Large Firms, Consumer Heterogeneity and the Rising Share of Profits*

Robert C. Feenstra† — UC Davis and NBER

Luca Macedoni‡ — Aarhus University

Mingzhi (Jimmy) Xu§ — Peking University

January 3, 2022

Abstract

We examine the relationship between large firms and the rising profit share in a model that features oligopolistic competition and consumer heterogeneity. Conditional on the sales distribution, the presence of consumer heterogeneity increases the profit share because it increases firm-level markups. Using data on purchases at the household-barcode level from Nielsen, we quantify the role of consumer heterogeneity, finding that the aggregate markup and the profit share are 8 and 3 percentage points larger than those predicted by a model of a representative consumer. Furthermore, we find that the profit share has been increasing over time and that firm targeting of consumer types plays a role in explaining this rise.

Keywords: Firm Heterogeneity; Multiproduct firms; Consumer Heterogeneity; Scanner data.

JEL Code: D12, L11, L25, O51.

---

*Researchers own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researchers and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. We thank Philipp Schröder, Frederic Warzynski, and Zhiyuan Chen for suggestions and feedback. We thank seminar participants at Aarhus University, ETOS, SFSU, and Renmin University.

†rcfeenstra@ucdavis.edu

‡lmacedoni@econ.au.dk

§mingzhixu@nsd.pku.edu.cn

1 Introduction

Since the 1980s, the labor share of income has declined in the United States and globally (Elsby et al., 2013; Karabarbounis and Neiman, 2014) while the profit share has risen (Barkai, 2020). Autor et al. (2020) attribute this pattern to competition-induced allocation of production in favor of superstar firms, which are large in size and have high markups and profit margins. In a model of monopolistic competition, they show that an increase in competition, due to international trade or technological progress, causes a reduction in markups of all surviving firms and a reallocation of production towards high-markup firms, leading to an increase in the profit share.[1]

Despite the robust empirical evidence confirming the role of large firms in explaining the rising profit share in Autor et al. (2020), there are challenges in constructing a model consistent with these facts. First, while the empirical analysis of Autor et al. (2020) focuses on superstar firms, their theory relies on monopolistic competition between "small" firms, who take market conditions as given. Second, their theory depends on a sufficiently skewed firm productivity distribution (i.e., a log-convex productivity distribution), and thus, the results require a stronger condition on productivity distribution than commonly imposed, such as the Pareto distribution (which is log-linear).[2] Therefore, it is natural to explore the case of large firms in a more general setting, that allows for increases in the profit share without strong assumptions on the productivity distribution.

This paper introduces consumer taste heterogeneity into a model with large oligopolistic firms. Taste differences are modeled as heterogeneity in the demand shifters that consumers have across disaggregate products. This dimension is is motivated by Hottman et al. (2016), who document that 50-70% of firms size heterogeneity is driven by heterogeneity in product-specific demand shifters. While Hottman et al. (2016) consider a representative consumer, Neiman and Vavra (2020) investigate consumer heterogeneity, and argue that this dimension leads to an extra source of consumer gains from variety. They do not find, however, a significant impact of consumer heterogeneity on aggregate market power. This finding may be related to their assumption that consumer preferences across products are distributed as

[1] There has been a general consensus that the declining labor share and rising profit share is significant and robust (Autor et al., 2020). The pattern is also broadly consistent with other perspectives associated with the declining labor share, such as the treatment of depreciation and capitalization when measuring capital (Smith et al., 2019; Koh et al., 2020), increased cost of housing (Rognlie, 2016), the underestimated labor income due to self-employment (Elsby et al., 2013) and the increasing human capital among top earners (Smith et al., 2019).

[2] While it has been common to use the Pareto distribution in the Melitz model since Chaney (2008), recent research has shown that the truncated Pareto or the log-normal distribution for productivity is a better fit to the trade data: see Fernandes et al. (2018) and Adão et al. (2020), for example. Neither of these distributions are globally log-convex, as required by Autor et al. (2020).

Pareto, but with different preferred products for each consumer. So the Pareto assumption on taste parameters – like the analogous Pareto assumption on firm productivities – combined with the assumption that there is no product that is preferred by a positive mass of consumers, appears to rule out any impact on the aggregate markup and profit share.[3]

In contrast, we suppose that discrete groups of consumers – which we refer to as "consumer segments" – have identical preferences, as defined by their demand shifters which differ across segments. We impose no special assumption on the distribution of tastes across segments beyond its discrete nature. We find that markups are in general higher in this heterogeneous consumer setting than the markups of a firm with identical aggregate market share selling to a representative consumer. Hence, conditional on the same firm level shares, the implied markups in a model with heterogeneous consumers are larger than in a model with a representative consumer. This is because a firm facing heterogeneous consumers has different demand elasticities across segments so that it can exploit the (relatively) lower demand elasticity in one segment, charging higher prices for all consumers.

To determine the extent to which consumer heterogeneity contributes to high and rising profits of large, multi-product firms, we leverage rich scanner data from Nielsen that contains information on sales, prices, and quantities at the barcode level in US supermarkets. We further combine the data with information on purchases at the household-retailer-barcode level for a representative sample of households. Using a nested-CES demand system, we derive a segment-specific demand shifter for each barcode associated with the appeal of the barcode for each segment. In this manner, we divide consumers into segments according to their their inferred demand shifters.[4]

Allowing for this consumer heterogeneity, implied markups are 8 percentage points larger on average than predicted by the formula for a representative consumer model (such as Hottman et al. (2016)). Furthermore, the percentage difference is significantly greater when considering the largest firms: for the top five and one firms, respectively, markups are 13 and 18 percentage points higher than the counterparts in a model with the representative consumer. A similar result occurs when we consider total profits. The aggregate profit share is 3 percentage points larger under consumer heterogeneity, although there are differences across product groups: the effect of consumer heterogeneity on the profit share is small for bread (0.08 percentage point difference) and juices (0.1), and large for batteries (10.7) and

[3]In a trade model with heterogeneous firms that differ in their productivities, using the Pareto distribution for productivities rules out any impact of import competition on the profit share, even with very general consumer preferences and variable firm markups (Arkolakis et al., 2019); this result does not hold, however, with alternative distributions such as the truncated Pareto (Feenstra, 2018).

[4]We do not consider how product characteristics (such as flavor or functionality) may affect demand shifters due to data limitations. For an application of such case to the ice cream product group, see Draganska et al. (2007).

diapers (15.5).

The aggregate profit share predicted by our model increases from 28.6% to 32.3% from 2004 to 2014, and the increase is concentrated in the last three years. The aggregate profit share in 2004 is 2 percentage points higher than what predicted by a model of small firms, and the difference increases to 5 percentage points in 2014. We decompose such differences into the component due to consumer heterogeneity and the component due to firm size. We find that both components play a role in the increase of the profit share. The contribution of consumer heterogeneity, which is the difference in profit share predicted by our model and a model with large firms and a representative consumer, increases from 1 to 2 percentage points. The contribution of firm size, which is the difference in prediction from a model with large firms and a representative consumer and a model with small firms, increases from 1 to 3 percentage points.

To further investigate the role of consumer heterogeneity in explaining the increase in the aggregate profit share, we focus on three key aspects: changes in preferences across segments; changes in the strategies followed by firms to deal with heterogeneous consumers; and changes in the fixed costs associated with targeting segments. We find that all three dimensions contribute to explaining rising firm markups and profit share. Specifically, we document a rise in consumer heterogeneity and a better ability of firms to exploit such heterogeneity using their product mix. Furthermore, we find a rise in the fixed costs, which can generate further market concentration.

Related Literature Our paper relates to the growing body of literature on the determinants of firm size heterogeneity. The seminal work by Hottman et al. (2016) showed that among the possible determinants of firm size, product appeal and product scope are the most quantitatively relevant.[5] Product appeal has been further examined by Faber and Fally (2021), who find that larger firms tend to cater to the taste of the richer consumers within a country, and Aw-Roberts et al. (2018), who document that the taste component of product appeal could be quantitatively relevant in several industries.[6] Failure to account for the heterogeneity in such demand shifters can lead to substantial bias in estimating firms' productivity heterogeneity (Foster et al., 2008; Blum et al., 2018).[7] This paper examines the

[5] Macedoni and Xu (2018) explore the role of supply factors such as productivity and flexibility in determining firm product scope.

[6] The authors assume that differences in terms of demand shifters of one product across countries of destination can be attributed to differences in taste across countries. As they use customs data and a product has a higher level of aggregation, we consider their findings as industry-level evidence.

[7] In the context of global markets, demand heterogeneity has been considered as an important factor for the selection of export destinations by firms and for export performance. Khandelwal (2010), Johnson (2012) and Feenstra and Romalis (2014) estimate demand shifters across countries using country and product-level

contribution of consumer taste heterogeneity in explaining the size heterogeneity of firms. We are agnostic about the sources of the taste heterogeneity.[8]

The empirical literature has predominantly assumed that differences in taste manifest clearly when considering consumers from different geographical areas. For instance, Coşar et al. (2018) examine taste differences across countries in the car market and Jäkel (2019) in the chocolate market. Atkin (2013) study taste heterogeneity for food products across regions in India, and builds a model of habit formation to endogenize the development of taste heterogeneity. Our paper relaxes the assumption that geography or income is the only determinant of taste heterogeneity, by dividing consumers into segments according to their demand shifters. Our approach stands in line with the review by Nevo (2011), who argues that standard consumer attributes (e.g., income, education) are only partially correlated with consumer choices and that unobserved variables (at least to the researcher) are quantitatively more relevant.[9]

As already mentioned, a closely related paper is Neiman and Vavra (2020) who also assume heterogeneous consumers. They document that over the last 15 years, households have increasingly concentrated their expenditure on specific barcode items with the preferred items differing across then, or what they call "niche" consumption. That pattern is rationalized by increasing product variety along with an increase in the fixed costs of learning about each variety. Brand (2021) finds a reduction in the price sensitivity of consumers in supermarkets during the last decade, which he also attributes to greater "niche" consumption. He argues that improvements in supply-chain management has allowed producers to increasingly focus on newer products that target particular consumers, and the reduced price sensitivity is associated with rising markups. In our paper, we also study how firms exploit consumer heterogeneity to charge higher markups and evaluate the effects of this pattern on the aggregate profit share, finding that consumer heterogeneity has increased the profit share over time. We also find an increase in the fixed marketing costs of targeting a segment.

