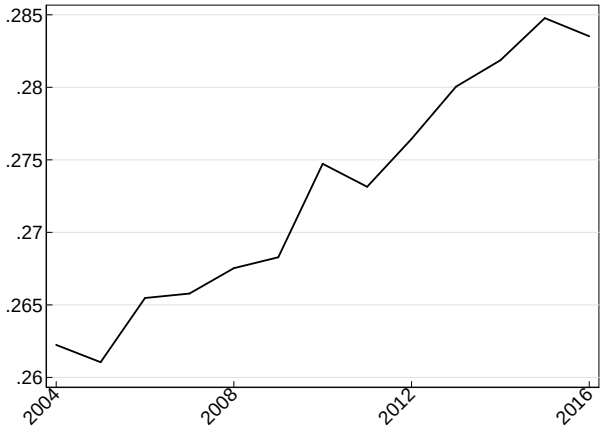
where $share_c$ is the average share of category $c$ in total spending across our entire sample.

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 occuring within sectors. Our results are qualitatively robust, though a bit noisier, when we allow compositional shifts across categories to influence our concentration measures.

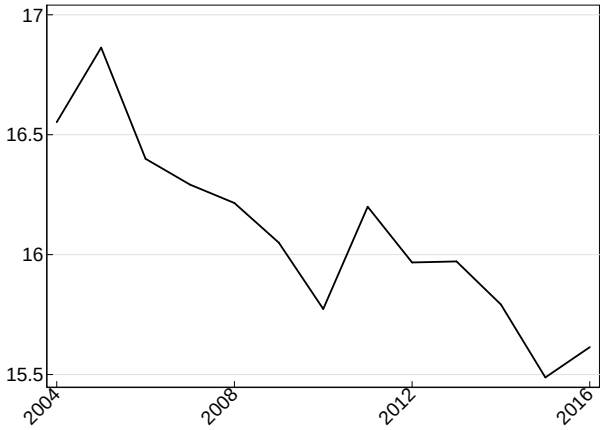
Figure 1a plots $\mathcal{H}_t^{\text{HH}}$ and reveals a nearly monotonic increase in Household concentration from 2004-2016.[11] Figure 1b shows that this increase in concentration is also associated with a decline in the average number of products consumed per household within a product category.[12] We delay interpreting the quantitative magnitude of these changes until we develop our model in Section 3 but note now that fitting these data with a linear trends yields precise and highly significant estimates.

**Figure 1: Household Product Concentration**

**(a):** Household Herfindahl




**(b):** Ave # Products Per Household




## 2.3 Aggregate Spending Concentration

What underlies this increase in the concentration of household expenditures? Motivated by the rise of "super-star" firms documented in David, Dorn, Katz, Patterson, and Van Reenen (2017), one possible explanation is that there has been an increase in the importance of "super-star" products that

[11] 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 3.

[12] Average product number counts are constructed using the same weights as were used in equations (3) and (5).

increasingly dominate their product categories. 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 aggregate market share of $j$ in $c$ as:

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

and the Aggregate Herfindahl in category $c$ as:

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

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 3, 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.[13]

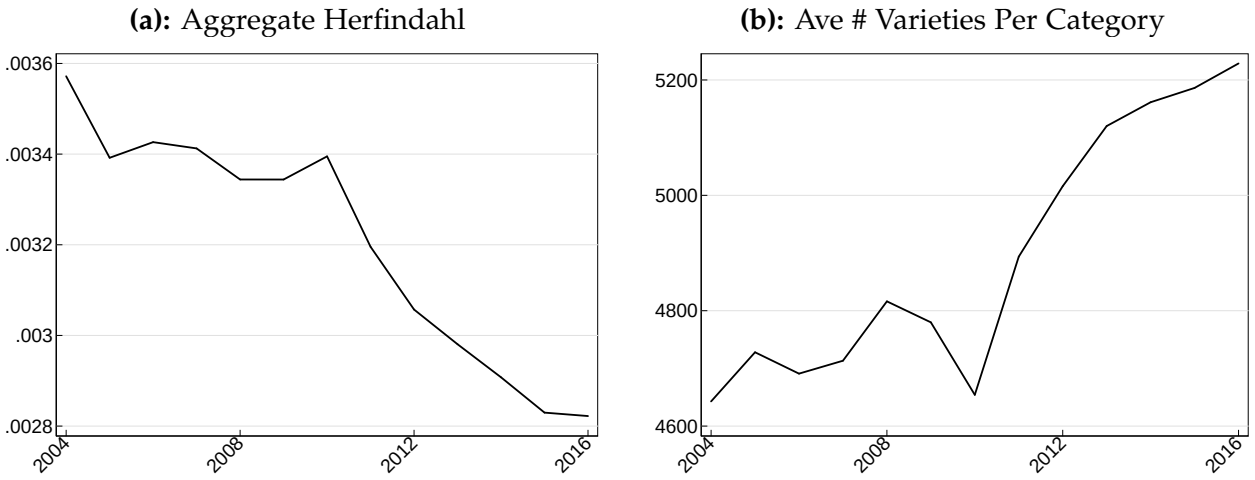
This result may, at first, seem at odds with the rise in sales concentration measured in Census data by papers including David, Dorn, Katz, Patterson, and Van Reenen (2017). Our aggregate concentration measure, however, captures expenditures at the product level whereas Census-based estimates aggregate products to the producer level.[14] The resulting trends may therefore differ significantly, particularly in the face of changes in the number of goods each manufacturer produces.

[13]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.

[14]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.

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.[15] We therefore conclude that our results are broadly consistent with the Census-based literature. Which of producer vs. product concentration is of interest depends, of course, on the question of interest. 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 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.

## 2.4 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 A1-A6 show that these divergent trends continue to hold if we exclude generics, compute

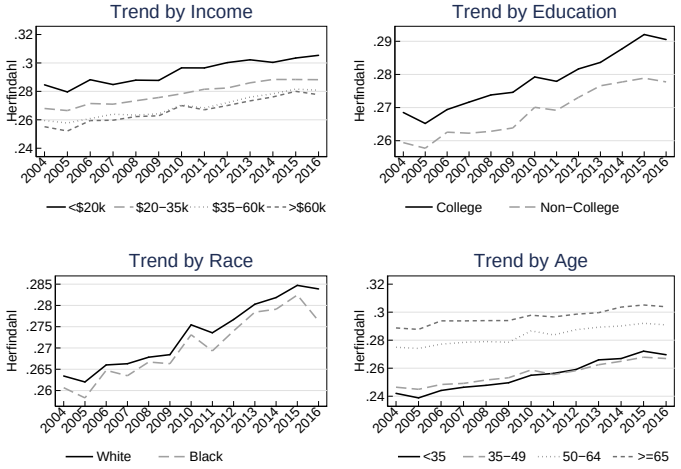
[15] We further measure in our data an increase in the share of the top retailers or stores, a finding unrelated to our focus but one that closely resembles the results for the Retail industry in David, Dorn, Katz, Patterson, and Van Reenen (2017).

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.

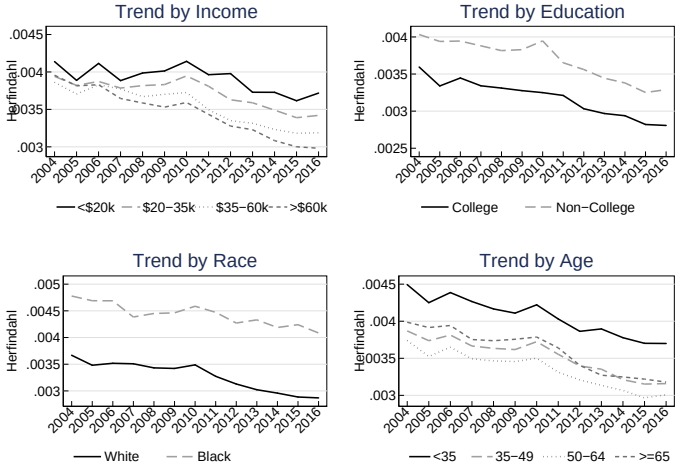
Is the rise of niche consumption driven by shifts in the importance of different groups, such as old and young or rich and poor? While there are differences in the level of concentration across different groups, the trends are primarily driven by *within* group variation.[16] To show this, we re-calculate 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 each of these groups. 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**

**(a):** Household Concentration




**(b):** Aggregate Concentration




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.[17]

As a simple summary statistic for the prominence of niche consumption, we consider the ratio of

[16] 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.

[17] Since $\mathcal{H}_i^{\text{HH}}$ 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 $\mathcal{H}^{\text{agg}}$ separately for each demographic group. For example, it is straightforward to regress $\mathcal{H}_i^{\text{HH}}$ on age $\times$ income $\times$ education bins and this requires no new calculation of $\mathcal{H}_i^{\text{HH}}$. In contrast, calculating how aggregate concentration varies simultaneously with age $\times$ income $\times$ education requires calculating new aggregate product market shares for each of these different groups. This makes these calculations substantially more computationally burdensome. Even more importantly, measurement error in aggregate market shares increases rapidly for more narrow demographic groups.

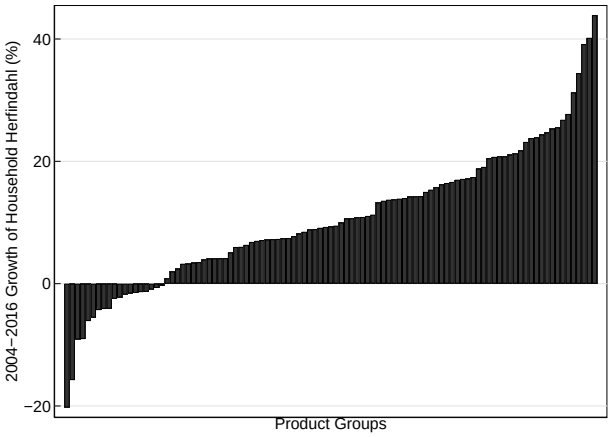
the Household Herfindahl to the Aggregate Herfindahl. A higher value for this "niche ratio" means that household consumption is more segmented into different niches. Figure 4 shows that the rise of niche consumption is pervasive across product categories, with three-quarters of product categories exhibiting increases in $\mathcal{H}_{c,t}^{\text{HH}}$, eighty percent of product categories exhibiting decreases in $\mathcal{H}_{c,t}^{\text{Agg}}$, and growth in the niche ratio in 92 percent of the categories.

In Appendix Figure A7, 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 A8 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. 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.

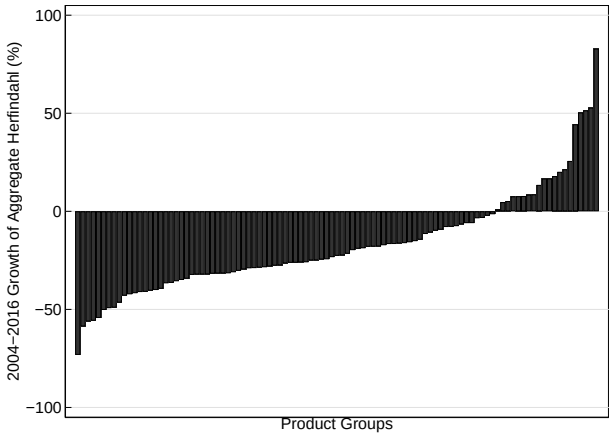
## 2.5 The Role of Product Churn

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

**(a):** Household Concentration




**(b):** Aggregate Concentration




While the rise in niche consumption is pervasive, 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".[18]

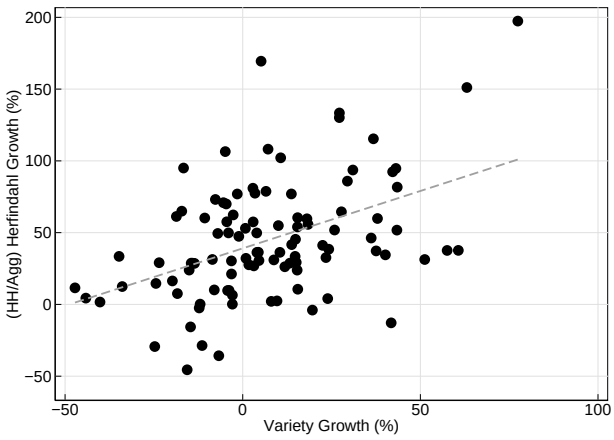
Interestingly, there is a common observable linking together the categories with the most rapid

[18]The level of the niche ratio is highest in "Cosmetics" and "Fragrances-Women" and is lowest for "Charcoal" and "Dough Products".

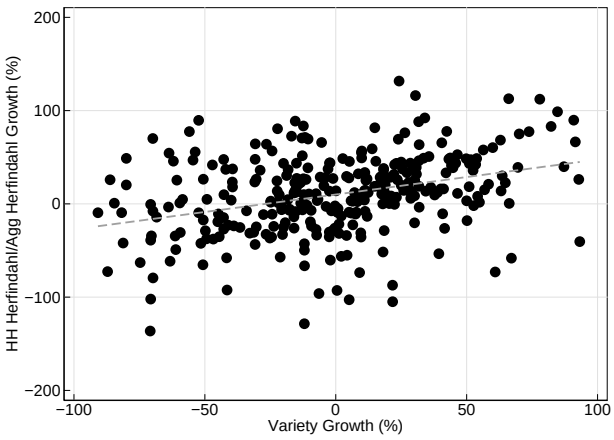
increases in niche ratios: they are also the categories with the fastest growth in the number of aggregate products, measured using the same assumptions as were used for Figure 2b above. We emphasize this relationship as it will be central to the mechanism in our model in Section 3 and its implications 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.

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

**(a):** Product Groups




**(b):** Retailers




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.[19] 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.[20] Figure 6a shows the upward trend in household concentration is much stronger when using all UPCs than when restricting to continuing products, implying a large role for product entry and exit in generating household concentration increases.[21] Figure 6b shows that when

[19] 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.

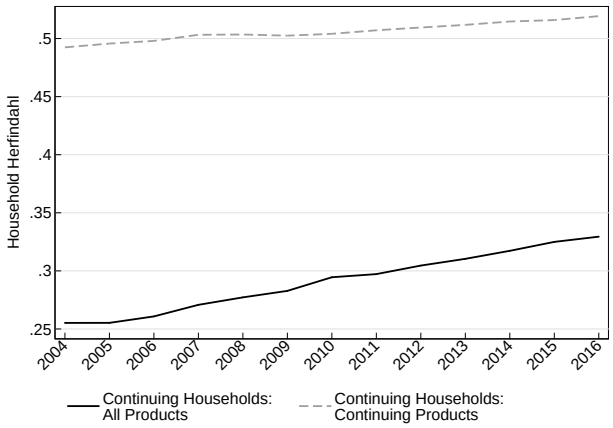
[20] We report cumulative changes rather than levels since this procedure results in two levels of the Herfindahl for each year (one initial and one continuation year), so the levels are not meaningful.

[21] Note that the trend for “All Products" is larger than in Figure 1, since this analysis is using within-household variation.