---

trade data; using firm-level data, Baldwin and Harrigan (2011), Kugler and Verhoogen (2012), Crozet et al. (2012), Hallak and Sivadasan (2013), Di Comite et al. (2014) and Roberts et al. (2018) document that heterogeneous demand (i.e., the firm-level demand residuals) are important to account for firms' exporting patterns. Several factors are responsible for the heterogeneity in demand shifters. For instance, product appeal can originate from the quality of a product that makes it perceived as intrinsically better (Fajgelbaum et al., 2011), or it can arise from the closeness of the product to the taste of consumers (Coşar et al., 2018). Furthermore, product appeal is influenced by the distribution channels, which are only partially controlled by manufacturers (Guan et al., 2019).

[8]The difference in consumer taste can be traced to genetic reasons according to Drewnowski et al. (1997).

[9]This is apparent when considering demographic characteristics of different segments of consumers. In the Appendix A.6.1, in which we provide a detailed analysis for the product group of oral hygiene products, we show that demographic characteristics only partially account for different purchasing patterns. Along these lines, Neiman and Vavra (2020) find that demographic characteristics such as income, race and education cannot explain the household expenditure patterns that they focus on.

That a change in the fixed cost technology can drive up profits is also a mechanism explored Hsieh and Rossi-Hansberg (2021) for service industries including retailing.

The paper is organized as follows. After briefly reviewing the data in section 2, section 3 describes the model that guides our empirical analysis. Section 4 outlines the strategy we use to divide consumers into segments and identify the segment taste shifter for each product. Section 5 applies the model to evaluate markups and the profit share, whiles section 6 focuses on how firms adjust their targeting strategies in response to consumer heterogeneity. Section 7 concludes and additional results are gathered in the Online Appendix.

2 Data

We use the Retail Scanner and the Consumer Panel (or the Home Scanner) databases collected by Nielsen (US) and provided by the Kilts Center for Marketing Data Center at the University of Chicago Booth School of Business.[10] The retailer scanner data provides price and quantity for each transaction of barcode product by a store, and stores are widely distributed across the United States. A barcode is a 12-digit Universal Product Code (UPC). The transaction is available on a weekly basis. Besides the sales data, we also observe detailed product characteristics at the barcode level, which we use to identify the firm (or brand) supplying the given product. The unique numeric barcode and brand identifiers allow us to calculate firm sales and product scope, i.e., the number of barcode products produced by a firm. From the Retail Scanner Database, we source our firm-specific variables.

We trace household consumption at the barcode level using the home scanner data, which is the household-level panel for 2004-2016 and contains about 700 million unique transactions on non-service retail spending made by 170,000 households distributed across 20,000 zip codes. We observe the associated price and quantity at the barcode level for each purchase by a consumer in different stores from this data.[11] In the analysis, we aggregate the expenditure and quantity (of each barcode item) across stores to focus on the consumers' annual purchases of each item. Along with the detailed budget shares for 60,538 households in 2012, we also have information on some characteristics such as race, age, education, marital status, income level, and presence of children. The products included by both databases cover a wide range of categories (i.e., there are 597 distinct product modules in 2012) such as food, liquor, snacks, household appliances, and personal care products.

[10] While we focus on 2012 for cross-section results, we also provide time-series evidence over 2004-2014. For more detailed information about the Retail Scanner and Consumption Panel Databases, see Hottman et al. (2016), Faber and Fally (2021), and Feenstra et al. (2020).

[11] Nielsen collects these data using scanner devices, with which households scan each transaction they have made after shopping.

3 Model

The aim of the model is twofold. First, we use the structure on the demand side of the model to identify consumer heterogeneity in the data. Second, we study the implications of consumer heterogeneity for markups, profits, and marketing costs. Our model follows Hottman et al. (2016) but with one important difference: there are segments of different consumers with different preferences. In particular, we assume that the nested-CES preference structure with demand shifters applies to each consumer segment, where demand shifters for each product differ across the segments. The elasticities of substitution are identical across consumer groups, however. Our model features a discrete number of multiproduct firms that compete oligopolistically: both prices and the set of products are decided by taking into account the strategic interactions across firms.

3.1 Demand

Consumers are divided (exogenously) into $n = 1, ..., N$ types or segments (we use these terms interchangeably). So we use the terms "consumer types" and "consumer segments" interchangeably. Each segment is made of $L_n > 0$ consumers with average per-capita income of $w_n > 0$ and identical tastes within each segment. The total number of consumers is $L = \sum_{n=1}^{N} L_n$. We stress that it is not the average income of consumers that distinguishes segments, but rather, their demand shifters for UPC items, as we describe now.

All consumers have a nested preference structure with three layers. The first layer consists of a continuum of product groups indexed by g, over which consumers have the following Cobb-Douglas utility function:

$$\ln U_n = \int z_g \ln U_{gn} dg, \qquad \int z_g dg = 1, \tag{1}$$

where z_g is a product group shifter that is common across consumers of all segments. The sub-utility U_{gn} is a CES aggregate over the bundle of products produced by $f = 1, ..., F_g$ firms:

$$U_{gn} = \left[\sum_{f=1}^{F_g} \left(z_{fg} Q_{fgn} \right)^{\frac{\sigma^g - 1}{\sigma^g}} \right]^{\frac{\sigma^g}{\sigma^g - 1}}, \tag{2}$$

where $\sigma^g > 1$ is the elasticity of substitution across firms in product group g and z_{fg} is a firm-specific demand shifter. The firm-specific demand shifter is also common across segments.

The third and most disaggregate level of the utility function is the consumption Q_{fgn} which is a CES aggregate over the varieties that the firm f produces and sells to a consumer segment n. We identify a "variety" as a UPC item, denoted indexed by $i \in I_{fg}$, where I_{fg} is the set of all UPC's sold by firm f in product group g. We allow consumer segment n to potentially demand a subset of these UPC's, denoted by $I_{fgn} \subseteq I_{fg}$. The consumption Q_{fgn} is then given by:

$$Q_{fgn} = \left[\sum_{i \in I_{fgn}} \left(z_{ifgn} q_{ifgn} \right)^{\frac{\eta^g - 1}{\eta^g}} \right]^{\frac{\eta^g}{\eta^g - 1}}, \tag{3}$$

where z_{ifgn} is a UPC-firm-good demand shifter that by assumption is *common* to all consumers within the segment n, while q_{ifgn} is the per-capita quantity consumed in that segment, and $\eta^g > \sigma^g$ is the elasticity of substitution between UPCs of a firm.

With this demand structure, the per-capita demand for a variety identified by its UPC-firm-group-segment is given by:

$$q_{ifgn} = z_g w_n P_{gn}^{\sigma^g - 1} z_{fg}^{\sigma^g - 1} P_{fgn}^{\eta^g - \sigma^g} z_{ifgn}^{\eta^g - 1} p_{ifg}^{-\eta^g}, \tag{4}$$

where p_{ifg} is the price of the UPC and it is common across segments. In other words, we rule out the possibility that firms can price discriminate across consumer segments. The CES price indexes will still depend on the segment n, because they are defined inclusive of the demand shifters:

$$P_{gn} = \left[\sum_{f=1}^{F_g} \left(\frac{P_{fgn}}{z_{fg}} \right)^{1 - \sigma^g} \right]^{\frac{1}{1 - \sigma^g}}, \tag{5}$$

$$P_{fgn} = \left[\sum_{i \in I_{fgn}} \left(\frac{p_{ifg}}{z_{ifgn}} \right)^{1 - \eta^g} \right]^{\frac{1}{1 - \eta^g}}. \tag{6}$$

For the remainder of the model section, we drop the product group index g to ease the notation. The elasticity of the price indexes with respect to prices can be expressed as appropriately defined market shares:

$$\frac{\partial P_n}{\partial P_{fn}} \frac{P_{fn}}{P_n} = s_{fn} \equiv \frac{\left(\frac{P_{fn}}{z_f} \right)^{1-\sigma}}{\sum_{f'=1}^{F} \left(\frac{P_{f'n}}{z_{f'}} \right)^{1-\sigma}}, \tag{7}$$

$$\frac{\partial P_{fn}}{\partial p_{if}} \frac{p_{if}}{P_{fn}} = s_{ifn} \equiv \frac{\left(\frac{p_{if}}{z_{ifn}} \right)^{1-\eta}}{\sum_{j \in I_f} \left(\frac{p_{jf}}{z_{jfn}} \right)^{1-\eta}}, \tag{8}$$

where s_{fn} is firm f's revenue share within the expenditures of consumers in segment n, with

$\sum_{f=1}^{F} s_{fn} = 1$, and s_{ifn} is the revenue share of barcode product i within the expenditure of consumers in segment n on firm f's UPCs, with $\sum_{i \in I_f} s_{ifn} = 1$.

These preferences allow us to nest the results of a representative consumer, in which case all segments are identical. We formalize the difference between that case and the heterogeneous consumers by reintroducing the goods subscript g and using the definition:

Definition 1 *The representative consumer (RepC) model has $z_{ifgn} = z_{ifg}$ $\forall$ g and n, while $\forall$ g the heterogeneous consumer (HetC) model has $z_{ifgm}/z_{ifgn} \neq z_{jf'gm}/z_{jf'gn}$ for at least two firms $f \neq f'$ each selling a UPC item $i \neq j$ to at least two consumer segments $m \neq n$, with $i \in I_{fgm} \cap I_{fgn} \neq \emptyset$ and $j \in I_{f'gm} \cap I_{f'gn} \neq \emptyset$.*

In other words, the *HetC* model has at least two firms f and f' selling the UPCs i and j, respectively, to consumer segments m and n with relative demand parameters that differ. In contrast, if the demand parameters are equal across all consumer segments, then we are in the *RepC* model, which is equivalent to a single consumer segment.

3.2 Technology

The production of one unit of UPC i by firm f requires c_{if} units of a numeraire resource, which is meant to represent the combinations of factors of production used by firms. UPCs within a firm can vary in their unit costs to capture the idea of core competence, which is a common and verified assumption in the literature (Eckel and Neary, 2010; Arkolakis et al., 2021). Furthermore, we assume that c_{if} is independent of the quantity sold and of the consumer segment that is purchasing.

Our assumption that unit costs per UPC are the same for all consumer segments means that we assume away differences in variable trade costs associated with reaching different segments. Hence, our model best represents the case in which firms face different consumer types within the same geographical area. This assumption is mainly motivated by data availability: since we do not know the geographic origin of UPCs, it is difficult to measure any differences in trade costs by segment. In addition, as price discrimination (or lack thereof) across different geographic locations is a well-studied phenomenon (Atkeson and Burstein, 2008; Simonovska, 2015; DellaVigna and Gentzkow, 2019), we prefer to focus here on optimal pricing across heterogeneous consumers when discrimination is not allowed.

3.3 Pricing

Initially, we do not consider any fixed costs of the firm, but introduce those in section 6.3. Firm f's "variable" profits – meaning the excess of revenue over the variable costs of production – within a product group are equal to:

$$\pi_f = \sum_{i \in I_f} \sum_{n=1}^{N} L_n (p_{if} - c_{if}) q_{ifn}. \tag{9}$$

Firms maximize their profits by choosing p_{if} $\forall i \in I_f$, as well as the product set I_f (see section 6.3). Firms take as given the vector of UPC-segment specific demand shifters z_{ifn}, and the marginal costs of production c_{if}.

In a standard multiproduct firm setting with Bertrand competition, a firm internalizes two effects from raising the price of a UPC. First, raising the price increases its own price index P_{fn}, which increases the demand for all its other UPCs. Second, it increases the aggregate price index P_n, further raising the demand for all its other UPCs depending on how large the firm is. In our framework, there is an extra dimension due to the segment heterogeneity — firms internalize the fact that raising the price of one variety has different effects across segments.

The equilibrium markups $\mu_{if} \equiv p_{if}/c_{if}$ for all UPC-firm (if) pairs are implicitly defined by the first-order condition:[12]

$$(\sigma - 1)\left[1 - \frac{\sum_{n=1}^{N} \alpha_n s_{fn}^2 s_{ifn}}{\sum_{n=1}^{N} \alpha_n s_{fn} s_{ifn}}\right] = \frac{\eta}{\mu_{if}} - \frac{\sum_{n=1}^{N} \alpha_n s_{fn} s_{ifn} \left[(\sigma - 1) s_{fn} + (\eta - \sigma)\right] \sum_{i \in I_f} \frac{s_{ifn}}{\mu_{if}}}{\sum_{n=1}^{N} \alpha_n s_{fn} s_{ifn}} \tag{10}$$

where $\alpha_n = L_n w_n / \sum_{n=1}^{N} L_n w_n$ denotes the share of income of consumers in segment n. Several features of these markups should be noted.

First, we have indexed the markups by the UPC item i sold by firm f, because markups can differ across the items sold by a multiproduct firm even if the marginal costs of those items happen to be equal. This feature of *variable markups* with heterogeneous consumers is also found by Neiman and Vavra (2020) and will be demonstrated empirically in our model. It stands in marked contrast to the case of a multiproduct firm in a *RepC* model. The *RepC* model is equivalent to the multiproduct firm selling to one segment, i.e., $N = 1$ in the summations within (10). The markup is then immediately solved from (10) as:

$$\mu_{\text{RepC}} = \frac{\sigma - (\sigma - 1) s_f}{(\sigma - 1)\left(1 - s_f\right)}, \tag{11}$$

where $\sum_{i \in I_f} s_{ifn} = s_{fn} = s_f$ is the *total market share* of the firm since there is only one

[12] The derivation of (10) as well as the proof of Proposition 1 are in Appendix A.1.

segment, $N = 1$. The markup of a firm in (11) is increasing in its total market share s_f, and *only* depends on that total share in the *RepC* model. Even though a firm might be selling multiple products with different individual market shares, because it internalizes the cross-demand effects of changing any individual price, it ends up charging the same markup on all of its products (regardless of their marginal costs).[13]

We can contrast the *RepC* markup in (11) with the markup charged by a firm with the same total market share, but facing heterogeneous consumers. For simplicity, consider the markup in (10) for single-product firms in the *HetC* model satisfying Definition 1. Then we have the following result:

Proposition 1 *The HetC model with single product firms has markups that are at least as high as the markups for single or multiproduct firms in the RepC model, conditional on having the same firm-level market shares, and strictly higher for at least one firm.*

To establish this result, consider the optimal markup in (10) of a firm f with only one UPC, i.e., $I_f = \{i\}$. As a result, $s_{ifn} = 1$. Firm f's markup is then computed from (10) as:

$$\mu_{\text{HetC}} = \frac{\sigma - (\sigma-1)\frac{\sum_{n=1}^{N}\alpha_n s_{fn}^2}{\sum_{n=1}^{N}\alpha_n s_{fn}}}{(\sigma-1)\left[1 - \frac{\sum_{n=1}^{N}\alpha_n s_{fn}^2}{\sum_{n=1}^{N}\alpha_n s_{fn}}\right]}. \tag{12}$$

We see that for a single-product firm in the *HetC* model, the markup in (12) is increasing in $\sum_{n=1}^{N}\alpha_n s_{fn}^2 / \sum_{n=1}^{N}\alpha_n s_{fn}$, which is the sum of squared market shares across segments, weighted by α_n and divided by the firm's total market share $s_f = \sum_{n=1}^{N}\alpha_n s_{fn}$. In contrast, the markup in (11) is increasing in s_f. By Jensen's inequality, $\sum_{n=1}^{N}\alpha_n s_{fn}^2 / \sum_{n=1}^{N}\alpha_n s_{fn} \geq \sum_{n=1}^{N}\alpha_n s_{fn} = s_f$, and in the proof of Proposition 1 we show that this inequality holds strictly for at least one firm given the heterogeneity in demand shifters in Definition 1. For the same firm-level market shares, therefore, a single product firm that sells to multiple segments will have a higher markup than a firm selling to a single segment. Since the markup in (12) increases with s_{fn}^2, we can also conclude that the larger the variance in s_{fn} across segments, the higher the firm's markups.

The intuition for this result is as follows. In the presence of consumer taste heterogeneity, firms face a different perceived demand elasticity across segments. Rather than charging a markup consistent with an average of the perceived demand elasticities, however, firms optimally use a greater weight on low-elasticity segments. The higher profits obtained by

[13]This result is demonstrated in the CES case for a representative consumer by Hottman et al. (2016) and Feenstra (2004), p. 267.

charging higher markups to low-elasticity and all other segments more than offsets the loss in demand on the high-elasticity segments.

The Aggregate Profit Share. We have shown above that markup of a single-product firm selling to heterogeneous consumers is larger than the markup of a firm, with identical aggregate market share, selling to a representative consumer. Hence, conditional on the firm level market share, the implied profit share of a model with heterogeneous consumers is larger than a model of a representative consumer. In addition, since markups under oligopoly are larger than markups under the assumption of atomistic firms, the implied profit share of the representative consumer model with large firms is larger than that implied by a model of small firms. So we obtain a ranking of markups across these three models, conditional on the firm-level market shares.

We will measure the aggregate share of variable profits in each product group as:

$$\frac{\Pi}{R} = \frac{\sum_{f=1}^{F}\sum_{i\in I_f}\pi_{if}}{\sum_{f=1}^{F}\sum_{i\in I_f}r_{if}} = \sum_{f=1}^{F}\sum_{i\in I_f}\frac{\mu_{if}-1}{\mu_{if}}\frac{r_{if}}{\sum_{f=1}^{F}\sum_{i\in I_f}r_{if}} = \sum_{f=1}^{F}\sum_{i\in I_f}\frac{\mu_{if}-1}{\mu_{if}}s_{if}s_f, \tag{13}$$

where r_{if} denotes the revenues of a firm f in product i and $\frac{\mu_{if}-1}{\mu_{if}}$ is the ratio between profit margin and markups at the product-firm level, which is the profit share at the product-firm level. The aggregate share of variable profits is the sum of each product-level profit shares, weighted by the product revenue share within a firm s_{if} and firm market share s_f.

In our quantitative exercise, we observe the market shares at high level of disaggregation and apply the model to measure the implied markups, thus estimating the aggregate share of variable profits as shown above. We stress that in order to estimate the markups from (10) we need to have the UPC-firm share (if) sold within each product group g *to each* consumer segment n, or s_{ifgn}. Measuring these shares will depend on our identification of meaningful consumer segments, which we consider in the next section.

4 Consumer Heterogeneity

As a first step, we describe the strategy we follow to estimate the segment-specific demand shifters. Reintroducing the product group notation g, we taking logs of (4) to decompose the demand for a UPC as:

$$\ln q_{ifgn} = \underbrace{\ln\left(z_g w_n P_{gn}^{\sigma^g-1} z_{fg}^{\sigma^g-1} P_{fgn}^{\eta^g-\sigma^g}\right)}_{\text{Firm-segment}\equiv\ln\Phi_{fgn}} + (\eta^g-1)\times \underbrace{\ln z_{ifgn}}_{\text{UPC-Firm-segment}} - \eta^g \ln p_{ifg}. \tag{14}$$

After controlling for the price, the quantity demanded of a UPC by a segment is a function of a firm-segment component, which we label Φ_{fgn}, and a firm-segment-UPC component, denoted by z_{ifgn}. Let us assume that the demand shifter z_{ifgn} can be decomposed into two components: first, $\bar{\phi}_{ifg}$ is a firm-UPC component that captures the quality or appeal of the product and is globally recognized across consumers' segments; and second, ϕ_{ifgn} is a firm-UPC-segment component that captures the taste of consumers in segment n for the UPC i produced by firm f. In particular:

$$\ln z_{ifgn} = \ln \bar{\phi}_{ifg} + \ln \phi_{ifgn}. \tag{15}$$

Because $\bar{\phi}_{ifg}$ captures the demand shifter that is commonly recognized across consumers' segments, without loss of generality we normalize the taste component (firm-UPC-segment) of the demand shifter so that it has mean zero:

$$\frac{1}{N_g}\sum_{n=1}^{N_g} \ln \phi_{ifgn} = 0. \tag{16}$$

Hence, differences in taste *across segments* are represented by the difference of ϕ_{ifgn} from zero. To empirically identify such demand shifters, we follow a two-step procedure, which we describe in the following sections. First, we divide consumers into segments (section 4.1). Second, we estimate the segment-specific demand shifters using (14) (section 4.2).