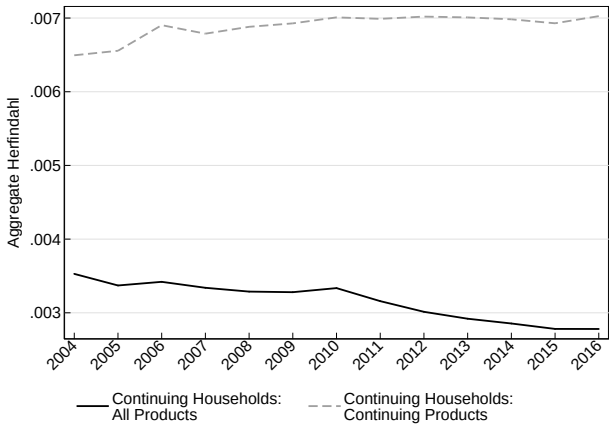
focusing only on continuing products, aggregate concentration actually rises instead of declines.[22]

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

**(a):** Household Herfindahl




**(b):** Aggregate Herfindahl




## 3 Modeling the Rise of Niche Consumption

In this section, we develop a model able to match the rise of niche consumption documented in Section 2 and we use it to identify key driving forces and understand the resulting implications for welfare and 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 are identical, assume that all households consume a single product, or assume that household tastes are symmetric across products. 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) and Arkolakis, Demidova, Klenow, and Rodriguez-Clare (2008), 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". We assume each household's tastes for products, adjusted for price, are distributed Pareto, which allows us to write the Household Herfindahl analytically. Further, while we assume each household's tastes are distributed Pareto, we introduce a rank function that implies the preference ordering of products will differ across households and allows us to write an analytical expression for the Aggregate Herfindahl.

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 emphasize

See Appendix C.1 for discussion.

[22] 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 A10 shows that changes in quantity rather than price are more important.

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 adjustment 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.

### 3.1 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:[23]

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

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 measure of varieties consumed.[24]

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:

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

where $y \geq b > 0$ and where we assume $\theta > 2\left(\sigma - 1\right)$. 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. Household $i$ will consume the set of goods with $\tilde{\gamma}_k \in [\tilde{\gamma}^*, \infty)$ for some $\tilde{\gamma}^* \geq b$. We note that with this setup, all households will get the same marginal utility out of their top-choice (in a price-adjusted sense) product, their second-choice product, and so on.

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

$$P_i = P \quad = \quad \left( \int_{k \in \Omega_i} \left( \tilde{\gamma}_{i,k} \right)^{\sigma-1} dk \right)^{\frac{1}{1-\sigma}}$$

[23] 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.

[24] We also assume $\epsilon > 1/(1-\sigma) - 1/\theta$, which implies that higher fixed costs $F$ lead to less purchased products $|\Omega|$.

$$= \underbrace{\left(1+\frac{1-\sigma}{\theta}\right)^{\frac{1}{\sigma-1}} b^{-1}}_{\text{Average Price}} \times \underbrace{\left(|\Omega_i|\right)^{\frac{1}{1-\sigma}}}_{\text{Variety Gains}} \times \underbrace{\left(\frac{|\Omega_i|}{N}\right)^{\frac{1}{\theta}}}_{\text{Selection Effects}} . \tag{9}$$

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 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.[25]

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

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

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

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

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

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

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

[25] The index reduces to more standard expressions in special cases. For example, consider $\theta \to \infty$, which implies that households value all products identically, with $\tilde{\gamma}_{i,k} = b$. 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.

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

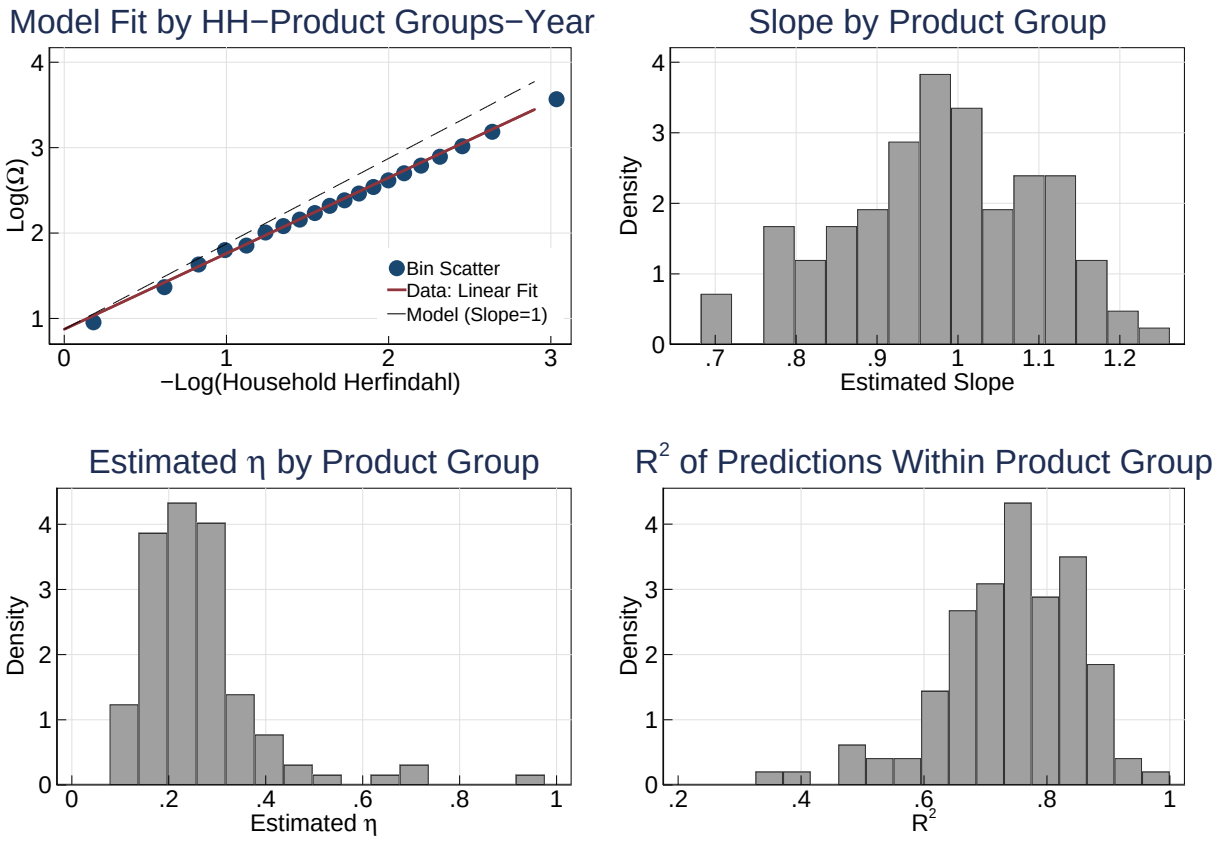
### 3.2 Household Herfindahls

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

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

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, or $\partial\mathcal{H}^{\text{HH}}/\partial|\Omega| < 0$. And for fixed $|\Omega|$, concentration declines monotonically with $\eta$, or $\partial\mathcal{H}^{\text{HH}}/\partial\eta < 0$.

**Figure 7: Model Fit on Household-Category Data**




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 $\mathcal{H}_{i,c}^{HH}$ 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. and 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.[27] In the upper right panel, we estimate these regressions separately for each category in 2016 and 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).[28] 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.

### 3.3 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 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 $i$:

$$r_{i,j} = (1 - \alpha)j + \alpha x_{i,j}, \tag{12}$$

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 as their aggregate ranking. If $\alpha$ approaches 1, tastes are purely idiosyncratic and resulting aggregate spending will be evenly distributed over all 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

[27] This fixed effect specification allows for heterogeneity in $\eta$ across category-years but 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. This specification has large explanatory power even though $\eta$ varies across category-years but is common across households.

[28] 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. This shows that the flexibility afforded by our Pareto taste distribution (which is what allows for $\eta < 1$) is crucial for matching differences in concentration across categories

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 $R(r)$. Without loss of generality, we assume $\alpha < 1/2$ and write:[29]

$$R(r) = \begin{cases} \frac{1}{2}\left(\frac{r}{N}\right)^2 \frac{1}{\alpha(1-\alpha)}, & 0 \leq r < \alpha N \\ \frac{r}{N}\frac{1}{1-\alpha} - \frac{1}{2}\frac{\alpha}{1-\alpha}, & \alpha N \leq r < (1-\alpha)N \\ -\frac{1}{2}\left(\frac{r}{N}\right)^2 \frac{1}{\alpha(1-\alpha)} + \frac{r}{N}\frac{1}{\alpha(1-\alpha)} - \frac{1}{2}\left(\frac{\alpha}{1-\alpha} + \frac{1-\alpha}{\alpha}\right), & (1-\alpha)N \leq r \leq N. \end{cases} \tag{13}$$

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. $R(r)$ has three regions depending on whether the CDF has no mass points (case 1), strictly positive mass at 0 (case 2) or strictly positive mass at 0 and $N$ (case 3). The first region corresponds to parameters in which there is no product that is purchased by all households in the economy while the third region corresponds to the case when even the worst possible product is consumed by some household (and most products will be consumed by all households).[30]

As long as $0 \leq \frac{|\Omega|}{N} < \frac{\alpha}{2(1-\alpha)} < \frac{1}{2}$, it can be shown that we will only be in 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 $\tilde{\gamma}_{i,j} = G^{-1}\left(1 - R\left(r_{i,j}\right)\right)$, the spending share of household $i$ that is dedicated to product $j$ can be written as a function of $j$'s rank:

$$s_{i,j} = P^{\sigma-1}\tilde{\gamma}_{i,j}^{\sigma-1} = (Pb)^{\sigma-1}\left(R\left(r_{i,j}\right)\right)^{-\frac{\sigma-1}{\theta}} = \frac{\eta+1}{2}N^{\frac{\eta-1}{2}}|\Omega|^{-\frac{\eta+1}{2}}\left(R\left(r_{i,j}\right)\right)^{\frac{\eta-1}{2}}, \tag{14}$$

if $R\left(r_{i,j}\right) \leq |\Omega|/N$ and zero otherwise. To determine the products for which the share $s_{i,j}$ in equation (14) 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

[29] Replacing $\alpha$ with $1-\alpha$ in all instances in equation (13) 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.

[30] 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.

(13) under the assumption that $0 \leq \frac{|\Omega|}{N} < \frac{\alpha}{2(1-\alpha)}$, we get:

$$r^* = (2\alpha(1-\alpha)|\Omega|N)^{\frac{1}{2}}. \tag{15}$$

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:

$$j^* = \left(\frac{2\alpha|\Omega|N}{1-\alpha}\right)^{\frac{1}{2}}. \tag{16}$$

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 (12) into the expression (15), and using the definition of $j^*$ in equation (16), we can write the highest value or "cutoff" random draw $x_j^*$ that yields positive consumption of $j$ as:

$$x_j^* = \frac{1-\alpha}{\alpha}(j^* - j). \tag{17}$$

### 3.4 The Aggregate Herfindahl

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

$$\begin{aligned}
s_j &= \frac{1}{\int_i E di}\int_i E s_{i,j} di \\
&= \frac{\eta+1}{2} N^{\frac{\eta-1}{2}} |\Omega|^{-\frac{\eta+1}{2}} \int_0^{\frac{1-\alpha}{\alpha}(j^*-j)} 2^{\frac{1-\eta}{2}} N^{1-\eta} \left(\alpha(1-\alpha)^{\frac{1-\eta}{2}}\right) \left((1-\alpha)j + \alpha x_{i,j}\right)^{\eta-1} \frac{dx}{N} \\
&= (\eta+1)(2N|\Omega|)^{-\frac{\eta+1}{2}} (\alpha(1-\alpha))^{\frac{1-\eta}{2}} \int_0^{\frac{1-\alpha}{\alpha}(j^*-j)} \left((1-\alpha)j + \alpha x_{z,i}\right)^{\eta-1} dx \\
&= \frac{\eta+1}{\eta}\left(\frac{2\alpha N|\Omega|}{1-\alpha}\right)^{-\frac{\eta+1}{2}} \left((j^*)^\eta - j^\eta\right) \\
&= \frac{\eta+1}{\eta j^*}\left(1 - \left(\frac{j}{j^*}\right)^\eta\right).
\end{aligned} \tag{18}$$

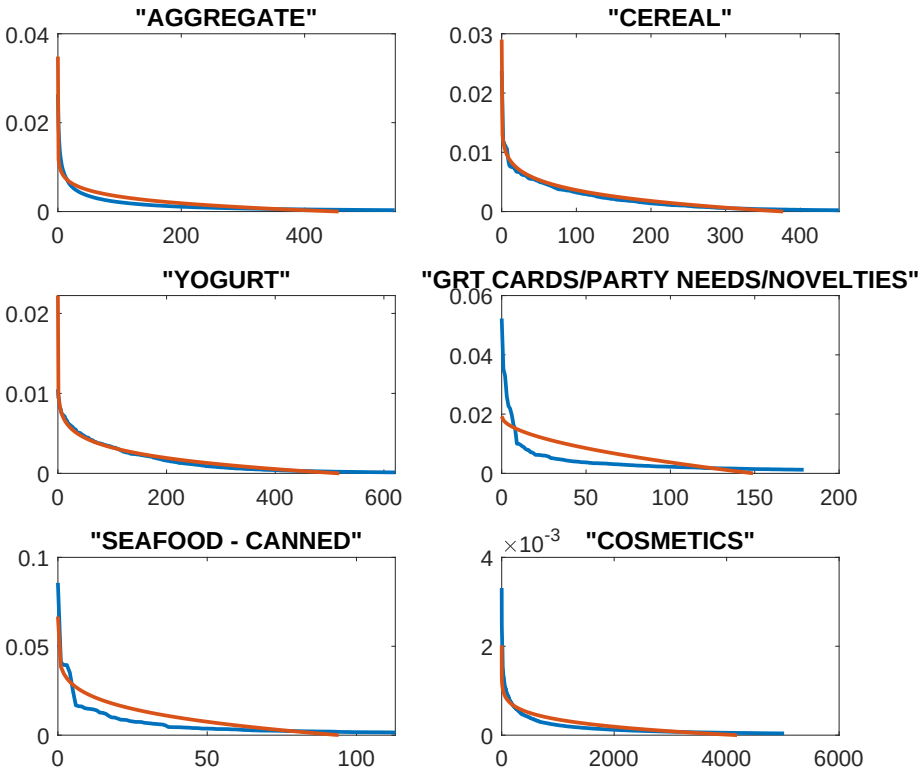
Using equation (18), we immediately obtain the Aggregate Herfindahl:

$$\mathcal{H}^{\text{Agg}} = \int_{j=0}^{j^*} s_j^2 dj = \left(\frac{\eta+1}{\eta j^*}\right)^2 \int_{j=0}^{j^*} \left(1 - \left(\frac{j}{j^*}\right)^\eta\right)^2 dj$$

$$
\begin{aligned}
&= \left(\frac{\eta+1}{\eta j^*}\right)^2 j^* \left(1 - \frac{2}{\eta+1} + \frac{1}{2\eta+1}\right) \\
&= \frac{2(\eta+1)}{(2\eta+1)} \frac{1}{j^*} \\
&= \frac{2(\eta+1)}{(2\eta+1)} \left(\frac{1}{2\tilde{N}|\Omega|}\right)^{\frac{1}{2}}, \qquad (19)
\end{aligned}
$$

where we define $\tilde{N} = N\alpha/(1-\alpha)$. Concentration declines monotonically with $\tilde{N}$ as $\partial \mathcal{H}^{\text{Agg}}/\partial|\tilde{N}| < 0$. For fixed $\theta$ and $\sigma$, aggregate concentration declines monotonically with the number of consumed products as $\partial \mathcal{H}^{\text{Agg}}/\partial|\Omega| < 0$. And for fixed $|\Omega|$, concentration declines monotonically with $\eta$ as $\partial \mathcal{H}^{\text{Agg}}/\partial\eta < 0$. Importantly, the signs of these $|\Omega|$ and $\eta$ effects are identical to those obtained for the household Herfindahl, $\mathcal{H}^{\text{HH}}$. 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 $\tilde{N}$.