4.1 Assignment of Consumers to Segments

As noted earlier, consumer types are defined as having identical demand shifters for UPC items within a product group. Empirically, we allocate each of the L household into one of the non-overlapping sets H_n, where $\cup_{n=1}^{N} H_n = \{1, ..., L\}$, according to the similarity of their budget shares. Let x_{ifh} denote the budget share of household h on UPC-firm if relative to the entire sample budget share on the same UPC, so x_{ifh} is defined as:

$$x_{ifh} = \frac{q_{ifh}p_{if}}{w_h} \Big/ \frac{\sum_{h' \in H_n} q_{ifh'}p_{if}}{\sum_{h' \in H_n} w_{h'}}.$$

where w_h is the total expenditures of consumer h in a given product group. We apply this definition to each product group separately and, thus, x_{ifh} is defined for all consumers and products within a product group. The relative budget share of a consumer captures a simple measure of its demand shifter, or tastes, relative to the aggregate demand shifter.[14]

Empirically, we use the k-means clustering algorithm to assign each consumer h in our data to a segment, or cluster H_n. The k-means clustering is a statistical algorithm for classifying objects in clusters based on their attributes in a certain number of categories (Everitt, 1993).[15] The method employs the distance between objects as a dissimilarity measure when forming the clusters such that each object belongs to the cluster with the nearest mean.

The assignment of consumers into a segment is done by minimizing the within-cluster dissimilarity between individuals in terms of the relative budget shares x_{ifh}.[16] The objective function of the algorithm is:

$$\min \sum_{n=1}^{N} \sum_{h \in H_n} \sum_{f=1}^{F} \sum_{i \in I_f} ||x_{ifh} - \mu_{ifn}||,$$

where μ_{ifn} is the segment-specific average of the relative budget share x_{ifh}, and $||x_{ifh} - \mu_{ifn}||$ is therefore a measure of the dissimilarity between the relative budget shares of individual h and the mean for the segment n.

The k-means clustering algorithm requires two choices: i) the number of segments N and ii) the measure of the dissimilarity defined by the operator $||\ ||$. The number of segments N is chosen in order to maximize the Calinski-Harabasz pseudo F statistics. A higher value of the F statistics is associated with a larger dissimilarity across clusters. In practice, our algorithm requires two loops. In the inner loop, given N the algorithm divides consumers into segments and computes the Calinski-Harabasz pseudo F statistics. We run the inner loop for $N = 2, ..., 15$, and choose the number of segments N that maximizes the Calinski-Harabasz pseudo F statistics. Finally, we use the Minkowski distance of order one, i.e. the absolute value, as the measure of dissimilarity.[17]

To demonstrate the above method, let us consider the case of oral hygiene products. The k-means clustering algorithm divides households into five types (where the Calinski-

[14] We have verified that the results are robust to using the absolute purchases on goods rather than the relative budget shares.

[15] K-means clustering has been widely used in economic analysis to classify states into regions with similar cycles (Crone and Clayton-Matthews, 2005), to group countries based on economic performance (Levy-Yeyati and Sturzenegger, 2005; Caballero, 2014; Ghosh et al., 2014), and to classify firms into technological clusters (Chyi et al., 2012).

[16] As a robustness check, we also consider two additional criteria to divide consumers into segments based on the geographical location of consumers, and their preferred distribution channel, i.e. type of store: see Appendix A.5. We focus in the main text on the consumer type as defined by the relative budget shares because, in our model, firms cannot price discriminate across consumers and there are no differences in costs to reach different segments. Empirically, while is little evidence of price discrimination within retail chains across geographic locations, though different chains charge different prices (DellaVigna and Gentzkow, 2019).

[17] In Stata, this is achieved with the command cluster k-means **x**, start(firstk) measure(L(1)) k(5). Our results remain robust to using Minkowski distance of different order to cluster consumers.

Harabasz pseudo F statistics is maximized). Table 1 summarizes the main characteristics of each segment. For exposition purposes, we rank segments by their aggregate expenditure shares in the toothpaste market: segment 1 is the largest segment, as it captures 38% of the market while segment 5 is the smallest, as it captures 2% of the market. Segment 3 exhibits the largest expenditure per household on toothpaste ($19) while segment 5 the smallest ($6). Segments 1, 2, and 4 are over-represented by young households with children. In contrast, segments 3 and 5 are over-represented by older households, while segment 5 is also over-represented by low-income households (41%) and by households with a male head. The remaining four segments have a similar level of income, and race and ethnicity are fairly similar across segments (see Appendix A.6).

Table 1: Segments characteristics - Oral Hygiene

	#	Exp per Household	Total Share	Main Brand	Distinctive Features
1	18437	$14	0.38	Colgate	Young, with Child
2	15843	$13	0.32	Crest	Young, with Child
3	7157	$19	0.20	Sensodyne	Older, High-educated
4	4385	$13	0.08	Aquafresh	Young, with Child
5	2587	$6	0.02	AIM	Low Income/Education, Older, Male Head

In the application of the method to all product categories, we use the relative expenditure shares of UPCs that are purchased by at least 1% of the households in the data so that we have heterogeneous consumption purchases within a segment. Table A.1 in the Appendix reports the number of segments by consumer type per product category. The number of segments according to this criterion ranges from two (Prepared foods) to seven (Beer).

4.2 Estimation of the Taste Parameter

Having divided consumers into segments, we estimate product and product-segment specific demand shifters. To achieve this, suppose that (14) represents the demand for consumer h in segment n. That is, we replace segment n in (14) by consumer h, and then for each UPC we take the average across all consumers in the same segment n to obtain:

$$\overline{\ln q_{ifgh}} = \ln \Phi_{fgn} + (\eta^g - 1) \ln z_{ifgn} - \eta^g \ln p_{ifg}, \tag{17}$$

where $\overline{\ln q_{ifgh}} \equiv \sum_{h \in H_n} \ln q_{ifgh}/L_n$ is the average quantity demand by consumers in segment n. By assumption, $z_{ifgh} = z_{ifgn}\ \forall h \in \{H_n\}$, hence, $\sum_{h \in n} \ln z_{ifgh}/L_n = \ln z_{ifgn}$. We estimate the firm-segment demand shifter $\ln \Phi_{fgn}$ in (17) by running this regressions separately for each firm-segment pair fn in each product group g. In other words, for each sample consisting of UPC products supplied and consumed by the firm-segment pair, $(\eta^g-1)\ln z_{ifgn}$ is obtained as the residual of a regression of $\overline{\ln q_{ifgn}} + \eta^g \ln p_{ifg}$ on a constant that captures the value $\ln \Phi_{fgn}$. We use the elasticity of substitution η^g from Hottman et al. (2016), so from this procedure we obtain the estimates $\hat{z}_{ifgn}$.

With the demand shifters $\hat{z}_{ifgn}$ for each UPC item, we apply the normalization outlined in (15) and (16) to obtain the common component of the demand shifter $\ln \hat{\bar{\phi}}_{ifg}$ and the segment-specific taste parameter $\ln \hat{\phi}_{ifgn}$. The segment's taste for a particular UPC item is captured by the deviation of that demand shifter from the average demand shifter across consumer segments for that UPC. The identifying assumption is that, conditional on prices and on the firm-segment component of demand, differences in the consumption patterns across segments reflect different preferences. In other words, we condition on a given firm-UPC item, which is sold at the same price and the same characteristics across different segments, and then demand differences reflect the heterogeneity of tastes.

5 Markups and the Profit Share

Combining our data on UPC-segment specific sales, we evaluate the implications of consumer heterogeneity for firm markups and profits. We first examine the cross-section across firms in 2012 and then examine the change in average profits and markups over time.

5.1 Markups and Profits Across Firms

For each product group, we compute segment, firm-segment, and UPC-firm-segment specific market shares. We use the elasticities of substitution from Hottman et al. (2016) to solve for UPC-specific markups using equation (10). Panel (a) of Figure 1 shows the relationship between the implied UPC markups in our model and the markups that would be implied by the *RepC* formula (11). There is significant heterogeneity of markups within a firm. This is in contrast to the *RepC* result that markups are constant within a firm. For instance, for the top-selling firm in shaving needs, markups $\mu_{if} = p_{if}/c_{if}$ vary from 1.66 to 2.73.

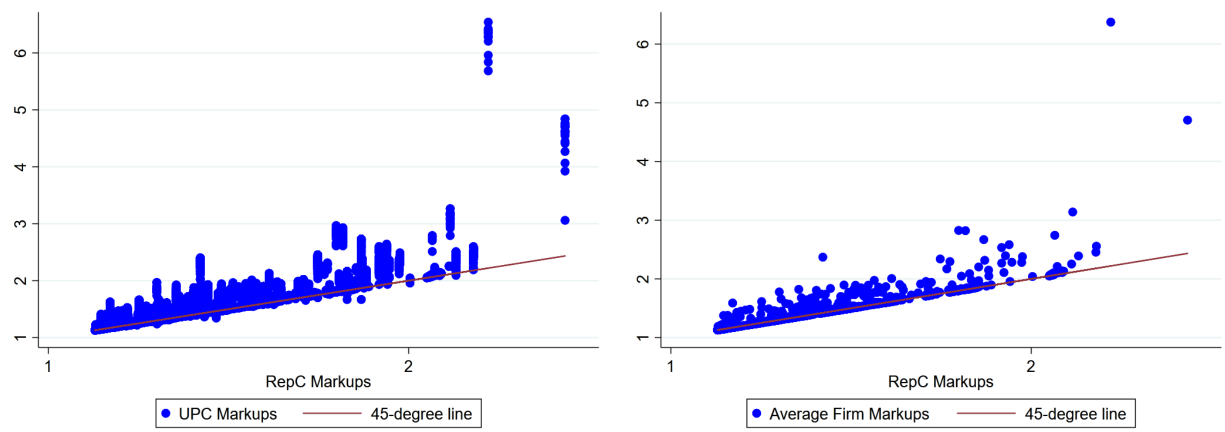
While most UPCs' markups are larger than the *RepC* formula, there are some UPCs where markups are lower. For the top-selling firm in shaving needs, as mentioned just above, the *RepC* formula predicts markups of 1.87 and there are a few products sold by that firm with markups below that value.[18] Across product groups, UPCs with markups below the *RepC* prediction account for 3.6% of UPCs, and they are UPCs with low market share

[18]Notice that having a lower markup than the *RepC* formula for some items sold by a multiproduct firm in the *HetC* model does not contradict Proposition 1, because that result assumes that firms in the *HetC* model are selling only a single product.

and low dispersion of market share across consumer segments. In panel (b) of Figure 1, we report the weighted average markup by firm, where the weights are the UPCs market share within a firm. The presence of consumer heterogeneity generates larger markups than those implied by the *RepC* formula. For the top firm in shaving needs, average markups are 2.67 against the *RepC* formula of 1.87.