**Figure 8: Model Fit on Aggregate Category Data**




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 $\mathcal{H}^{\text{Agg}}$ and $\mathcal{H}^{\text{HH}}$. Figure 8 then plots the market share distribution

implied by our model in equation (18) (in red) against the actual market share distribution in the data (in blue) for a number of product categories as well as for the aggregate across Nielsen categories. 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 to validate our use of the model, particularly given the distributions are fully determined by only three parameters and reflect parametric assumptions and functional forms chosen largely for analytical convenience.

### 3.5 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:

$$\begin{aligned} \mathcal{H}^{\text{HH}} &= \frac{(\eta+1)^2}{4\eta}\frac{1}{|\Omega|} \\ \mathcal{H}^{\text{Agg}} &= \frac{2(\eta+1)}{(2\eta+1)}\left(\frac{1}{2\tilde{N}|\Omega|}\right)^{\frac{1}{2}}. \end{aligned}$$

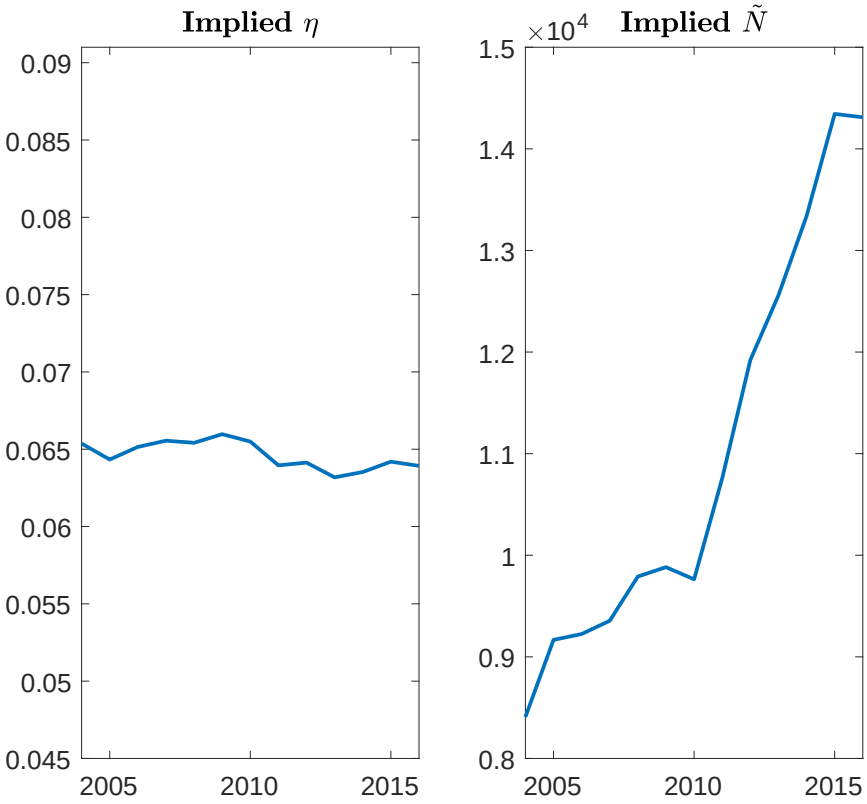
Since $\mathcal{H}^{\text{HH}}$, $\mathcal{H}^{\text{Agg}}$, and $|\Omega|$ are directly observable in the data, this produces a system of two equations that can be solved to determine $\eta$ and $\tilde{N}$ for each year. Figure 9 shows the time-series for $\tilde{N}$ and $\eta$ necessary to hit these 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. There are theoretical combinations of changes in $|\Omega|$ and $\eta$ that leave $\mathcal{H}^{\text{HH}}$ increasing with $\mathcal{H}^{\text{Agg}}$ decreasing, even without a change in $\tilde{N}$. However, the data imposes restrictions on $|\Omega|$, and $\eta$. While it is obviously possible that the trends in the data might have some measurement error, we can show that even if we allow for substantial flexibility in the trends on $\Omega$ and $\eta$ to account for this type of empirical uncertainty, it is still the case that matching the quantitative divergence of concentration measures requires some growth in $\tilde{N}$. Since multiple factors can cause the observed changes in $|\Omega|$ but all explanations of the data will require an increase in $\tilde{N}$, we focus our analysis on this force.

Mechanically, increases $\tilde{N}$ can in turn arise from increases in $\alpha$ or $N$. Changes in $\alpha$ are straightforward to interpret, since $\alpha$ is simply an exogenous parameter governing preference heterogeneity.

**Figure 9: Modeling the Rise of Niche Consumption**




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.[31] Based on this logic, we assume there were no changes in $\alpha$ and start with an exploration of increases in $N$ in the model, holding constant all other parameters, including the price-adjusted taste distribution. We will later discuss the implications of additionally changing other parameters, which is necessary to fully match the empirical patterns.

[31] Jaravel (2019) uses Bartik style variation to disentangle the effects of increased demand for new products from increased supply side innovation in the context of household inflation differences. Directly using his empirical methodology isolates the effects of broad demand shifts across demographic groups and product categories, but it does not deliver exogenous shocks to heterogeneity within-groups, which is the relevant movement we would need to isolate. However, it seems possible that some variation on this approach might be able to isolate this type of heterogeneity in the data and this is an interesting avenue for ongoing research.

### 3.6 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.[32] We then calculate the change in household 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:

$$U_{2016} = \frac{E}{P_{N_{2016}}} - F \times \left(|\Omega_{N_{2016}}|\right)^{\epsilon} = 1.095 \times \frac{E}{P_{N_{2004}}} - F \times \left(|\Omega_{N_{2004}}|\right)^{\epsilon}, \tag{20}$$

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 (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 $\mathcal{H}^{\text{Agg}}$.[33] 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 actually 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 $\left(E\left(1+\left(1-\sigma\right)/\theta\right)^{\frac{1}{1-\sigma}} b\right)$, increases in effective costs per number of products consumed ($F$ or $\epsilon$), or declines in an "effective curvature" of utility term $\left(\frac{1}{1-\sigma} - \frac{1}{\theta}\right)$.[34] Expenditures, however, increase in the data, and while changes in either $\sigma$ or $\theta$ could change the curvature term, they would

[32] 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 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 $\Omega$.

[33] We also consider alternative decompositions of the 70 percent growth of $\tilde{N}$ into growth of $\alpha$ vs. $N$. Alternative decompositions matter for welfare because $N$ impacts the extent of selection effects while $\alpha$ does not. As such, the more that growth in $\tilde{N}$ is driven by growth in $\alpha$, the smaller are the welfare gains. We find, however, that unless changes in idiosyncratic preferences within narrow demographic groups are the overwhelming force driving the rise of niche consumption, our model implies that greater niche consumption is associated with large welfare gains, mostly arising from selection effects.

[34] Real measured expenditures equal $E$ divided by the expression labeled "Average Price" in equation (9).

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 $\epsilon$.[35]

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 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.

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

Concentration measures 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 also on an endogenous "extensive margin" force that arises from the possibility that products would gain new customers (or lose existing ones). Since these forces are differentially important to 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 would not be purchased, which pushes their markups down. 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. As a result, our model delivers large changes in aggregate and

[35] 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. Our model implicitly assumes that it is costless for households to rank all products, while in reality the process of ranking products and determining the ideal basket is likely more costly when there are more products. Furthermore, even if a household knows exactly what product they want to purchase and that it is carried by a particular store, physically finding it in the store is likely more costly if the store carries a much larger number of products. An increase in $F$ or $\epsilon$ can then be interpreted as a simple proxy for these forces when it is accompanied by the increase in $N$ which we previously argued is necessary to match the data.

household concentration without any change in aggregate market power.[36]

To solve for the price elasticity of aggregate demand for product $j$, we express its total sales as the integral of spending on $j$ by each households. Since $j$ will only be purchased by those households with a sufficiently high idiosyncratic taste for it $x_{i,j}$, we need only integrate from households drawing $x_{i,j} = 0$ to the marginal household that draws $x_{i,j} = x_j^*$. We then differentiate with respect to $p_j$ to get:

$$\frac{\partial s_j}{\partial p_j} = \frac{1}{N}\frac{\partial \int_0^{x_j^*} s_{i_j,j} dx}{\partial p_j} = \frac{1}{N}\left(\int_0^{x_j^*} \frac{\partial s_{i,j}}{\partial p_j} dx + s_{i_j^*,j}\frac{\partial x_j^*}{\partial p_j}\right), \tag{21}$$

where the right hand side of equation (21) follows from Leibniz's rule. The first term can be solved as:

$$\frac{\partial s_{i,j}}{\partial p_j} = \frac{\partial P^{\sigma-1} p_j^{1-\sigma} \gamma_{i,j}^{\sigma-1}}{\partial p_j} = (1-\sigma)\frac{s_{i,j}}{p_j}, \tag{22}$$

where we take the aggregate price index $P$ as fixed. Moving on to the second term, we have:

$$s_{i_j^*,j}\frac{\partial x_j^*}{\partial p_j} = \frac{\eta+1}{2} N^{\frac{\eta-1}{2}} |\Omega|^{-\frac{\eta+1}{2}} \left(R\left(r^*\right)\right)^{\frac{\eta-1}{2}} \frac{\partial x_j^*}{\partial p_j}. \tag{23}$$

Substituting everything back into equation (21), we get:

$$\begin{aligned}
\frac{\partial s_j}{\partial p_j} &= \frac{1}{N}\frac{\partial \int_0^{x_j^*} s_{i,j} dx}{\partial p_j} = \frac{1}{N}\left(\int_0^{x_j^*} \frac{\partial s_{i,j}}{\partial p_j} dx + s_{i_j^*,j}\frac{\partial x_j^*}{\partial p_j}\right) \\
&= (1-\sigma)\frac{1}{p_j}\frac{1}{N}\int_0^{x_j^*} s_{i,j} dx + \frac{1}{N} s_{i_j^*,j}\frac{\partial x_j^*}{\partial p_j} \\
&= (1-\sigma)\frac{s_j}{p_j} + \frac{1}{N}\frac{\eta+1}{2} N^{\frac{\eta-1}{2}} |\Omega|^{-\frac{\eta+1}{2}} \left(R\left(r^*\right)\right)^{\frac{\eta-1}{2}} \frac{\partial x_j^*}{\partial p_j} \\
&= (1-\sigma)\frac{s_j}{p_j} + \frac{1}{N}\frac{\eta+1}{2} N^{\frac{\eta-1}{2}} |\Omega|^{-\frac{\eta+1}{2}} \left(\frac{1}{2} N^{-2} \alpha^{-1} (1-\alpha) \left(j^*\right)^2\right)^{\frac{\eta-1}{2}} \frac{\partial x_j^*}{\partial p_j} \\
&= (1-\sigma)\frac{s_j}{p_j} + \frac{\eta+1}{2N|\Omega|}\frac{\partial x_j^*}{\partial p_j}.
\end{aligned} \tag{24}$$

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

[36] 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 might call into question some 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.

price increases, we start with the relationship:[37]

$$b^{\theta}\gamma_{i,j}^{-\theta}p_j^{\theta} = 1 - G\left(\frac{\gamma_{i,j}}{p_j}\right) = R\left((1-\alpha)\,j + \alpha x_{i,j}\right), \tag{25}$$

and totally differentiate to yield:

$$\theta b^{\theta}\gamma_{i,j}^{-\theta}p_j^{\theta-1}dp_j = \frac{r_{i,j}}{1-\alpha}\frac{1}{N^2}dx_{i,j}. \tag{26}$$

We evaluate at $r_{i,j} = r^*$ and $\tilde{\gamma}_{i,j} = \tilde{\gamma}^*$ 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, and get:

$$\frac{dx_j^*}{dp_j} = -\frac{\theta}{j^*}|\Omega|N\frac{1}{p_j}. \tag{27}$$

Inserting this into equation (24), we have:

$$\begin{aligned}\frac{\partial s_j}{\partial p_j} &= (1-\sigma)\frac{s_j}{p_j} - \frac{\eta+1}{2N|\Omega|}\frac{\theta}{j^*}|\Omega|N\frac{1}{p_j} \\ &= (1-\sigma)\frac{s_j}{p_j} - \frac{\eta\theta}{2\left(1-\left(\frac{j}{j^*}\right)^{\eta}\right)}\frac{s_j}{p_j}.\end{aligned} \tag{28}$$

Equation (28) implies that product $j$'s price elasticity of demand $\varepsilon_j$ can be written as:

$$\begin{aligned}\varepsilon_j &= 1 - \frac{\partial s_j}{\partial p_j}\frac{p_j}{s_j} \\ &= \underbrace{\sigma}_{\text{Intensive Margin}} + \underbrace{\left(1-\left(\frac{j}{j^*}\right)^{\eta}\right)^{-1}\left[\theta/2-(\sigma-1)\right]}_{\text{Extensive Margin}} > \sigma,\end{aligned} \tag{29}$$

where 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. High $j$ or "niche" products are consumed by few households, so the extensive margin is increasingly important. As a result, the elasticity of demand increases as market share falls.[38] As $j \to j^*$ and a product approaches the point where it is dropped from the aggregate consumption bundle, the elasticity approaches infinity, i.e. $\varepsilon \to \infty$. The markup $\mu_j$

[37] In Appendix E we use numerical simulations which do not require any distributional assumption on tastes to verify that the approximation that follows is extremely accurate.

[38] 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.

can then be written as:

$$\mu_j = \frac{\varepsilon_j}{\varepsilon_j - 1} = \frac{\sigma + \frac{\theta(\eta+1)}{2j^* s_j}}{\sigma + \frac{\theta(\eta+1)}{2j^* s_j} - 1}, \tag{30}$$

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

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 (18) and (29), it can be written as:

$$\mu^{Agg} = \frac{\int_0^{j^*} s_j dj}{\int_0^{j^*} s_j \frac{\varepsilon_j - 1}{\varepsilon_j} dj} = \left[ \frac{\theta + (\sigma - 1)^2}{\sigma^2} - \frac{1}{2} \frac{\eta \theta^2}{\sigma^2} \left( \frac{\eta + 1}{2 + \theta} \right) \times {}_2F_1 \left( 1, \frac{1}{\eta}; 1 + \frac{1}{\eta}; \frac{2\sigma}{2 + \theta} \right) \right]^{-1}, \tag{31}$$

where ${}_2F_1(\cdot)$ is the hypergeometric function.[39] 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 have 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 $j^*$ is not in the expression for aggregate markups. 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. As $j^*$ increases, any particular $j$ becomes further from $j^*$ and sees a decline in the extensive margin of demand, which is a force which increases aggregate market power and profit shares.[40] At the same time, as $j^*$ increases, the distribution of sales flattens.[41] Holding markups for individual $j$ constant, this increases the relative weight on low markup products in overall spending and is a force for declining aggregate profit shares.

Both forces are individually quite powerful, but in our model they exactly offset and there are no changes in aggregate profit shares. However, this does not mean that the distribution of profits in the economy is unchanged. Indeed, as just discussed, changes in the distribution of profits over $j \in (0, j^*)$ are crucial for understanding the invariance of aggregate markups. Figures 10-12 change $\tilde{N}$ and $F$ to fit the rise of niche consumption and illustrate the effects on distributions.

[39] 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 $n = 0$.

[40] The interpretation of this decline in the importance of the extensive margin will depend on the underlying structural shock which raises $j^*$. If $N$ increases, products face a more selected set of households who are less likely to switch. If $\alpha$ increases, idiosyncratic tastes for particular products are more extreme and if $F$ decreases, the overall basket households consume is larger and so given $j$'s become more inframarginal.

[41] When driven by $N$ this can be interpreted as arising from greater competition with new products. When driven by $\alpha$ it can be interpreted as greater competition from "tail" products, and when arising from declines in $F$ it just occurs from households expanding the basket over which they allocate spending to include a more diverse set of products.

While we do not have data on product markups to directly test these distributional implications, it is important to make a particular caveat about the interpretation of these distributional changes that would be relevant for taking them to the data. Our results do not require taking a stand on whether the change in $\tilde{N}$ is driven by $N$ or $\alpha$, but the distinction between changes driven by $N$ and those driven by $\alpha$ would be important for testing certain implications in the data. In response to changes in $\alpha$ or $F$, it is natural to assume that a particular product $j$'s aggregate rank does not change; that is $j^{new} = j^{old}$. In response to increases in $N$, this is no longer clear, since the set of available products changes, and the extent to which these new products are better or worse than existing products will depend on assumptions about entry. In order to maintain a Pareto distribution, it is natural to assume that new products have a quality which is uniformly distributed over the product space, so that in expectation $j^{new} = j^{old} \frac{N^{new}}{N^{old}}$. However, alternative assumptions about entry would lead to alternative changes in incumbent product rankings. For example if new products enter at the highest possible quality, then $j^{new} = j^{old} + N^{new} - N^{old}$.

This means that in response to increases in $N$, our model cannot pin down the effect on any particular incumbent product without making auxiliary assumptions on the nature of entry. However, even with no assumptions on entry, our model does pin down directly the distribution over $j \in (0, j^{*,old})$ and $j \in (0, j^{*,new})$. For example, when $N$ increases, our model implies that the median ranked product (relative to $j^*$) in the low $N$ environment will have the same markup as the median ranked product (relative to $j^*$) in the high $N$ environment. But without additional assumptions on entry, our model does not pin down how the rank of any particular incumbent product changes when $N$ increases, and so has no predictions for markup or market share changes for fixed products in the data. Instead, testable implications obtain for the distributions of demand, markups and profits.[42]

# 4 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

[42]For example, Figure 10 shows that higher j-ranked products have more market share after an increase in $N$ or $\alpha$. If this arises from $\alpha$, one can interpret this as saying that particular low-ranked products gain market share when. If this arises from an increase in $N$, one should not interpret this as saying that greater competition raises market share for any particular products. If one assumes the natural benchmark in which products enter uniformly over the product space, one can show that if one focuses on any particular product, an increase in competition always reduces market share.

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

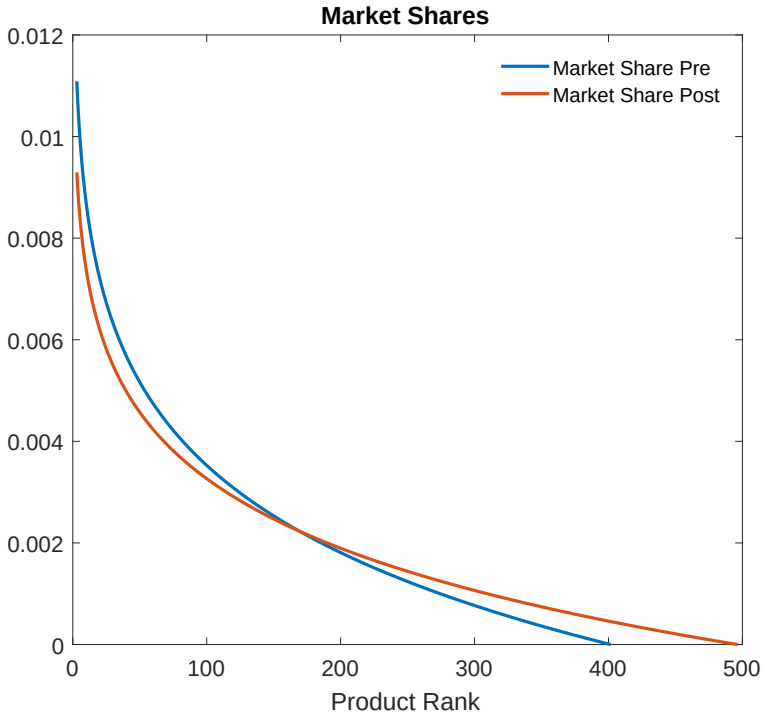



**Figure 11: Effects of $\tilde{N} \uparrow$ and $F \uparrow$ on Markup Distribution**

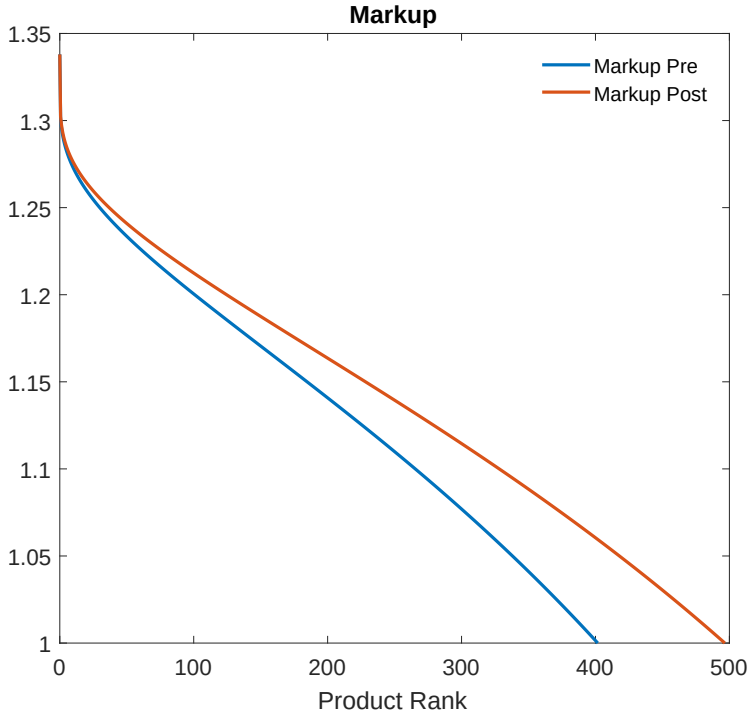
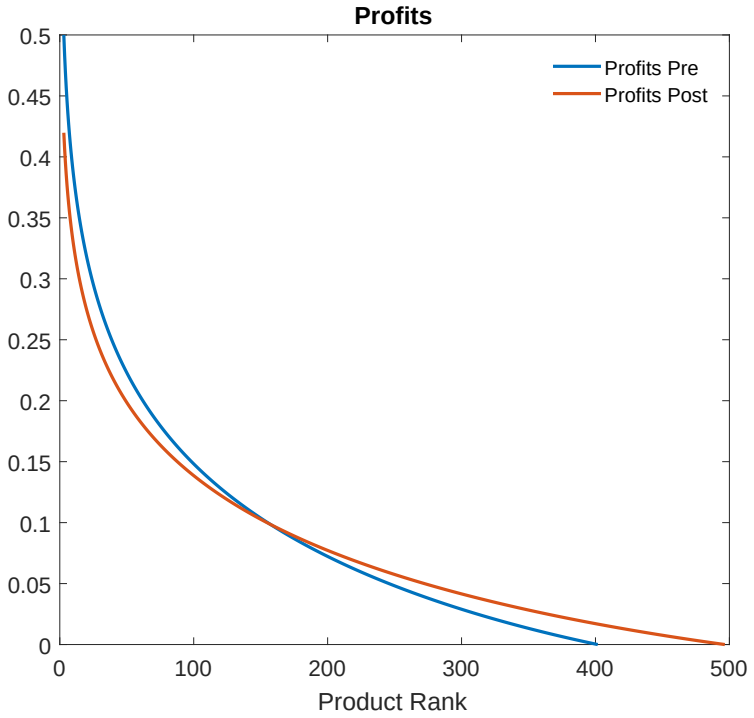


**Figure 12: Effects of $\tilde{N} \uparrow$ and $F \uparrow$ on Profit Distribution**




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 best satisfy 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.

While our paper highlights the growing prominence of niche products in US retail spending, associates these changes with substantial increases in overall product availability, and argues that the interaction of this greater product choice with household heterogeneity has important welfare implications, our model treats the set of products available as an exogenous parameter. We think better understanding the underlying drivers of greater product growth is an important avenue for future empirical and theoretical work. Recent decades have seen continued global integration of supply chains, increasing sophistication of inventory management and IT technology, more and more sophisticated tracking of individual household preferences and targeted advertising, rising consolidation of the retail sector, and growing importance of online competition. Some or all of these changes are likely

important for understanding the incentives for manufacturers to create and retailers to stock new products. Understanding the relative importance of these channels and their potential interactions with household heterogeneity is a fruitful avenue for future research.

## 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.

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.

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.

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.

David, H., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2017): "The fall of the labor share and the rise of superstar firms," .

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 âĂIJA Just SocietyâĂİ 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.

Hansen, K., and V. Singh (2015): "Choice Concentration," Discussion paper.

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.

Levin, A., and T. Yun (2008): "Reconsidering the microeconomic foundations of price-setting behavior," *Unpublished Manuscript, Board of Governors of the Federal Re serve 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*.

Olivier Coibion, Yuriy Gorodnichenko, D. K. (2017): "Consumption Inequality and the Frequency of Purchases," Discussion paper.

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.

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*.