Figure 1: Implied Markups

(a) UPC-specific Markups (b) Average Firm Level Markups

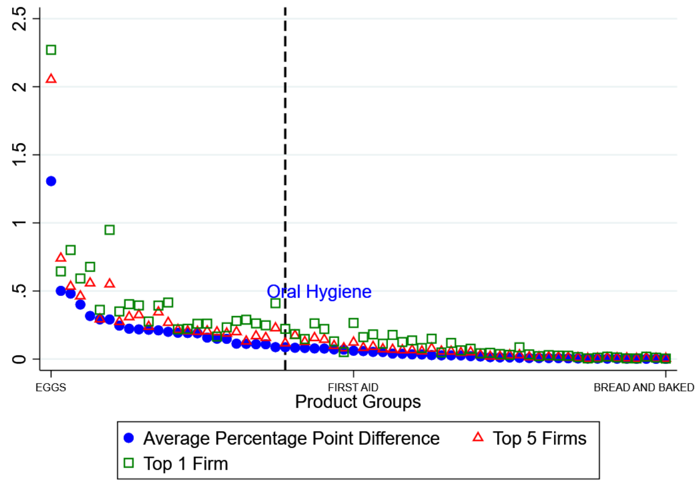


The percentage difference between our markup formula and the *RepC* result differs considerably across product groups, and across firms within a product group. On average, the presence of taste heterogeneity across consumer types implies that markups are 8 percentage points larger across product groups. For the top five and top one firm within each product group, average markups across groups are 13 percentage points and 18 percentage points larger. Therefore, the presence of consumer heterogeneity generates markups that are higher relative to the case of consumer homogeneity, and these differences are magnified for large firms.

Figure 2 shows the average percentage point markup difference between our *HetC* model and the *RepC* formula by product group, as well as the difference for the top five and top one firms. The most significant differences are found in dairy product groups (e.g. eggs, butters, fresh cheeses) and personal care (e.g. shaving needs, diapers). These are product groups with a relatively larger number of consumer segments and relatively higher concentrations. For shaving needs, the average markup in the *RepC* model is 1.68 and it rises to 2.16 in the *HetC* model, while for the top firm these markups are 1.87 and 2.67. The smallest markup differences are found in wine, health aids, and bread and baked goods.[19]

[19]In Table A.2, we report the differences by product group.

Figure 2: Difference in Markups by Product Group: *HetC* versus *RepC*



Similar results emerge when we compute the profit share as the ratio between variable profits and revenues, as in (13). The aggregate profit share in 2012 is 31.5% and this value is in line with the results of Karabarbounis and Neiman (2014), who consider a wider sample of firms than just retailers. In Figure 3, we compare the profits implied by our model to those that emerge in the *RepC* model of a representative consumer. Panel (a) shows that profits are always larger in our *HetC* model, regardless of firm size. The aggregate profit share (31.5%) is 3 percentage points larger than with a representative consumer (28.6%) in 2012.[20] There is considerable heterogeneity across product groups, as the largest difference is in the disposable diaper category, and the smallest in bread and baked products. In a model with small firms, and constant markups, the implied profit share is 26%, which is almost 6 percentage point smaller than our *HetC* case.

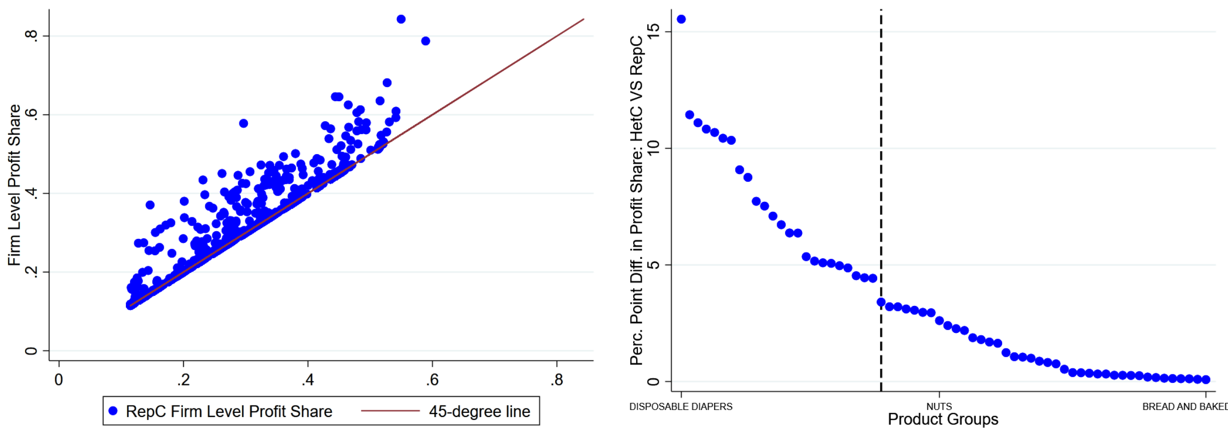
5.2 Profits Over Time

In this section, we study the change in the aggregate and product group specific profit share from 2004 to 2014. We apply the clustering algorithm to the households in the sample of each year. Note that the Consumer Data are more akin to a repeated cross-section than to

[20] In computing the profit share and the aggregate markup for each product group, we use the same weights across UPC items and firms as shown in (13). The only difference between these two calculations, therefore, is that the profit share is averaging the Lerner index $(\mu_{if}-1)/\mu_{if}$, whereas the aggregate markup is averaging μ_{if}. To get a rough idea of the difference in these calculations, consider the average markups for the *HetC* model (1.51) and *RepC* model (1.43), which differ by 8 percentage points. The percentage point difference between the corresponding Lerner indexes is 3.7 (i.e. $(1-1.51^{-1})-(1-1.43^{-1})=0.037$).

Figure 3: Total Profits: *HetC* versus *RepC*

(a) Firm-Level (b) Product Group-Level



a panel data. We then apply our model to infer markups.[21]

Panel (a) of Figure 4 reports the aggregate profit share across models. Our *HetC* model of large firms and consumer heterogeneity predicts the largest profit share across years. From 2004 to 2011, the profit share is around 28-29% and then increases to 31% in 2012 and further increases in the following years until the 2014 value of 32.3%. Relative to a model of small firms, which features constant markups, our baseline model has two additional margins: large firms and consumer heterogeneity. Before 2012, our baseline model generates an aggregate profit share which is 2 percentage points larger than a model of small firms, and consumer heterogeneity accounts for 1 percentage point (see red area in panel (b)).

The increase in the profit share after 2012 is due to three different channels. First, there is a change in the expenditure composition across product groups, as product groups with higher profit share grow more than product groups with low profit share. This is shown in the increase in the aggregate profit share of a model with small firms, where markups within product groups are constant and so is the product group specific profit share. Furthermore, there is an increase in the role of firm size, since the difference between a model of large

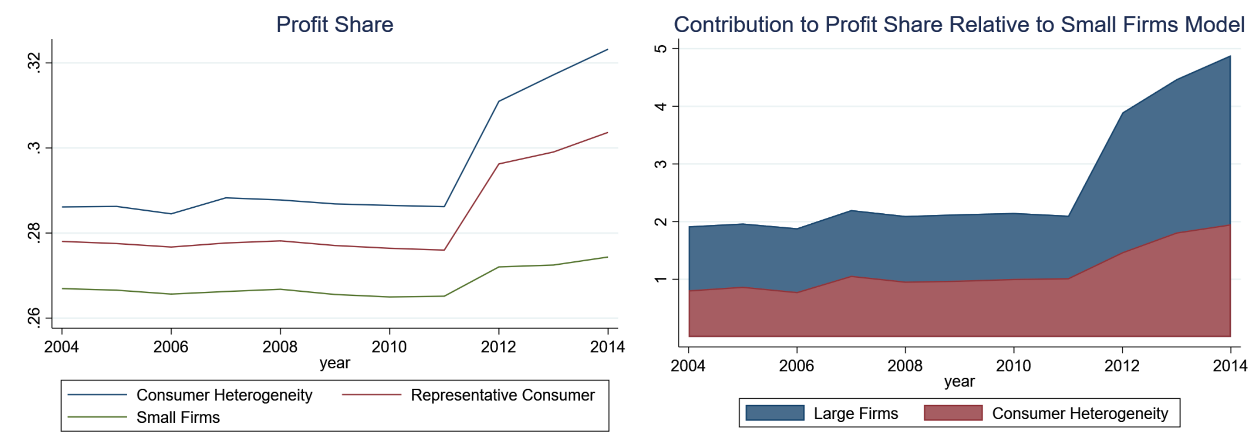
[21]Relative to the cross-section results of the previous section, there is a change in the definition of firms, as now we apply the Nielsen definition of brand ID. This measure is a narrow definition of firm, since the same firm has different brand IDs. For instance, the firm Colgate is made of several brand IDs. For the year 2012 we applied a manual procedure of aggregating the brand IDs provided by Nielsen to the firm level. Such a procedure cannot be applied consistently across all years, but the difference in the definition of firm does not lead to substantial differences in the aggregate predictions. In fact, the aggregate profit share for 2012 is 31.5% using the baseline definition of firms, and 31.1% using the narrower definition of this section. Similarly, the aggregate profit share in the *RepC* model is 28.6% using the baseline definition, and 29.6% using the narrower definition. Hence, the difference between the *HetC* and *RepC* profit share is 3 percentage points under the baseline definition and 2 percentage points under the narrower definition using brands.

firms and a representative consumer and the model of small firms increases from 1 percentage points to almost 3 percentage points in 2014 (see blue area in panel (b)). Finally, there is a smaller increase in the role of consumer heterogeneity as the difference between our model and that of a representative consumer increases from 1 percentage points to almost 2 percentage points in 2014. Notice that in this exercise we do not investigate the sources of the increase in the role of firm size, which could be due to changes in product-segment specific appeal, i.e. to increased consumer heterogeneity.