# Appendix A. Additional Empirical Results

**Figure A1: Concentration Trends: Excluding Generics**

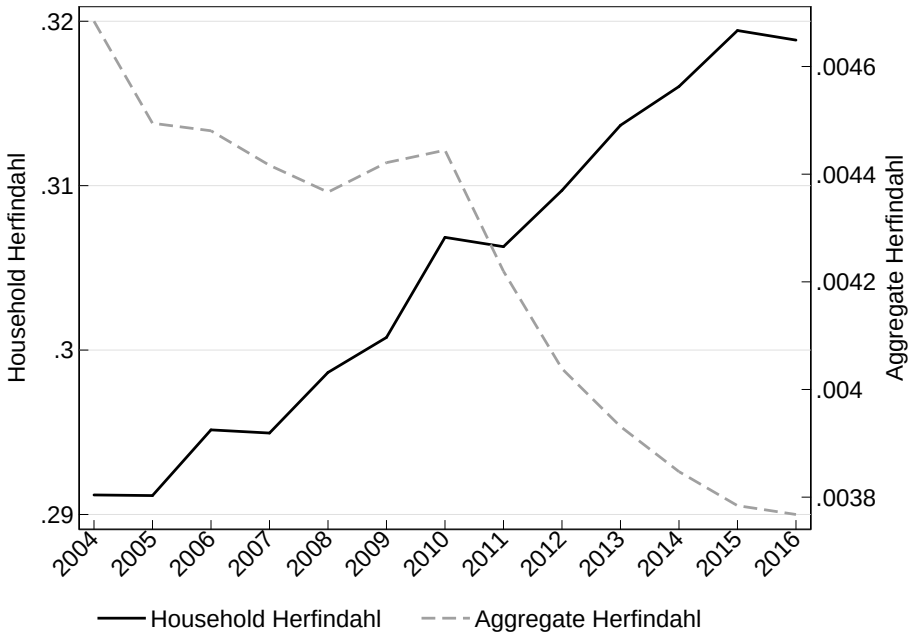


**Figure A2: Concentration Trends: Including Category Composition Changes**

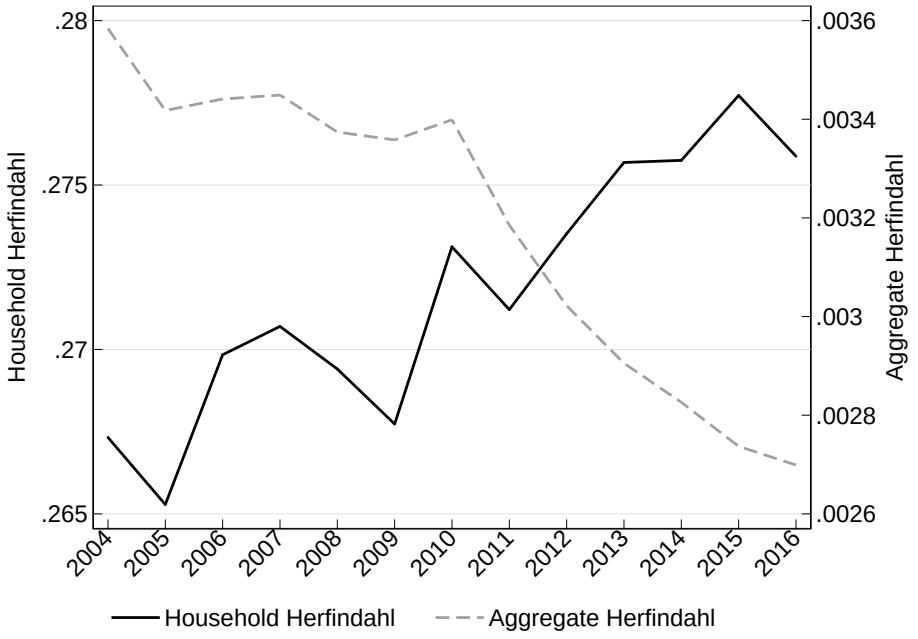



**Figure A3: Concentration Trends: Brand Instead of UPC**

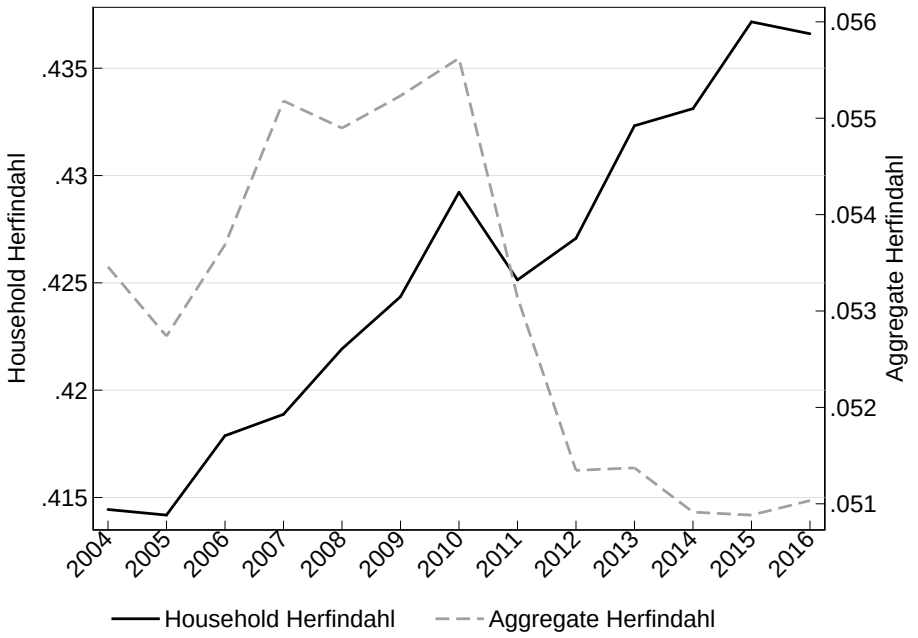


**Figure A4: Concentration Trends: Product Module instead of Group**

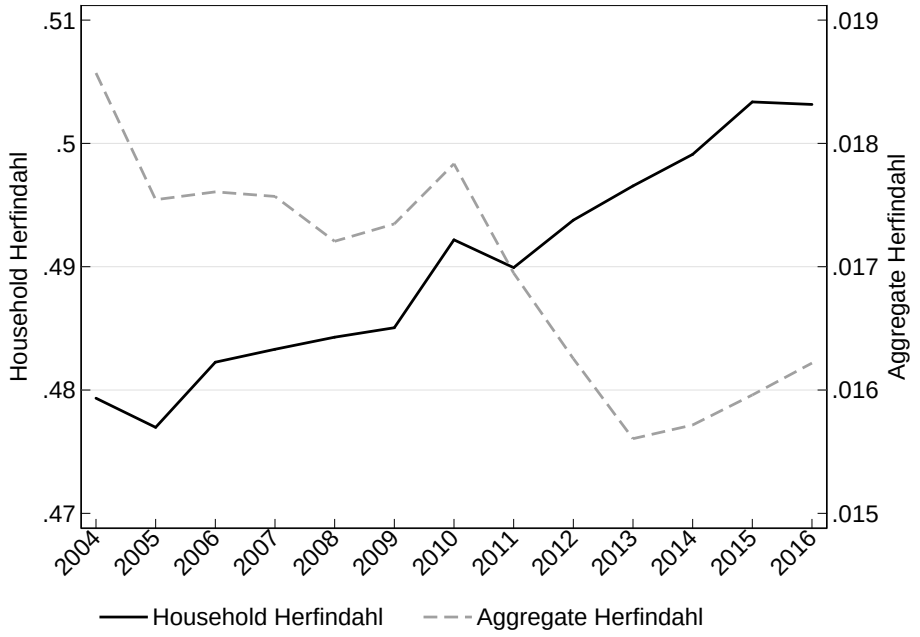
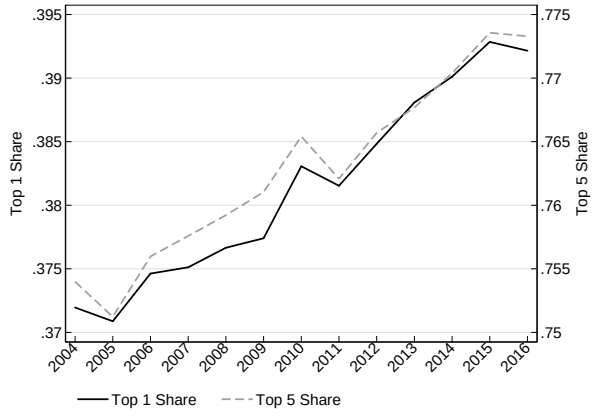



**Figure A5: Alternative Concentration Measures**

**(a):** Household Shares on Top Products




**(b):** Aggregate Shares on Top Products

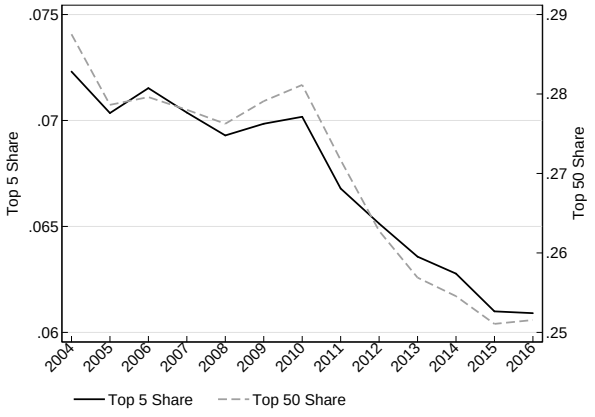
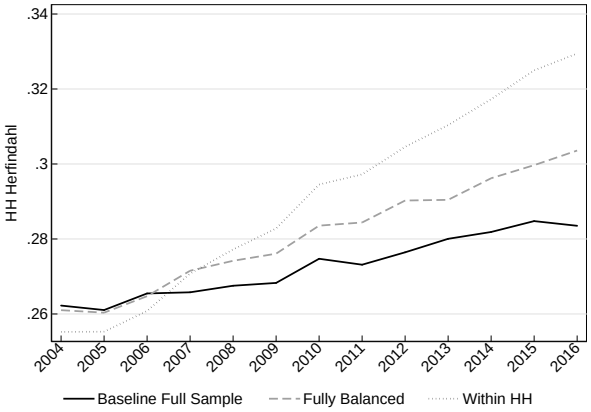


**Figure A6: Concentration Trends for Different Samples**

**(a):** Household Herfindahl




**(b):** Aggregate Herfindahl

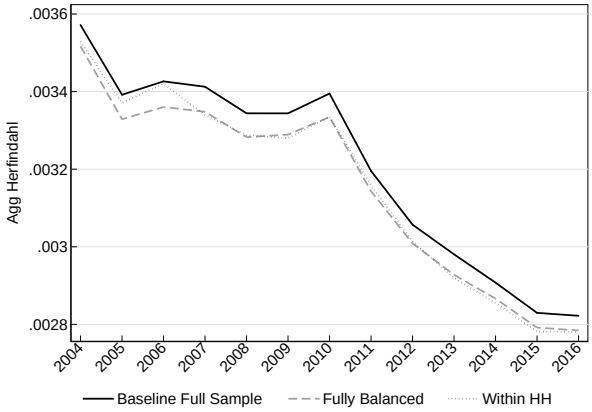
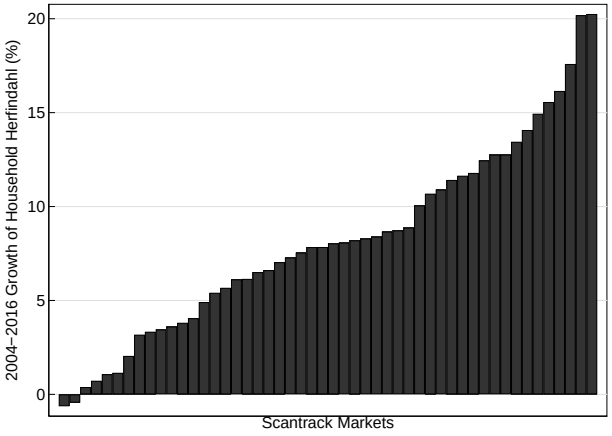



**Figure A7: 2004-2016 Concentration Growth Within Location**

**(a):** Household Concentration




**(b):** Aggregate Concentration

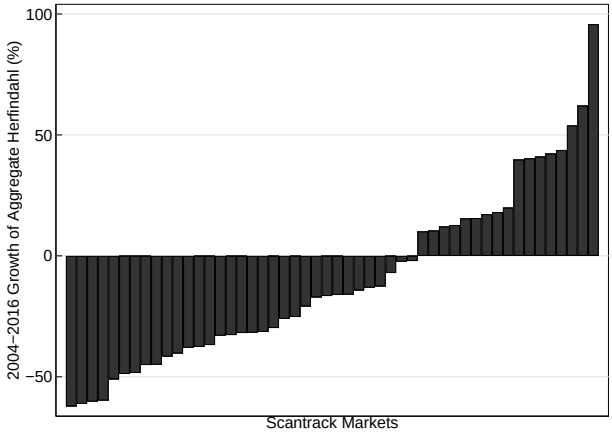
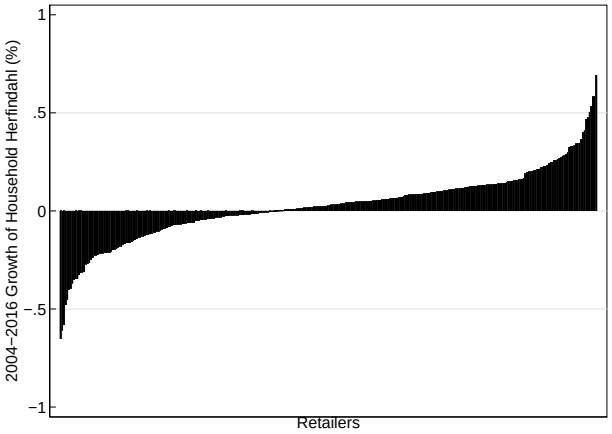


**Figure A8: 2004-2016 Concentration Growth Within Retailer**

**(a):** Household Concentration




**(b):** Aggregate Concentration

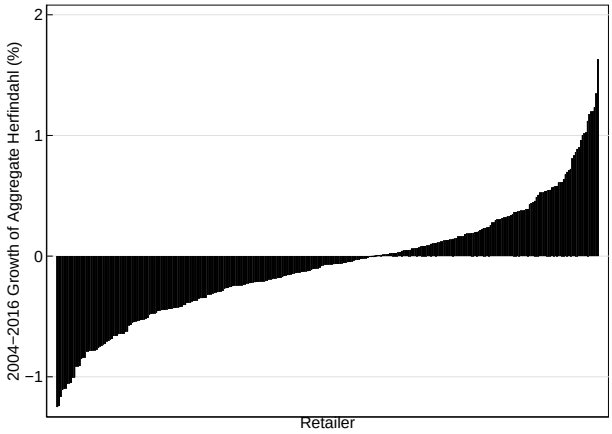
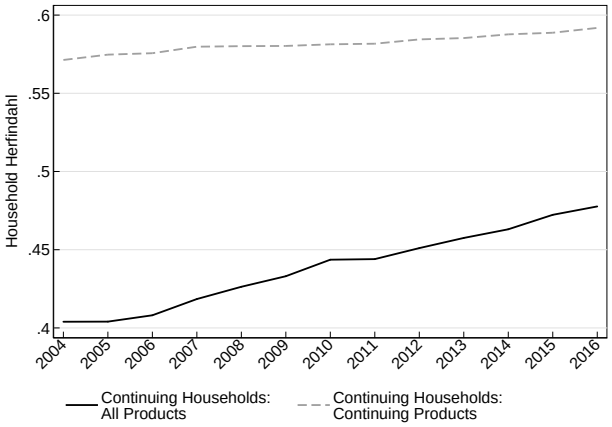



**Figure A9: 2004-2016 Concentration growth for continuing vs. all products (brands)**

**(a):** Household Herfindahl




**(b):** Aggregate Herfindahl

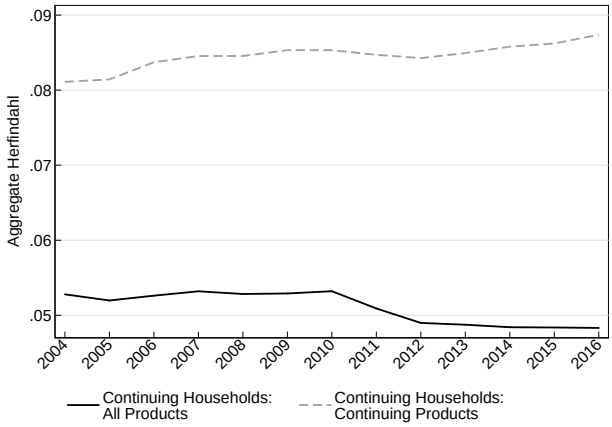


**Figure A10: Intensive Margin P v. Q effects for UPCs**

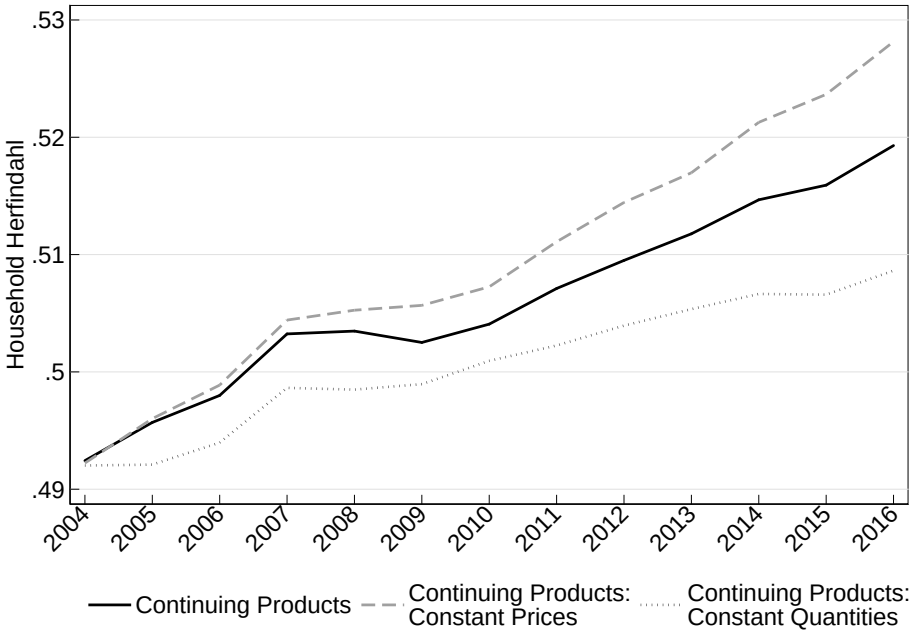


**Table A1:** Effect of Demographics on Household Concentration Trends

| | (1) HH herf | (2) HH herf | (3) HH herf | (4) HH herf | (5) HH herf | (6) HH herf |
|---|---|---|---|---|---|---|
| t | 0.0026*** | 0.0025*** | 0.0024*** | 0.0024*** | 0.0034*** | 0.0035*** |
| | (0.00038) | (0.00037) | (0.00037) | (0.00037) | (0.00035) | (0.00035) |
| household size 2 | | -0.045*** | -0.044*** | -0.043*** | -0.032*** | -0.031*** |
| | | (0.0029) | (0.0029) | (0.0031) | (0.0029) | (0.0028) |
| household size 3 | | -0.066*** | -0.062*** | -0.061*** | -0.048*** | -0.047*** |
| | | (0.0050) | (0.0043) | (0.0046) | (0.0043) | (0.0043) |
| household size 4 | | -0.076*** | -0.070*** | -0.069*** | -0.051*** | -0.050*** |
| | | (0.0062) | (0.0051) | (0.0054) | (0.0047) | (0.0046) |
| household size 5 | | -0.083*** | -0.077*** | -0.076*** | -0.054*** | -0.053*** |
| | | (0.0068) | (0.0057) | (0.0061) | (0.0058) | (0.0057) |
| age 30-34 | | | -0.0020 | -0.0018 | 0.00066 | 0.00034 |
| | | | (0.0013) | (0.0012) | (0.0013) | (0.0013) |
| age 35-39 | | | -0.0033* | -0.0032* | 0.0030* | 0.0031* |
| | | | (0.0019) | (0.0019) | (0.0016) | (0.0016) |
| age 40-44 | | | -0.0017 | -0.0016 | 0.0093*** | 0.0098*** |
| | | | (0.0029) | (0.0030) | (0.0024) | (0.0023) |
| age 45-49 | | | -0.000069 | 0.000031 | 0.014*** | 0.015*** |
| | | | (0.0037) | (0.0039) | (0.0033) | (0.0030) |
| age 50-54 | | | 0.0035 | 0.0036 | 0.018*** | 0.019*** |
| | | | (0.0046) | (0.0048) | (0.0041) | (0.0038) |
| age 55-59 | | | 0.0070 | 0.0071 | 0.023*** | 0.024*** |
| | | | (0.0052) | (0.0054) | (0.0046) | (0.0042) |
| age 60-64 | | | 0.0087 | 0.0087 | 0.027*** | 0.028*** |
| | | | (0.0060) | (0.0061) | (0.0049) | (0.0046) |
| age 65-69 | | | 0.0078 | 0.0078 | 0.025*** | 0.026*** |
| | | | (0.0062) | (0.0063) | (0.0049) | (0.0045) |
| age 70-74 | | | 0.017** | 0.017** | 0.031*** | 0.031*** |
| | | | (0.0070) | (0.0070) | (0.0057) | (0.0053) |
| Income $20000-34999 | | | | -0.0021 | -0.00036 | -0.00078 |
| | | | | (0.0023) | (0.0021) | (0.0018) |
| Income $35000-59999 | | | | -0.0034 | 0.0014 | 0.000016 |
| | | | | (0.0035) | (0.0034) | (0.0029) |
| Income >$60000 | | | | -0.0026 | 0.0071 | 0.0024 |
| | | | | (0.0052) | (0.0052) | (0.0041) |
| log *spend* | | | | | -0.065*** | -0.065*** |
| | | | | | (0.0055) | (0.0055) |
| R-squared | 0.31 | 0.32 | 0.32 | 0.32 | 0.38 | 0.38 |
| N | 49988358 | 49988358 | 49988358 | 49988358 | 49988358 | 49987989 |

**Note:** Table shows results from a regression of household herfindahls on various demographic variables and a time trend. Omitted categorical variables are household size 1, age<=29, and income<$20000. In column (5) additional controls are: dummy variables for education, employment status, occupation, scantrack markets, marital status, type of residence, race, presence of children, presence of household internet, cable/non-cable tv, and various indicators for the presence of major kitchen appliances. The unit of observation is a household-year, and observations are weighted using Nielsen sampling weights. Standard errors shown in parantheses are clustered by household. Significance levels: * ($p<0.10$), ** ($p<0.05$), *** ($p<0.01$).

## Appendix B. Detailed Data Description

Our primary data set is the AC Nielsen Homescan data, which we use to measure household-level shopping behavior.[1] As discussed in the text, our panel contains weekly household-level product spending 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. Roughly half of expenditures are in grocery stores, a third of expenditures are in discount/warehouse club stores, and the remaining expenditures are split among smaller categories such as pet stores, liquor stores, and electronics stores.

While panelists are not paid, Nielsen provides incentives such as sweepstakes to elicit accurate reporting and reduce panel attrition. Projection weights are provided to make the sample representative of the overall U.S. population.[2] A broad set of demographic information is collected, including age, education, employment, marital status, and type of residence. Nielsen maintains a purchasing threshold that must be met over a 12-month period in order to eliminate households that report only a small fraction of their expenditures. The annual attrition rate of panelists is roughly 20 percent, and new households are regularly added to the sample to replace exiting households.

Households report detailed information about their shopping trips using a barcode scanning device provided by Nielsen. After a shopping trip, households enter information including the date and store location and scan the barcodes of all purchased items. 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, and "fabric softeners-liquid" is a typical product module within the "laundry supplies" product group within the "non-food grocery" department.

In our baseline analysis, we define a product as a UPC. UPCs are directly assigned by the manufacturer and will typically change any time there is any change in product characteristics. However, we also compute results instead defining a product as a "brand". Information on brands is constructed by Kilts/Nielsen and is more aggregated than UPCs but still very disaggregated: for example, "Pepsi" and "Caffeine Free - Pepsi" are two different brands, as are "Pepsi" and "Mountain Dew", despite the latter being produced by the same parent company. However, different flavors of Pepsi are typically all listed under the same Pepsi brand. We focus on UPCs as our baseline product definition for several reasons: 1) Most importantly, UPCs are directly assigned by the manufacturer, while the brand variable is constructed by Kilts/Nielsen. Which UPCs are grouped into more aggregate brands involves some subjective judgment, and this aggregation is not necessarily consistent across categories or time.

[1] 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.

[2] We use these projection weights in all reported results, but our results are similar when weighting households equally.

2) UPCs are the most fine-grained definition available and will capture relevant product changes like the introduction of new flavors which will typically not be captured with the brand-definition. 3) In order to preserve anonymity of the stores in the Nielsen sample, all generic UPCs are assigned the same brand code. This means that analysis of brand-level spending can only be done on the subset of name-brand products and must exclude the large and growing share of generic products from the sample. (see e.g. Dube, Hitsch, and Rossi (2018)).

However, there is legitimate concern that UPCs may be too fine a notion of product when considering the concentration of household purchases, since households may view certain UPCs (for example minor differences in size or packaging for otherwise equivalent UPCs) as identical products.[3] For this reason, we show robustness to instead defining a product as a brand rather than a UPC.

Our baseline analysis focuses on annual spending and computes household market shares across products within product groups, but all results are robust to calculating household product market shares in more disaggregated product modules or more aggregated department codes. There is substantial heterogeneity across product modules in the degree of household concentration, so our analysis focuses on a set of balanced product modules. This eliminates spurious changes in concentration which might otherwise arise from changes in the set of goods sampled by Nielsen (which do not represent real changes in household's actual consumption and instead merely changes in the categories of consumption reported in Nielsen). This focus on balanced product modules reduces our sample from 118 to 107 product groups. Our analysis excludes fresh produce and other "magnet" items without barcodes since products in these categories cannot be uniquely identified and products with identical product codes in these categories can potentially differ substantially in quality. Our baseline sample includes all households and weights each household using sampling weights provided by Nielsen which are designed to make the Nielsen demographically representative of the broader U.S. population. Appendix Figure A12 shows that aggregate spending growth in our sample tracks government data on aggregate spending growth in comparable categories. Our conclusions are even stronger when instead using a balanced panel of households to eliminate household composition changes.

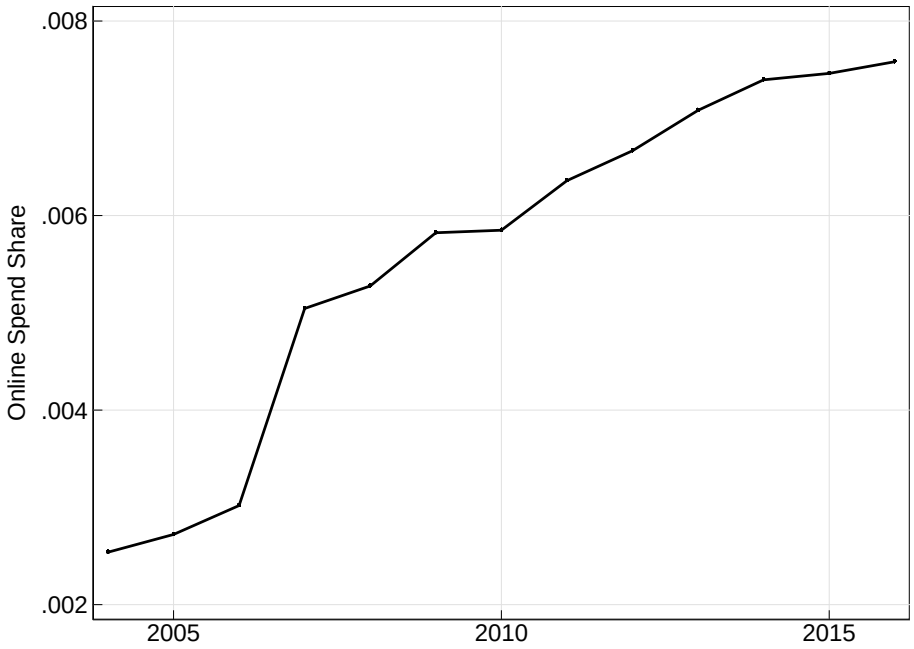
While our baseline sample includes all UPCs, we also show that our results hold when excluding generic/private-label products. In order to preserve anonymity of the stores in the Nielsen sample, the exact identity of generic brands in the Nielsen data is masked. There has been an increase in the private label share of all purchases over the last decade (see e.g. Dube, Hitsch, and Rossi (2018)) so including generic spending which cannot be properly allocated to constituent UPCs might lead to spurious concentration trends. However, we show that excluding generics and calculating concentration trends for branded products produces nearly identical results.

Finally, it is also useful to discuss the potential role of online shopping for our measurement.

[3]It is not clear that we want to classify a switch from spending $10 on Brand-X 64 oz laundry detergent and $10 Brand-X 60 oz laundry detergent to instead spending $20 on Brand-X 64 oz laundry detergent as a large increase in concentration. If UPCs become more homogeneous across time, using UPCs as our notion of product may lead to spurious changes in concentration with no substantive change in household behavior.

Households in the Nielsen Homescan sample are supposed to scan barcoded purchases of purchases from online retailers in addition to the items they scan from brick-and-mortar retailers. Indeed the Nielsen panel shows a growing share of online spending across time (Figure A11). However, for the categories covered in Nielsen data, online spending is relatively unimportant, so even by the end of the sample these spending shares remain low.[4]. Breaking results out further for particular categories where online spending is likely to be more and less relevant delivers no obvious interaction with concentration trends. For these reasons, we conclude that online shopping is unlikely to be of direct importance for understanding the diverging trends that we document.

**Figure A11: Online Spending Shares**



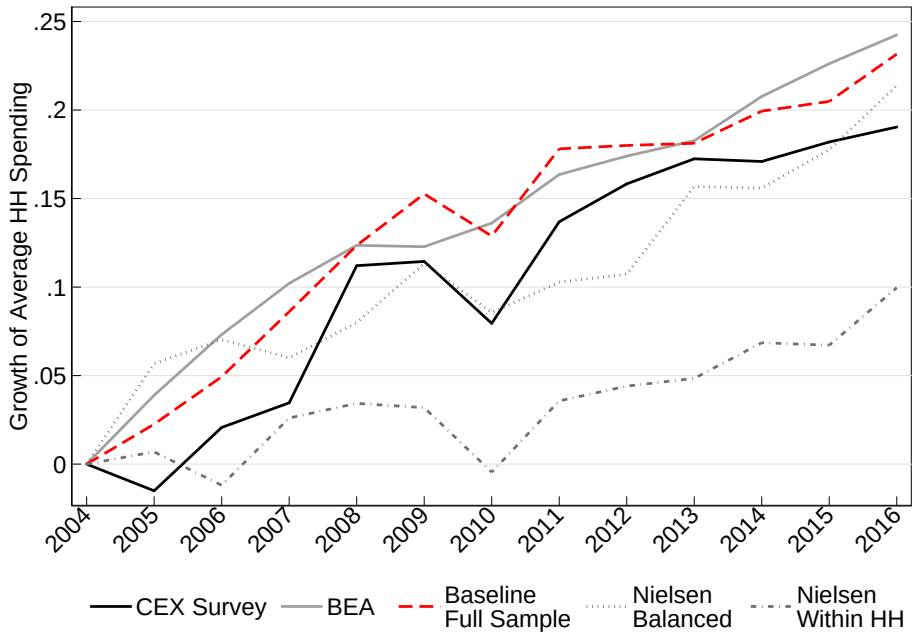

[4]Online vs. brick-and-mortar spending is classified at the level of the retail chain. This means that our measure captures spending at online only retailers such as Amazon but misses online spending associated with traditional retailers such as spending at Walmart.com.

# Appendix C. Relationship to External Data

## C.1 External Spending Data

Figure A12 shows that aggregate Nielsen spending lines up well with spending growth measures from the Consumer Expenditure Survey and BEA national accounts for similar categories.[5]

**Figure A12: Household Spending in Nielsen vs. Consumer Expenditure Survey**




However within-household spending growth is substantially less strong than overall household spending. This is likely driven by two forces: 1) The panel dimension of Nielsen is not representative of all households. The continuing households in the sample are substantially older than the overall Nielsen sample and the overall population, and we know from other research that households around retirement have declining food spending. While Nielsen provides sampling weights to make the overall sample representative of the U.S., they do not provide weights to make the panel dimension representative of the overall US, and the requisite demographic variables in the data to construct them ourselves do not exist. 2) There is likely attrition bias and households probably report a declining share of spending across time. This attrition bias may be particularly strong in the final year in which a household is in the sample, which could potentially explain the difference between the fully balanced

[5]It is well-known that the consumer expenditure captures a lower level of spending than the BEA and this "missing spending" has a positive trend. However this growth in missing spending mostly occurs prior to our sample period. Throughout our sample period, the CEX captures a relatively constant share of aggregate spending. This means CEX spending growth is slightly lower but broadly similar to aggregate spending growth from the BEA.

and within-household spending growth patterns. If reduced reporting tends to proceed exit, then one would expect attrition bias to be less severe for households who remain in the sample for the full 12 years. Consistent with this, the balanced sample exhibits stronger spending growth than the within household sample.

For these reasons, our baseline results use the entire Nielsen homescan panel rather than focusing on a balanced panel of households. However, it is useful to compare our basic trends in the full sample to those computed using within-household variation. Figure A6 thus redoes Figure 1 using a fully balanced panel and with a specification using only the within-household changes specification.

Clearly trends are even stronger than our baseline results when using the fully balanced panel or when identifying off of within-household variation, so in this sense our baseline is conservative. We now describe several forces that might spuriously increase the within-household trend as well as some alternative forces which might spuriously flatten the full sample trend. This makes it difficult to know whether our baseline sample is likely to be understated or whether it is instead the balanced panel specification that is overstated. However, in either case, the trend is robustly positive, and our baseline sample is the one which generates more conservative results.

More specifically, the full sample trend could potentially be biased downwards because the Nielsen sampling technology changes across time, and these changes are implemented when households enter the sample. These changes in technology could obscure underlying trends in the data, but would be stripped out when using within-household variation. More generally, households have very different concentration levels, as shown above, so that random household entry and exit in the sample could make it more difficult to pick up underlying trends. These are both forces that might lead our baseline full sample to understate the true increase in concentration across time.

Conversely, we have shown above both that increases in spending are strongly negatively correlated with increases in concentration and that the within-household sample has spending growth much lower than in the consume expenditure survey. To the extent that the within-household sample has spurious declining spending due to sample attrition, there is then a concern that using within household variation might lead to an upward biased trend. However, if we redo all our regression results using within household variation *controlling* for within household changes in spending, we continue to find upward trends which are stronger than in the full sample. This suggests that the stronger upward trend in the within-household results is not driven solely by the lower reported spending growth in this sample. In addition, we can also recompute results using only households in the first year in the sample. By construction, attrition bias in spending across time cannot drive any trend, since this sample has no within-household time-series variation but it still delivers an upward trend. Finally, attrition bias is less likely to be a concern for the fully balanced sample: The upward trend in the fully balanced panel is roughly linear across time, so if this upward trend was explained by attrition bias and progressive under reporting, this under reporting would need to grow at a constant

rate, which seems unlikely, especially because Nielsen tries to drop households from the sample who are not reporting accurately. It seems much more likely that the biggest under reporting would occur in the first year or two in the panel as households are likely to be most enthusiastic about scanning purchases initially and then reduce scanning as it becomes more tedious across time. It would be quite surprising if enthusiasm waned at a constant linear rate across time but that households continued to participate in the homescan panel.