Figure 4: Profit Share Over Time

(a) Aggregate Profit Share Across Models

(b) Percentage Point Difference Relative to Small Firms Model



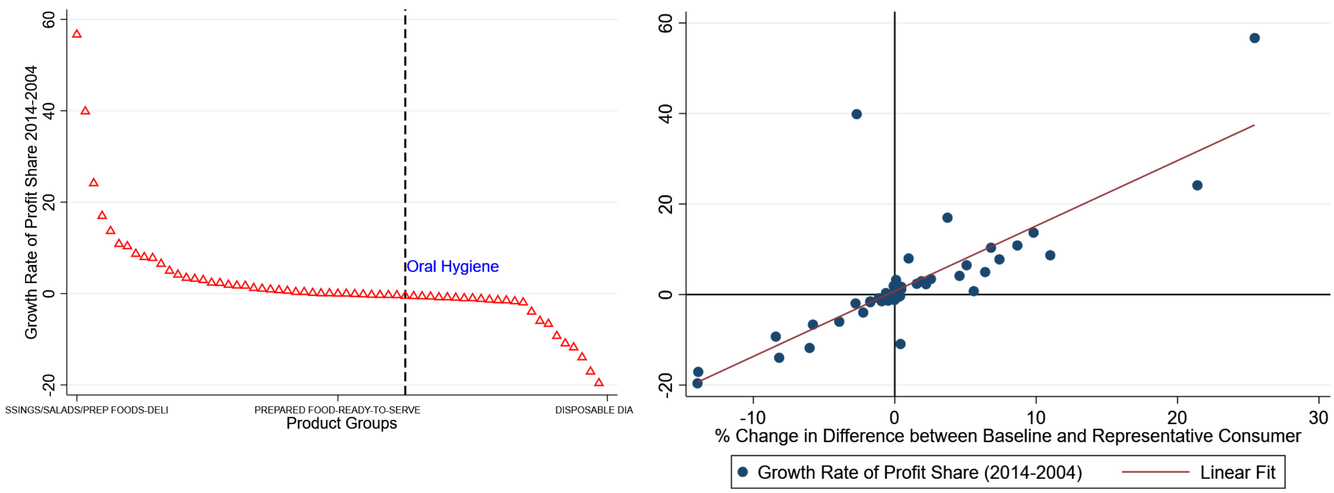
As shown in the cross-section results, there is considerable heterogeneity across sectors, which we illustrate in Figure 5. In some product groups, the profits share has increased, as in salads, while in others, such as diapers, it has decreased. The effects of consumer heterogeneity also varies across groups. In panel (b) of Figure 5, we show that there is a positive correlation between the change in profit share and the change in the difference between our prediction and the prediction of the representative consumer model. This result further highlights the importance of consumer heterogeneity in driving the changes in the profit share.

6 Firm Strategy to Target Segments

In this section, we investigate possible determinants of the rise in the profit share documented above. In particular, we consider changes in consumer preferences, in the strategies with which firms deal with consumer heterogeneity, and in the fixed costs of targeting segments. We find that the heterogeneity of segment preferences has increased and that firms have

Figure 5: Change in Profit Share Across Product Groups

(a) Profit Share Change (b) Contribution of Consumer Heterogeneity



progressively improved their ability to create their own niches and target several segments. Finally, we record an increase in the fixed costs of targeting segments which coincides with the rise in the profit share.

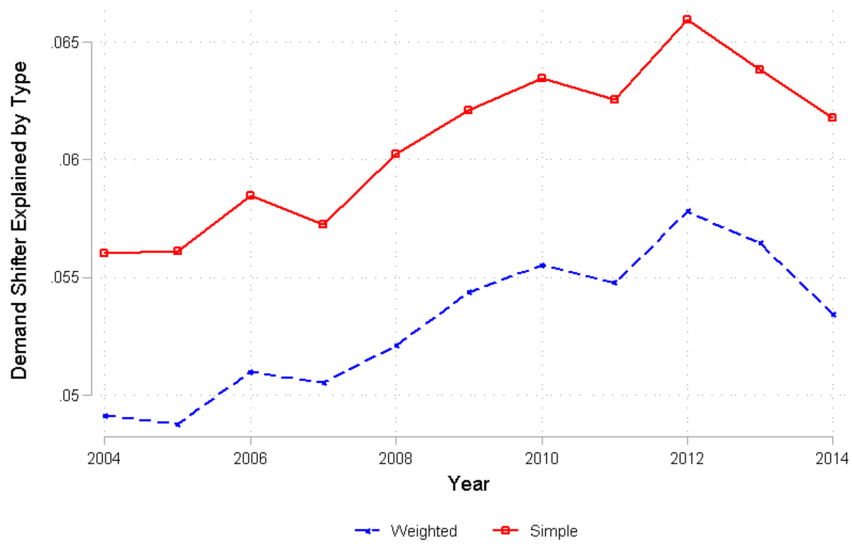
6.1 Variance Decomposition

To understand the quantitative contribution of appeal heterogeneity across segments to the UPC specific demand shifters, we follow the strategy used in the Hottman et al. (2016) and Bernard et al. (2021) and decompose the variance of the UPC demand shifter $\ln z_{ifgn}$ into the common component $\ln \bar{\phi}_{ifg}$ and the segment-specific component $\ln \phi_{ifgn}$. Since the log of the demand shifter $\ln z_{ifgn}$ is the sum of the two components $\ln \bar{\phi}_{ifg} + \ln \phi_{ifgn}$, the variance decomposition can be done exploiting the OLS properties and separately regress $\ln \bar{\phi}_{ifg}$ and $\ln \phi_{ifgn}$ on the demand shifter $\ln z_{ifgn}$. The coefficients from such regressions represent the contribution of each of the two components in explaining the variance of the demand shifters, and the two coefficients sum to one.

We run the variance decomposition by product group and year and report the average result across product groups in Figure 6, using a simple average or the weighted average, where the weights are the expenditure shares on product groups. The average explanatory power of the segment-specific component $\ln \phi_{ifgn}$ ranges between 5% to 6.5%.[22] There is

[22] In a robustness exercise, in which we consider a division into segments by state, retail chain, and consumer type (see Appendix A.5), we find that the segment-specific demand shifter accounts for approximately 39% of the variance of demand shifters. About 24% of the variance is due to heterogeneity in demand shifters

Figure 6: Contribution of Segments $\ln \phi_{ifgn}$ to Explain Product Appeal



significant heterogeneity across product groups: for crackers, the explanatory power of the segment demand shifter is almost 15%, while for yogurt it is around 2% (see Figure A.2).

On average, the explanatory power of the segment-specific component of the demand shifter is increasing over time. This indicates that the dispersion of demand shifters around the mean across segment has been increasing over time. In other words, segments have grown more heterogeneous in the years considered. There are several explanation for this increasing consumer heterogeneity that have been proposed in the literature. Neiman and Vavra (2020) argue the rise in "niche" products is explained by a large increase in the variety available to consumers, combined with a small increase in the fixed costs of consumers learning about each product.[23] An alternative explanation from Brand (2021) is that firms have adopted practices (such as improved supply-chain management) that enable them to offer more products that are targeted to particular consumers. We will explore the second of these explanations, by empirically describing the margins over which firms can target consumers.

across retail chains, while heterogeneity across states and consumer types account for 8-7%. Hence, focusing only on the division across segments by consumer type is likely to yield conservative estimates for the effects of consumer heterogeneity on markups, profits and other firm performance variables.

[23] Neiman and Vavra (2020) are very clear, however, that the total variety available to consumers cannot be determined by counting product in the Nielsen's data, and instead, the rise in variety is implied by fitting their structural model to other observed data.

6.2 Firm Personalization Strategies

In the presence of consumer heterogeneity, firms of different sizes may adopt different strategies in terms of targeting, or "personalization", strategies. There are two key margins to capture. First, how differently do consumer segments perceive the products of a firm? This is the *intensiveness of personalization* measured by how different the demand shifters across segments for a particular UPC are. It reflects the ability of a firm to produce varieties that are perceived differently across consumer segments. Let $\text{var}\phi_{ifgn}$ be the variance of the segment component of the demand shifter of a product across consumer segments. We measure the intensiveness of personalization as:

$$IP_{fgt} = \text{median}_i\{\text{var}\phi_{ifgn}\},$$

which is the median level of the within-product variance of ϕ_{ifgn}. In the extreme case in which all firm products have the same demand shifters across segments as in the *RepC* model, IP_{fg} assumes the minimum value of 0. Larger values of IP_{fg} imply that firm products are perceived differently by different segments.

How are products targeted to different consumer segments? This is the *extensiveness of personalization*, which captures the ability of a firm to design products that appeal to the taste of different consumer segments. In particular, for each UPC item i, we let $n(i)$ denote the consumer segment that a firm is targeting, by which we mean that $z_{ifn(i)} > z_{ifm}$ for all $m = 1, ..., N$ with $m \neq n(i)$.[24] In other words, in terms of appeal, such a UPC i ranks first in the segment $n(i)$. We measure the extensiveness of personalization as:

$$EP_{fg} = \frac{\#\text{ of segments with at least one product that ranks first}}{\#\text{ of consumer segments}},$$

where EP_{fg} ranges from $1/N$ to 1. If all firm products rank first in the same segment, the firm is only targeting one segment, and thus $EP_{fg} = 1/N$. If a firm has at least one product that ranks first in all segments, such a firm reaches the maximum level of $EP_{fg} = 1$. Small values of EP_{fg} do not imply that some segments do not purchase any products from firm f, but just that most products' demand shifters are the highest for only a few segments.

The two concepts of firm personalization are related to firm size. In particular, in Figures A.4 and A.5 of the Appendix, we plot the coefficients obtained from a regression of IP_{fg} and EP_{fg} on firm sales and scope within a product group. In terms of intensiveness of

[24] We implicitly assume that the demand shifter of the targeted segment is different from the demand shifter for the same UPC for all other segments. Hence, one UPC can only target one and one segment only. Given our empirical methodology for measuring the demand shifters (described in section 4), this assumption is always satisfied.

personalization, we find a statistically insignificant relationship between firm size and IP_{fg} for most product groups. In contrast, we find a positive and statistically significant coefficient for all product groups between extensiveness of personalization and firm size. The robust pattern indicates that larger firms target more segments than smaller firms.[25]

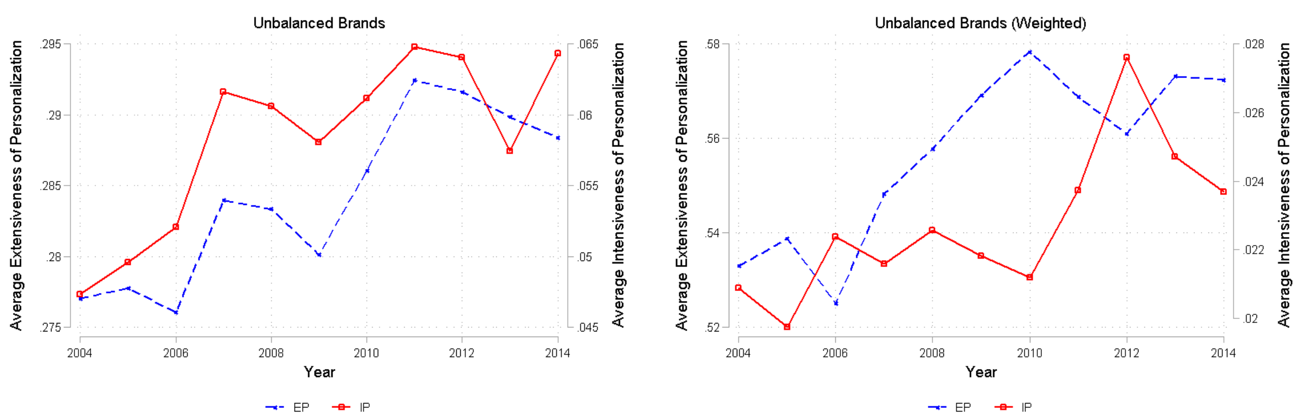
In order to study how the two measures have evolved over time, in Figure 7, we plot the average IP_{fg} and EP_{fg} across product groups over time. We consider both a simple average (panel (a)) and a weighted average (panel (b)). Both measures exhibit an upward trend from 2004 to 2014. This indicates that, on average, products of the same firm have in later years been perceived more differently than in the early part of the sample. In other words, firms have started to offer more niche products than mainstream products. Furthermore, firms have expanded the number of segments they target.