Together, we think that these results suggest the stronger upward trends using the balanced samples and the within-sample variation are not driven by spurious attrition bias. Nevertheless, we cannot fully rule out this concern. Furthermore, as discussed above the panel element of the sample is not representative since households who remain in the sample for progressive years are demographically different and not representative of the population leading total spending for this population to line up less well with aggregate spending inferred from the consumer expenditure survey. For these reasons and to be conservative, we focus on the full sample in all our baseline results but only note here that using other samples only strengthens our conclusions.

### C.2 Census Concentration of Production

A large and growing literature uses production data from the Census to show that the concentration of production has been broadly increasing from 1982-2012. For example, Autor, Dorn, Katz, Patterson, and Reenen (2017) calculates industry concentration within 4-digit industries, and averages this up to 6 major sectors and shows that various concentration measures have all increased when comparing 1982 to 2012. In this section we explore the relationship between the concentration measures in our paper and this large literature and argue that relevant comparisons from Nielsen data are broadly consistent with this Census based literature.

First, it is important to note that the concentration notions we emphasize in our paper are conceptually distinct along a number of important dimensions from the concentration of firms or establishments studied using census data. Most importantly, we are measuring the concentration of spending over very detailed UPCs (or slightly coarser but still highly disaggregated brands). This is a fundamentally much more disaggregated notion of concentration than that studied with production data, since firms can potentially produce tens, hundreds or even thousands of different products. For example, in our data Procter and Gamble produces over 40,000 unique UPCs, L'Oreal produces over 28,000 UPCs and General Mills, Unilever, and Kraft Heinz all produce 10,000-20,000 UPCs.[6]

Furthermore, the categories within which we calculate concentration are also more disaggregated than those in typical Census-based calculations and also cover a more narrow subset of production. For example, the broad manufacturing sector in Autor, Dorn, Katz, Patterson, and Reenen (2017) covers 86 4-digit industries within which concentration is computed. However, of these 86 industries

[6]It is also worth noting that our "household" concentration measures have no analogue in the Census literature even if we were measuring producer rather than product concentration.

only a small subset produce in categories which are covered by Nielsen (for example NAICS Code 3111 "Animal Food Manufacturing") while most are in production industries which have no overlap with Nielsen categories (for example NAICS Code 3336 "Engine, turbine, and power transmission equipment manufacturing" or NAICS Code 3365 "Railroad rolling stock manufacturing").

Finally, it is important to note that our sample covers the period 2004-2016 while census data starts in 1982 and is last available in 2012. The exact timing of concentration trends in Census data varies substantially, with many sectors exhibiting increases primarily in the period prior to our sample period.

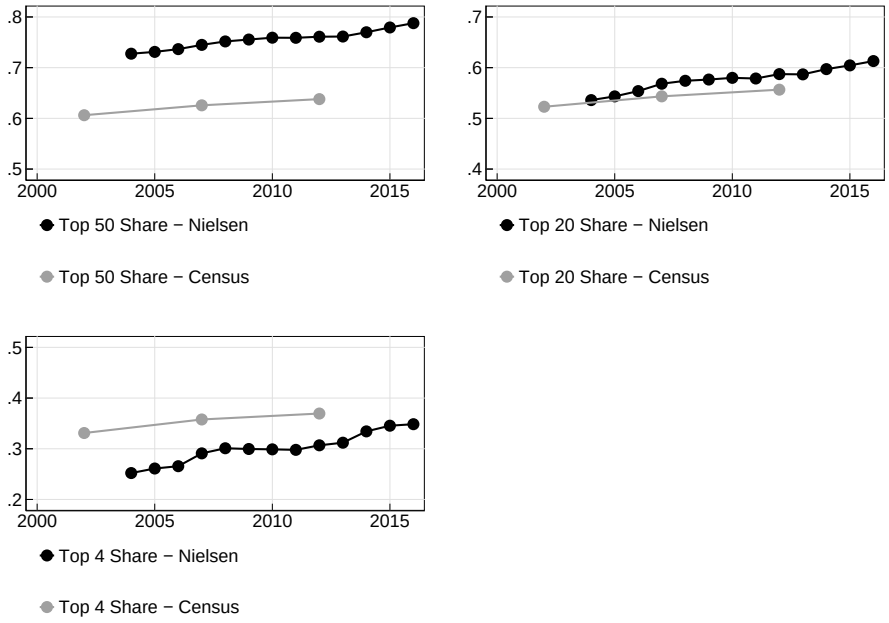
Since they are conceptually different notions, this means the aggregate product concentration trends which we emphasize in the body of the paper should not be directly compared to production concentration trends in Census. However, we can construct concentration measures using the Nielsen data which *are* more comparable with Census calculations and that can be used to explore the external validity of our data. We now explore these comparisons.

Since households in the Nielsen sample report the retail chain in which they shop, we can aggregate up total spending to compute a Nielsen based measure of spending at each retail chain and resulting retailer concentration. This can then be compared to the concentration of retail trade in Census data. Specifically, since the Nielsen sample is focused on grocery and drug store spending, in the Census we use firm concentration numbers only from NAICS Code 445 "Food and beverage stores" and 446 "Health and personal care stores" and weight the publicly available Census concentration numbers for these two sectors using their relative share of sales. This clearly does not provide a precise match between the retail establishments covered in Nielsen and Census so we should not expect numbers to line up exactly, but Figure A13 shows that that Nielsen data broadly matches the level of retail spending accounted for by the Top 4, Top 20 and Top 50 firms as well as the upward trend in retail concentration.

We can also perform a similar exercise by allocating UPC-level spending up to the manufacturer. When manufacturers produce a new product, it is assigned a barcode by the company GS1, which then maintains a database which can be used to link UPCs to manufacturers. This lets us aggregate product spending up to a measure of manufacturer spending, with two important caveats:

First, the link from UPCs to parent companies is sometimes inconsistent. For example, Gillette and Old Spice were both acquired in the past by Proctor and Gamble, and the UPCs for Gillette and Old Spice products both map to Proctor and Gamble. However, Ben and Jerry's was acquired by Unilever in 2000, yet UPCs for these products are assigned to the "Ben and Jerry's Homemade Inc" firm name rather than to the Unilever parent company. Similarly, Goose Island Beer UPCs are assigned to "Goose Island Beer Company" even though this firm was acquired by InBev in 2011. To the extent that some UPCs are assigned to subsidiaries rather than parent companies, our Nielsen based measure of manufacturer concentration will be biased downwards.

**Figure A13: Retail Trade Concentration**

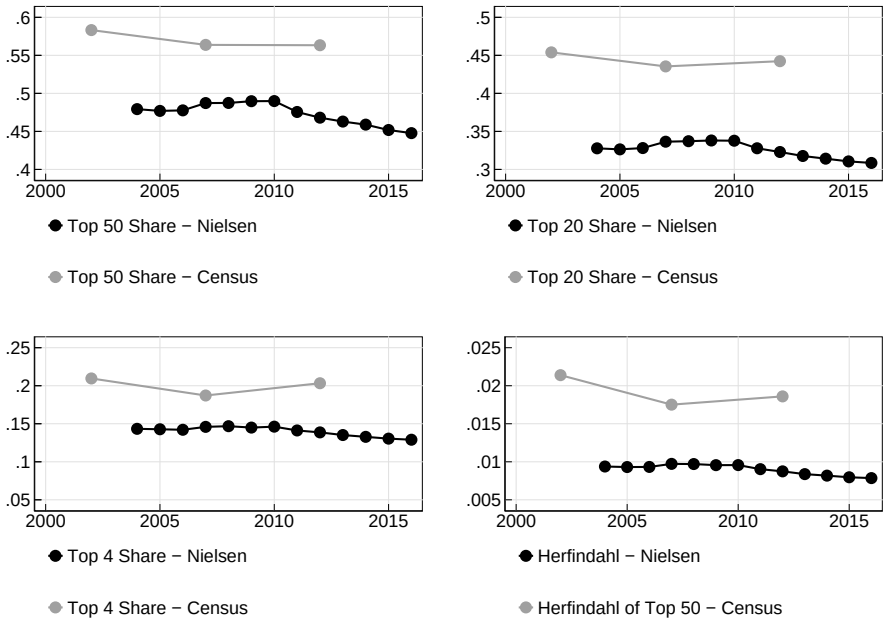



Second, UPCs for store-brand products map to the retailer rather than the actual manufacturer of the product. For example, Costco's "Kirkland" store-brand barcodes all map to "Costco", even though Costco does not actually produce most of these products. Although sometimes the actual producer can be identified (for example Kirkland Coffees are advertised as being roasted by Starbucks), this information is typically a trade-secret. This means that we cannot measure the producer for most generic products, and as a result we must drop these products when aggregating up UPCs to manufacturers and focus only on branded products. To the extent that the production of generic products is proportional to the production of branded products, this will have no effect on concentration. However, it is likely that generic products are disproportionately produced by larger manufacturers, so dropping generic products is likely a second force that will bias our Nielsen based measures of manufacturer concentration downwards.

To again focus the comparisons on the most relevant producers, we keep NAICS codes 311 and 312 "Food Manufacturing" and "Beverage and Tobacco Product Manufacturing" from the Census data and weight these concentration measures by their relative sales shares. Figure A14 shows that despite the above concerns, Nielsen data again broadly matches Census data, producing similar levels of manufacturer concentration and a flat to mild downward trend.

Overall the results in these two subsections give us confidence that the Nielsen data is largely in line with external evidence on aggregate spending and with Census data on producer concentration.

**Figure A14: Manufacturer Concentration**



## Appendix D. Robustness of Model Inference to Alternative Trends

In this appendix, we show that the need for increasing $\tilde{\alpha}$ to fit the rise in niche product consumption is very robust and does not depend importantly on the exact strength of this phenomenon in the data.

Given data on $\Omega$, $H^{HH}$, and $HH^{Agg}$, one can immediately solve for $\tilde{\alpha}$ at each date. Doing so implies that in order to rationalize the data, $\tilde{\alpha}$ must rise by 67.4%. Thus, when viewed through the lens of our model, the data can only be rationalized with a large increase in $\tilde{\alpha}$.

But how strong/robust is this conclusion? For example, if we had slightly different trends (or potentially some measurement error in trends), could we have reached a different conclusion? To show that the need for increasing $\tilde{\alpha}$ is a very robust qualitative conclusion that is not particularly sensitive to the exact empirical trends, suppose that we knew only that $\Delta H^{HH} \geq 0$ and $\Delta\Omega \leq 0$, as in our data, but we knew nothing about the strength of the trend. What decline in $HH^{Agg}$, if any, could be rationalized in a constant $\tilde{\alpha}$ environment if we only knew this weaker condition?

**Theorem 1** *Assume that $\Delta H^{HH} \geq 0$ and $\Delta\Omega \leq 0$. Then given initial $0 \leq \eta_0 \leq 1$, the maximum possible percentage decline in the aggregate Herfindahl without an increase in $\tilde{\alpha}$ is* $\max\left(0, \frac{2\eta_0+1}{(\eta_0)^{\frac{1}{2}}}\frac{1}{3} - 1\right)$. *This is equal to 0 for $\eta_0 < 1/4$ and the minimum over all $\eta_0$ is $2/3\sqrt{2} - 1 \approx -5.72\%$.*

**Proof.** We now compute the maximum possible percentage decline in $HH^{Agg}$ that can be achieved without increasing $\tilde{\alpha}$.

First, note that the maximum decline will be obtained when $\Delta H^{HH} = 0$, since if $\Delta H^{HH}$ were strictly positive, one could always find a different pair of $\eta$ and $\Omega$ that satisfy $\Delta H^{HH} \geq 0$ and $\Delta\Omega \leq 0$ but with lower $HH^{Agg}$. This means that we can solve for $\Omega$ as a function of $\eta$ and the initial household herfindahl: $H_0^{HH}$:

$$\Omega = \frac{(\eta+1)^2}{4\eta}\frac{1}{H_0^{HH}}. \tag{A1}$$

This relationship must hold for any possible pair of $\Omega$ and $\eta$. We label the initial values as $\Omega_0$ and $\eta_0$, and the new values as $\Omega_1$ and $\eta_1$

Since this is a declining function of $\eta$, $\Omega_1 \leq \Omega_0 \implies \eta_1 \geq \eta_0$.

We thus simply want to compute the minimum value of $HH_1^{Agg}/HH_0^{Agg}$ subject to $\eta_1 \geq \eta_0$ and A1.

$$\frac{HH_1^{Agg}}{HH_0^{Agg}} = \frac{\frac{(\eta_1+1)}{2\eta_1+1}}{\frac{(\eta_0+1)}{2\eta_0+1}}\left(\frac{\Omega_0}{\Omega_1}\right)^{\frac{1}{2}} \tag{A2}$$

$$= \frac{\frac{(\eta_1+1)}{2\eta_1+1}}{\frac{(\eta_0+1)}{2\eta_0+1}}\left(\frac{\frac{(\eta_0+1)^2}{4\eta_0}}{\frac{(\eta_1+1)^2}{4\eta_1}}\right)^{\frac{1}{2}} \tag{A3}$$

$$= \frac{2\eta_0+1}{(\eta_0)^{\frac{1}{2}}}\frac{(\eta_1)^{\frac{1}{2}}}{2\eta_1+1}. \tag{A4}$$

Since we take $\eta_0$ as given, the minimization problem is then solved by minimizing

$$\frac{(\eta_1)^{\frac{1}{2}}}{2\eta_1+1} \text{ subject to } \eta_1 \geq \eta_0 \tag{A5}$$

$\frac{(\eta_1)^{\frac{1}{2}}}{2\eta_1+1}$ is increasing with $\eta_1$ for $\eta_1 < \frac{1}{2}$ and decreasing for $\eta_1 > \frac{1}{2}$. This implies that $\eta_1 = 1$ will be a local minimum. There will be a second local minimum $\eta_1 = \eta_0$ if $\eta_0 < 1/2$. This means that for $\eta_0 > 1/2$ we know that the minimum is achieved by setting $\eta_1 = 1$, and

$$\frac{HH_1^{Agg}}{HH_0^{Agg}} = \frac{2\eta_0+1}{(\eta_0)^{\frac{1}{2}}}\frac{1}{3}. \tag{A6}$$

For $\eta_0 < 1/2$ the other local minimum implies that

$$\frac{HH_1^{Agg}}{HH_0^{Agg}} = 1, \tag{A7}$$

so we just need to solve for where

$$1 \leq \frac{2\eta_0+1}{(\eta_0)^{\frac{1}{2}}}\frac{1}{3}, \tag{A8}$$

Solving this equation implies that for it is satisfied if $\eta_0 < 1/4$. What does this imply? It means for $\eta_0 < 1/4$: $\Delta H^{HH} \geq 0$ and $\Delta\Omega \leq 0$ are inconsistent with *any* decline in $HH^{Agg}$ when holding $\tilde{\alpha}$ fixed. Note that we can also write $eta_0$ directly in terms of initial period observables, which makes it easier to interpret. From the equation for the household herfindahl, $\eta_0 < 1/4$ implies that

$$H_0^{HH}\Omega_0 \geq \frac{25}{16}. \tag{A9}$$

So if $H_0^{HH}\Omega_0$ satisfies this condition, it is not possible to increase household herfindahl, decrease aggregate herfindahl and household varieties without raising $\tilde{\alpha}$. For our data, $H_0^{HH}\Omega_0 = 4.24$, so we easily satisfy this condition and it would take very different empirical values (a roughly 300% lower $H_0^{HH}$ or $\Omega_0$) to violate this constraint and even be in a region of the parameter space where it's possible to get a decline in aggregate herfindahls. But even if we are in the region $\eta_0 > 1/4$, it is possible for

$HH^{Agg}$ to decline, but we have a strict bound on the maximum possible decline. In particular, the max possible percentage decline for a given $\eta_0$ is simply given by

$$\max\left(0, \frac{2\eta_0 + 1}{(\eta_0)^{\frac{1}{2}}}\frac{1}{3} - 1\right). \tag{A10}$$

This is minimized at $\eta_0 = 1/2$ and implies a max possible decline across all possible parameters of $2/3\sqrt{2} - 1 \approx -5.72\%$, which is much smaller than the observed percentage decline of -19.56% (and $\eta_0$ is totally inconsistent with the micro data; but even if willing to go to that unrealistic parameter, then still can't make things work without $\tilde{\alpha}$).

Finally, note that we get these bounds when imposing non-negative growth in the household herfindahl. If we used the actual value imposed in the data, we would get even stronger bounds on the feasible decline in aggregate herfindahl without a change in $\tilde{\alpha}$. So it is essentially impossible to get declines in the aggregate herfindahl when household concentration is rising and household varieties are falling, without an increase in $\tilde{\alpha}$.

■

## Appendix E. Model Simulation Results

In this section, we explore numerical simulations of our model to test the validity of our elasticity approximation as well as to explore how restrictive the assumption of a stable distribution of Pareto taste-adjusted prices is for our conclusion.

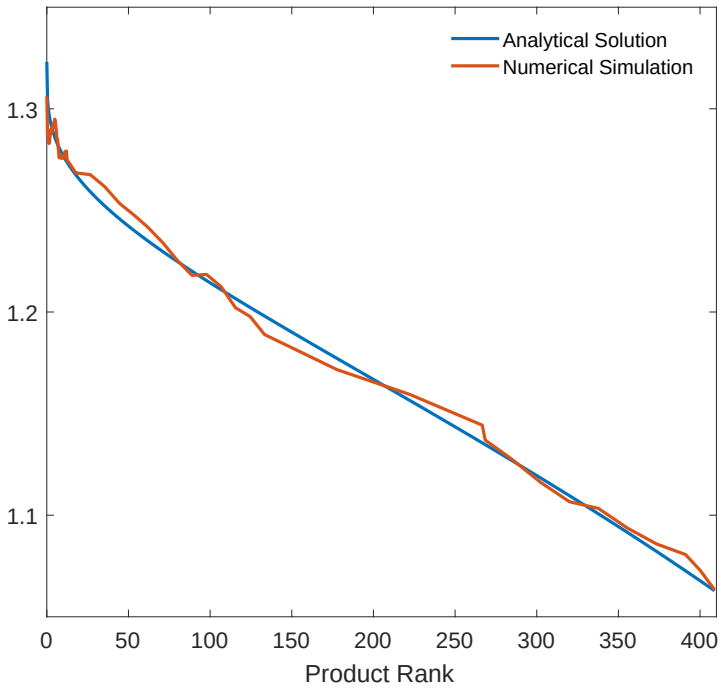
We simulate a discrete approximation to the main model in the paper by drawing a large random vector $\tilde{\gamma}^{rand}$ of price-adjusted tastes from a Pareto distribution for a large sample of households, using the same parameters as our baseline model. While $N = 15000$, our baseline random sample uses 2.25 million draws for each of 20,000 households since we are trying to approximate a continuum of products from [0,N]. However, rather than using analytical formulas to calculate household market shares, for each household we instead keep the random set $\Omega = \tilde{\gamma}^{rand} > \tilde{\gamma}^{*}$ and then calculate a numerical price index directly from $P = \left(\int_{k\in\Omega_i} (\tilde{\gamma}_{i,k})^{\sigma-1}\, dk\right)^{\frac{1}{1-\sigma}}$ and then compute market shares from Equation 11. These formulas hold for arbitrary distributions of taste, so even though we still simulate the taste draws from a Pareto distribution, in this simulation we are using no analytical results that rely on this assumption, which also means that we can also perform a similar procedure even if tastes do not follow a Pareto distribution.

In order to get aggregate market shares, we must identify the particular products that each household consumes. In order to do so, we use our rank function Equation 12 with a random uniform draw to compute for each household, the aggregate ranking of each of the 2.25 million possible products in [0,N] and then compute household $i's$ particular idiosyncratic rank for each of the 2.25 million $j$ products. We then sort $\tilde{\gamma}^{rand}$ and map the highest value to $r_{ij} = 0$, the second highest value to $r_{ij} = 1$ and so on to $r_{ij} = 2.25$ million. Finally, since for each $r_{ij}$, we know the value of $j$, this means that we then know household $i$'s taste draw and resulting individual spending for each aggregate product $j$. For example, the households highest $\tilde{\gamma}^{rand}$ draw will always map to their $r_{ij} = 0$, but the corresponding aggregate $j$ which household 1 ranks highest might be $j = 0$, the $j$ which household 2 ranks highest might be $j = 2043$, and the $j$ which household 3 ranks highest might be $j = 17$. Once we have these household specific spending shares for each product $j$, we can then numerically add up total spending on each product $j$ to calculate aggregate market shares.

Since these are computed entirely numerically, they do not rely on any of our closed form solutions for aggregate market shares and are thus again valid even under departures from the Pareto distribution. As we note in 3.7, our analytical market shares are only valid under the Pareto distribution so we must approximate the elasticity of demand by modeling a price change as a switch with another product in the aggregate ranking. Since these numerical results do not rely on the Pareto distribution, we can use this numerical model to simulate the aggregate elasticity of demand and resulting markup for a product $j$ by just raising all households' random taste draw for that product by a small amount. Note that calculating elasticities for each $j$ requires re-simulating a new set of aggregate market shares.

For these sample sizes, computing an elasticity for a single $j$ requires roughly 2 hours of computational time, so it is infeasible to simulate the elasticity of demand for all 2.5 million products. Instead, we compute the elasticity of demand and implied markups for 50 different values of $j$ distributed throughout the product space. Figure A15 compares this simulated markup to our analytical approximation and shows that the analytical approach produces essentially identical results (noting that there is still obvious numerical simulation error even with these large sample sizes).

**Figure A15: Simulated vs. Analytical Approximation for Markup**




As stressed throughout the paper, our analytical derivations and implications of changes in $N$ are only valid under the assumption that the distribution of price-adjusted tastes continues to follow a Pareto distribution as we vary $N$. If markups were fixed for all products, then assuming that the distribution of price-adjusted tastes is held fixed as $N$ varies would be a natural benchmark. However, our model instead implies that optimal markups do vary across products, and that the markups for individual products change as we vary $N$. This implies that if household tastes for products and their marginal costs held fixed, but we allow prices to change along with optimal markups when $N$ changes, then there will necessarily be a violation of the assumed Pareto distribution. Since all of our analytical results assume the Pareto distribution of price-adjusted tastes, this means that our analytical comparative statics to changes in $N$ and $F$ which induce changes in product markups are technically comparative statics in response to these parameter changes plus whatever implicit changes in household tastes (or marginal costs) are necessary to preserve a Pareto distribution of price-adjusted tastes after markups adjust. In practice, high turnover means that the set of products purchased in 2004 and

in 2016 is mostly disjoint, so one can primarily interpret these as taste shifts for new products rather than taste changes for existing products.[7] However, if the required taste shifts necessary to maintain the Pareto distribution under our counterfactuals were substantial, then this would potentially substantively change the interpretation of the welfare effects of changes in $N$.

However, we now use our numerical model to show that even though there are indeed implicit taste changes necessary to maintain the Pareto distribution as $N$ changes, in practice these required taste changes are quantitatively small and actually work against our conclusion that $N$ is welfare improving. We thus conclude that even though this is a large potential issue for the interpretation of our comparative statics, it is of little quantitative importance in practice. Specifically, we perform the following exercise: For the initial value of $N$ in 2004, we simulate our numerical model exactly as described above. Given household $i$'s resulting distribution of tastes for all $j$ products $\tilde{\gamma}_{i,j}^{rand}$, we can then compute a households actual (non-price adjusted) taste for product $j$ $\gamma_{i,j}^{rand} = \tilde{\gamma}_{i,j}^{rand}\mu_j$ using the analytical formula for $\mu_j$ from Section 3.7.[8] Note that as we explore above, even though our numerical model does not otherwise rely on analytical results, this analytical formula for the markup is valid since we are drawing the numerical distribution of price-adjusted tastes in the model from a Pareto distribution.

We then increase $N$ in the model but hold the particular random realizations of $\tilde{\gamma}_{i,j}^{rand}$ exactly fixed in the new simulation. Thus, by assumption, the values of *price-adjusted* tastes will be identical in the two simulations. However, as $N$ increases, the function $\mu_j$ and resulting prices will change. If price-adjusted tastes are fixed by assumption, but prices change then household tastes must change.[9].

How large are the required taste changes necessary to maintain an identical realization from a Pareto distribution of price-adjusted tastes as $N$ increases? Figure A16 shows that these changes are small. The left panel plots the implied taste draws as a function of initial aggregate product rank $j$ for a fixed household before and after a 70% increase in $N$.[10] Clearly the increase in $N$ induces some implied changes in tastes in order to maintain the Pareto distribution for price-adjusted tastes, but it is also clear that the requisite taste changes are small. The right panel of the plot shows a scatter plot of the realizations of taste before and after the increase in $N$. Overall the $R^2$ is above 0.999, so there is an almost perfect correlation of tastes under the two scenarios. In order to maintain an identical distribution of price-adjusted tastes, there is a modest *decline* in the implied average taste when $N$ increases, which lowers implied welfare by roughly 1.3%. This occurs because as $N$ increases, markups for incumbent products decline, which makes price fall and thus taste/price rise. In order to maintain a constant taste/price for that product, this means taste for those products must decline.

[7] Only 13.2% of UPCs purchased in 2004 are still purchased in 2016.

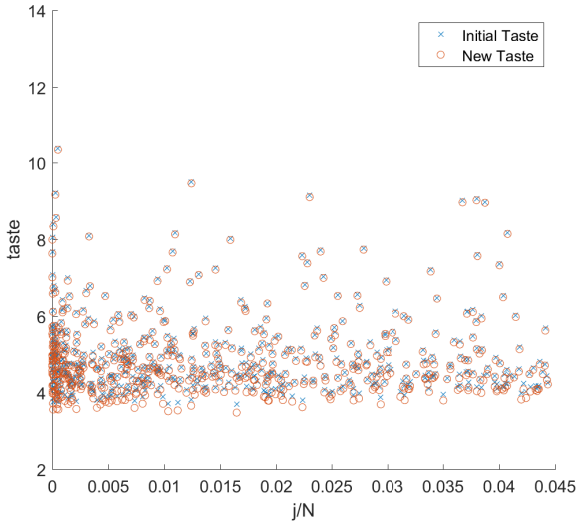
[8] For notational simplicity, we assume that marginal cost is 1 for all products. More generally this approach actually recovers the distribution of marginal cost adjusted tastes. As long as we assume marginal cost is constant as we vary $N$, one can interpret changes in taste and changes in marginal cost adjusted taste equivalently so these are equivalent exercises.

[9] Further, since markup changes are a monotonic function of $j$ but individual rankings of the $j$ products are non-monotonic when $\alpha > 1$, these price changes will be non-monotonic over individual households' consumption baskets.

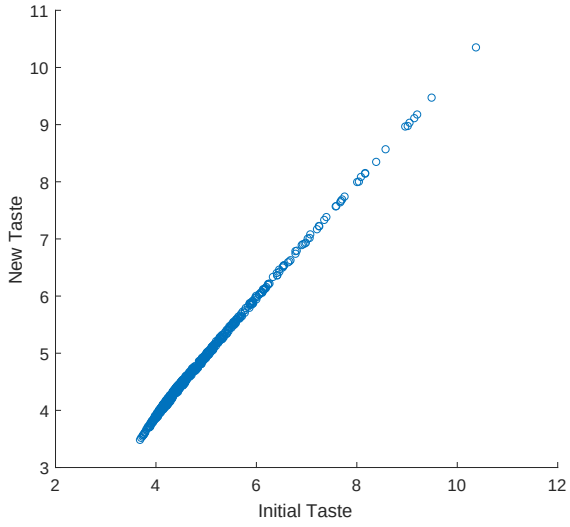
[10] Here we focus on products which are consumed in both scenarios so that such taste comparisons are relevant.

**Figure A16: Household Taste Changes Required to Maintain Pareto with N increase**

**(a):** Initial and New Tastes by Aggregate Product Rank




**(b):** Initial vs. New Tastes




However, the welfare conclusion in the body of the paper under the assumed constant Pareto distribution of price-adjusted tastes is that an increase in $N$ of 70% *raises* welfare by roughly 9.5%. The numerical results above show that that those welfare results are only valid if there is also a simultaneous modest decline in non-price adjusted tastes when $N$ rises, suggesting that if one instead held tastes fixed when increasing $N$ and departed from Pareto, the welfare increase would be slightly stronger. While such an exercise could potentially be performed numerically, it would require solving for the entire equilibrium distribution of the elasticity of demand and resulting markups numerically. As discussed above, the numerical calculation of the elasticity of demand (even for a single product in partial equilibrium) is very computationally costly.

Finally, we use this simulated model to also explore the role of potential measurement error in driving concentration trends. Although C.1 shows that the Nielsen data tracks aggregate spending measures fairly closely, the declining within-household spending patterns suggest there may be some role for attrition related measurement error across time. Furthermore, even though households are supposed to report online purchases and that Figure A11 shows that online spending is relatively unimportant for these sectors, it is possible that under-reported online spending might also drive increasing measurement error across time.

While it is difficult to analytically characterize the role of various forms of measurement error for concentration trends, we follow the indirect inference approach in Berger and Vavra (2015) and Berger and Vavra (2019) and *simulate* various flexible forms of measurement error in the numerical version of our model under the assumption that all other model parameters are held fixed. Specifically, we simulate the discrete version of our model and separately consider the effects of measurement error

on household and aggregate concentration. We focus primarily on measurement error arising from failing to report transactions entirely rather than from misreporting the size of a transaction, since the former is much more likely given the structure of the Homescan data collection. We consider three types of potential under-reporting encompassing various different extremes: 1) households failing to report some randomly chosen purchases, 2) households failing to report their smallest purchases and 3) households failing to report their largest transactions. Overall, we find that while measurement error can change both household and aggregate concentration, it pushes both household and aggregate concentration in the same direction and so is unlikely to be an important explanation for the observed rise in niche consumption. Unsurprisingly, the first and second form of measurement error raise both household and aggregate concentration while the third form of measurement error instead lowers both concentration measures. Furthermore, the second form of measurement error seems most plausible given the nature of the Nielsen data, since a household might fail to report a small one-off purchase which is likely to be a small share of that household's annual spending but is unlikely to consistently fail to report large, regular purchases that are likely to be a large share of annual spending. Since the third form of measurement error is especially unlikely, this means that measurement error is then also quite unlikely to explain a decline in aggregate concentration. As emphasized in Section D, a decline in aggregate concentration with flat household concentration would generally be sufficient to infer an increase in $N$. Overall, these simulation results strongly suggest that measurement error does not drive the rise of niche consumption.