The rise in the intensiveness and extensiveness of personalization is a possible determinant of the rise in the profit share. In fact, an increase in IP_{fg} is indicative of a rise in the dispersion of demand shifters at the firm level, and we have shown that consumer heterogeneity leads to higher markups. Furthermore, the rise in the EP_{fg} points to an increased ability of firms to exploit the consumer heterogeneity, by targeting more segments and, thus, limiting the within-segment cannibalization of products. Although the relationship between the two measures and the aggregate profit share is imperfect, since the aggregate profit share is flat until 2011, we find that in the years in which the profit share increases, both IP_{fg} and EP_{fg} exhibit their largest values.

Figure 7: Firm Personalization Strategies Over Time

(a) Simple Average

(b) Weighted Average



[25] In Appendix A.3, we report the relationship between the intensiveness and extensiveness of personalization and firm using alternative segmentation criteria. We also document a positive relationship between the share of products firms use to target a segment and the segment's size. For details see Appendix A.3.1.

6.3 Adding Fixed Costs

In this section, we extend our earlier model to allow for fixed costs for each UPC item. We think of these fixed costs as reflecting marketing costs for targeting an item to a particular consumer segment. As in the previous section, we let $n(i)$ denote the consumer segment with the highest demand parameter for item i, meaning that $z_{ifn(i)} > z_{ifm}$ for all $m = 1, ..., N$ with $m \neq n(i)$. each consumer segment. We assume that targeting a particular segment n is associated with product design and marketing costs FC_n, which can differ across n because some segments might be easier to reach, or appeal to, than others. For each UPC i, we then denote the fixed cost of targeting that associated consumer segment by $FC_{n(i)}$. In addition, there can be any other fixed costs associated with firm f that we denote by FC_f.

Including the fixed marketing costs, the profits from UPC i are written as:

$$\tilde{\pi}_{if} = \sum_{n=1}^{N} L_n y_n s_{fn} s_{ifn} \left(1 - \frac{1}{\mu_{if}}\right) - FC_{n(i)}. \tag{18}$$

Firm f's total profits over all its UPCs are then:

$$\tilde{\pi}_f(I_f) = \sum_{i \in I_f} \tilde{\pi}_{if}(I_f) - FC_f, \tag{19}$$

where FC_f are any additional fixed costs of the firm.

Note that the profits $\tilde{\pi}_{if}(I_f)$ appearing on the right of (19) refer only to the profits earned on UPC item i, but they depend on the entire set of products I_f sold by firm f, as well as on the set of products sold by other firms. This dependence means that there are multiple equilibrium in the optimal set of products offered by each firm, i.e. what one firm chooses to offer depends on what other have already offered. Regardless of this multiplicity, the set of products sold by each firm in equilibrium must satisfy three conditions. First, profits must be non-negative, so $\tilde{\pi}_f(I_f) \geq 0$ in (19). Second, profits cannot increase by dropping a product from the set being produced and sold (with no change in the products sold by other firms):

$$\tilde{\pi}_f(I_f) \geq \tilde{\pi}_f(I_f/\{i\}) \qquad \forall i \in I_f, \tag{20}$$

where $I_f/\{i\}$ denotes the set of products I_f but without item i. Third, profits cannot increase by adding a new product to the set being produced and sold (with no change in the products sold by other firms):

$$\tilde{\pi}_f(I_f) \geq \tilde{\pi}_f(I_f \cup \{j\}) \qquad \forall j \notin I_f.$$

It is convenient to express the second of these conditions using variable profits π_{if}, so that $\tilde{\pi}_{if}(I_f) = \pi_{if}(I_f) - FC_{n(i)}$. Then it follows from (19) that (20) is rewritten as:

$$\pi_{if}(I_f) - \sum_{j \in I_f/\{i\}} \left[\pi_{jf}(I_f/\{i\}) - \pi_{jf}(I_f)\right] \geq FC_{n(i)} \qquad \forall i \in I_f. \tag{21}$$

To interpret this expression, $\pi_{if}(I_f)$ is the variable profits earned from item i. The next term on the left reflects the *drop* in profits obtained on the other other $I_f/\{i\}$ items already sold by the firm from adding i and cannibalizing their sales, so that $\left[\pi_{jf}(I_f/\{i\}) - \pi_{jf}(I_f)\right] \geq 0$.[26] Using this inequality in (21), we obtain:

$$FC_{n(i)} \leq \pi_{if}(I_f). \tag{22}$$

Expression (22) states that an upper-bound on the marketing and design costs associated with segment n are the variable profits earned on each UPC item i targeting that segment. Let us denote the total set of UPCs that target segment n by I_n. We have allowed that the costs of targeting consumer segment $n(i)$ to depend on the characteristics of that set of households, but not on the item i itself. It follows that we can tighten the inequality in (22) as

$$FC_n \leq min_{i \in I_n} \left\{\pi_{if}(I_f)\right\}. \tag{23}$$

In practice, we estimate these fixed marketing costs by treating the inequality above as an equality. We stress that these marketing costs are only one component of the fixed costs of each firm, which also appear as FC_f in on the right of (19), and we do not attempt to estimate these broader firm-specific fixed costs.[27]

Using the estimated product-firm profits (see section 5), our procedure to estimate fixed marketing costs yields a vector of fixed costs (one per segment) for each product group in each year. For each product group in each year, we compute the average fixed marketing cost and the minimum and maximum across segments within the product group. We normalize these fixed costs by dividing them by the average fixed cost in 2004 for a given product group. Then, we take the average of these three measures across all product groups for each

[26] In the Appendix, we also consider the cannibalization effects that firms face. Following (Hottman et al., 2016), we focus on cannibalization rates, namely the elasticity of demand of a segment for the last UPC with respect to the number of UPCs of a firm. Intuitively, cannibalization rates vary across segments. For each firm, we compute the difference between the maximum cannibalization rate and its minimum, across segments, finding that the heterogeneity across segments for the largest firms is sizable, while for the smallest firms it is negligible.

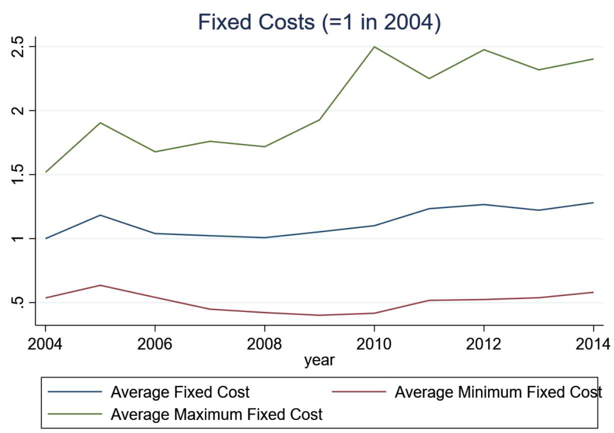
[27] While the resulting estimate of $F_{n(i)}$ from treating (23) as an equality constitutes an upper bound for the fixed marketing cost of targeting the segment, in practice we find that this upper bound is still quite small, simply because the minimum of the variable profits on the right of (23) is small. If we deduct these fixed costs $F_{n(i)}$ from the variable profits of the firm, we find that the profit shares computed as in section 5 differ only in the third significant digit.

year and plot the results in Figure 8. In 2004, the minimum fixed cost is 50% of the average and the maximum fixed cost is 50% larger than the average.

With the exception of a spike in 2005, the estimated average fixed costs remain fairly constant until 2011. In 2011, the average fixed cost increases, attaining a value which is 20% larger than its value in 2004. This increase in the fixed cost corresponds to the rise in the profit share. The relationship is intuitive: higher fixed costs associated with targeting segments lead to higher concentration and, thus, higher profit share. A similar pattern arises when we examine the minimum fixed cost (averaged over product groups) which increases by 24% from 2010 to 2011. The (average) maximum fixed cost also increases in the period considered although most of the growth occurs a year earlier, from 2009 to 2010.

Figure 8: Fixed Marketing Costs

(a) Expenditure Share



We also consider the cross-section variation of fixed costs across segments for the year 2012. In Figure A.7 of the Appendix, we plot the bounds for the fixed costs against the expenditure share of segments and the number of firms that have at least one product targeting the segment. We find a negative relationship: the marketing costs associated with targeting larger segments are lower than those for targeting niche segments. The result suggests that firms predominantly target those segments that have the lowest marketing costs. As a result, these segments are also the ones with the largest total revenues. Furthermore, the results rationalize the cross-sectional findings on the extensiveness of personalization. Smaller firms can only cover low marketing costs for the largest segments. Larger firms, in contrast, are able to afford the payment of larger marketing costs and thus can target smaller, niche segments, which tend to have larger marketing costs.

7 Conclusions

This paper studies how the interaction between large firms and consumer heterogeneity can contribute to the rising profit share. We build a model with oligopolistic multi-product firms, which sells to segments that exhibit different demand shifters. Our model predicts that the presence of consumer heterogeneity increases firm-level markups and profits.

We apply our model to the Nielsen dataset on household purchases at the barcode-level, and estimate product-segment-specific demand shifters for a large number of product groups. We find that the profit share has increased from 2004 to 2014, and we identify several channels that drove the increase in the profit share.

First, we quantify the contribution of segment-specific product appeal in explaining the total variation in product appeal. We find that the explanatory power of the segment-specific demand shifter has been increasing over time, indicating that preferences of segments have become more heterogeneous. This finding is consistent Neiman and Vavra (2020), who argue that consumers have increasing focused their purchases on “niche” products, though these authors do not predict any impact on the aggregate markup. Our results point to a stronger role for consumer heterogeneity in explaining the rising share of profits for the United States than has previously been found.

Second, we argue that changes in segment-specific demand shifters can be driven by firms’ strategies in targeting, or personalizing, their products to a given segment. A product targets a segment if it has the largest demand shifter in that segment relative to all other segments. We examine how *intensively* firms personalize their products for the taste of given segments, i.e. how differently are the products of a firm perceived by different segments. The ability of a firm to target segments also has an *extensive* component, i.e. the number of segments targeted. We find that both measures are increasing over time, indicating that firms have offered relatively more niche products, which are better able to cater to the tastes of particular segments, and that firms have expanded the number of segments they target.

Finally, using the structure of our model, we are able to estimate an upper bound for the segment-specific fixed costs associated with targeting a segment. Such fixed costs can capture the marketing costs associated with placing a product for a particular segment, such as brands choosing famous athletes for their campaigns. Our estimate indicates a rise in the fixed costs across segments, especially in the last three years. This finding is consistent with Brand (2021), who argues that firms have adopted strategies such as improved supply-chain management to offer more products targeted at particular consumer types, and more generally, with the adoption of new fixed-cost technologies in service industries (Hsieh and Rossi-Hansberg, 2021).

References

R. Adão, C. Arkolakis, and S. Ganapati. Aggregate Implications of Firm Heterogeneity: A Nonparametric Analysis of Monopolistic Competition Trade Models. *NBER Working Paper 28081*, 2020.

C. Arkolakis, A. Costinot, D. Donaldson, and A. Rodríguez-Clare. The Elusive Pro-Competitive Effects of Trade. *The Review of Economic Studies*, 86(1):46–80, 2019.

C. Arkolakis, S. Ganapati, and M.-A. Muendler. The Extensive Margin of Exporting Products: A Firm-level Analysis. *forthcoming in American Economic Journal: Macroeconomics*, 2021.

A. Atkeson and A. Burstein. Pricing-to-Market, Trade Costs, and International Relative Prices. *American Economic Review*, 98(5):1998–2031, 2008.

D. Atkin. Trade, Tastes, and Nutrition in India. *American Economic Review*, 103(5):1629–63, 2013.

D. Autor, D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen. The Fall of the Labor Share and the Rise of Superstar Firms. *The Quarterly Journal of Economics*, 135(2): 645–709, 2020.

B. Y. Aw-Roberts, Y. Lee, and H. Vandenbussche. Decomposing Firm-Product Appeal: How important is Consumer Taste? CEPR Discussion Papers 12707, C.E.P.R. Discussion Papers, 2018.

R. Baldwin and J. Harrigan. Zeros, Quality, and Space: Trade Theory and Trade Evidence. *American Economic Journal: Microeconomics*, 3(2):60–88, 2011.

S. Barkai. Declining Labor and Capital Shares. *The Journal of Finance*, 2020.

A. B. Bernard, E. Dhyne, G. Magerman, K. Manova, and A. Moxnes. The Origins of Firm Heterogeneity: A Production Network Approach. *forthcoming in Journal of Political Economy*, 2021.

B. S. Blum, S. Claro, I. Horstmann, and D. A. Rivers. The ABCs of Firm Heterogeneity: The Effects of Demand and Cost Differences on Exporting. Technical report, Working paper, University of Toronto, Toronto, 2018.

J. Brand. Differences in Differentiation: Rising Variety and Markups in Retail Food Stores. *University of Texas at Austin Working Paper*, 2021.

J. A. Caballero. Do Surges in International Capital Inflows Influence the Likelihood of Banking Crises? *The Economic Journal*, 126(591):281–316, 2014.

T. Chaney. The Intensive and Extensive Margins of International Trade. *American Economic Review*, 98(4):1707–1721, 2008.

Y.-L. Chyi, Y.-M. Lai, and W.-H. Liu. Knowledge Spillovers and Firm Performance in the High-technology Industrial Cluster. *Research Policy*, 41(3):556–564, 2012.

A. K. Coşar, P. L. Grieco, S. Li, and F. Tintelnot. What Drives Home Market Advantage? *Journal of International Economics*, 110(Supplement C):135 – 150, 2018.

T. M. Crone and A. Clayton-Matthews. Consistent Economic Indexes for the 50 States. *Review of Economics and Statistics*, 87(4):593–603, 2005.

M. Crozet, K. Head, and T. Mayer. Quality Sorting and Trade: Firm-level Evidence for French Wine. *The Review of Economic Studies*, 79(2):609–644, 2012.

S. DellaVigna and M. Gentzkow. Uniform Pricing in U.S. Retail Chains*. *The Quarterly Journal of Economics*, 134(4):2011–2084, 2019.

F. Di Comite, J.-F. Thisse, and H. Vandenbussche. Verti-zontal differentiation in export markets. *Journal of International Economics*, 93(1):50–66, 2014.

M. Draganska, K. Seim, and M. Mazzeo. Beyond Plain Vanilla: Modeling Joint Product Assortment and Pricing Decisions. *Quantitative Marketing and Economics*, 7, 11 2007.

A. Drewnowski, S. A. Henderson, and A. B. Shore. Taste responses to naringin, a flavonoid, and the acceptance of grapefruit juice are related to genetic sensitivity to 6-n-propylthiouracil. *The American Journal of Clinical Nutrition*, 66(2):391–397, 1997.

C. Eckel and J. P. Neary. Multi-Product Firms and Flexible Manufacturing in the Global Economy. *Review of Economic Studies*, 77(1):188–217, 2010.

M. Elsby, B. Hobijn, and A. Sahin. The Decline of the U.S. Labor Share. *Brookings Papers on Economic Activity*, 44(2 (Fall)):1–63, 2013.

B. Everitt. *Cluster Analysis*. London: Edward Arnold, 1993.

B. Faber and T. Fally. Firm Heterogeneity in Consumption Baskets: Evidence from Home and Store Scanner Data. *accepted at Review of Economic Studies*, 2021.

P. Fajgelbaum, G. M. Grossman, and E. Helpman. Income Distribution, Product Quality, and International Trade. *Journal of Political Economy*, 119(4):721 – 765, 2011.

R. Feenstra. *Advanced International Trade*. Princeton: Princeton University Press, 2004.

R. C. Feenstra. Restoring the Product Variety and Pro-competitive Gains from Trade with Heterogeneous Firms and Bounded Productivity. *Journal of International Economics*, 110:16–27, 2018.

R. C. Feenstra and J. Romalis. International Prices and Endogenous Quality. *The Quarterly Journal of Economics*, 129(2):477–527, 2014.

R. C. Feenstra, M. Xu, and A. Antoniades. What is the Price of Tea in China? Towards the Relative Cost of Living in Chinese and US Cities. *The Economic Journal*, 2020.

A. M. Fernandes, P. J. Klenow, S. Meleshchuk, M. D. Pierola, and A. Rodríguez-Clare. The Intensive Margin in Trade: How Big and How Important. *NBER Working Paper 25195*, 2018.

L. Foster, J. Haltiwanger, and C. Syverson. Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review*, 98(1):394–425, 2008.

A. R. Ghosh, M. S. Qureshi, J. I. Kim, and J. Zalduendo. Surges. *Journal of International Economics*, 92(2):266–285, 2014.

X. Guan, M. Mantrala, and Y. Bian. Strategic Information Management in A Distribution Channel. *Journal of Retailing*, 95(1):42–56, 2019.

J. C. Hallak and J. Sivadasan. Product and Process Productivity: Implications for Quality Choice and Conditional Exporter Premia. *Journal of International Economics*, 91(1): 53–67, 2013.

C. Hottman, S. J. Redding, and D. E. Weinstein. Quantifying the Sources of Firm Heterogeneity. *The Quarterly Journal of Economics*, 131:1291—1364, 2016.

C.-T. Hsieh and E. Rossi-Hansberg. The Industrial Revolution in Services. *NBER Working Paper 25968*, 2021.

I. C. Jäkel. Product appeal, Differences in Tastes, and Export Performance: Evidence for Danish Chocolate and Confectionery. *International Journal of Industrial Organization*, 63:417 – 459, 2019.

R. C. Johnson. Trade and Prices with Heterogeneous Firms. *Journal of International Economics*, 86(1):43–56, 2012.

L. Karabarbounis and B. Neiman. The Global Decline of the Labor Share. *The Quarterly Journal of Economics*, 129(1):61–103, 2014.

A. Khandelwal. The Long and Short (of) Quality Ladders. *The Review of Economic Studies*, 77(4):1450–1476, 2010.

D. Koh, R. Santaeulàlia-Llopis, and Y. Zheng. Labor Share Decline and Intellectual Property Products Capital. *Econometrica*, 88(6):2609–2628, 2020.

M. Kugler and E. Verhoogen. Prices, Plant Size, and Product Quality. *The Review of Economic Studies*, 79(1):307–339, 2012.

E. Levy-Yeyati and F. Sturzenegger. Classifying Exchange Rate Regimes: Deeds vs. Words. *European Economic Review*, 49(6):1603–1635, 2005.

L. Macedoni and M. Xu. Flexibility and Productivity: Towards the Understanding for Multi-product Exporters. *Aarhus University Working Paper*, 2018.

B. Neiman and J. S. Vavra. The Rise of Niche Consumption. *NBER Working Paper 26134*, 2020.

A. Nevo. Empirical Models of Consumer Behavior. *Annual Review of Economics*, 3(1):51–75, 2011.

M. J. Roberts, D. Yi Xu, X. Fan, and S. Zhang. The Role of Firm Factors in Demand, Cost, and Export Market Selection for Chinese Footwear Producers. *The Review of Economic Studies*, 85(4):2429–2461, 2018.

M. Rognlie. Deciphering the Fall and Rise in the Net Capital Share: Accumulation or Scarcity? *Brookings Papers on Economic Activity*, 2015(1):1–69, 2016.

I. Simonovska. Income Differences and Prices of Tradables: Insights from an Online Retailer. *The Review of Economic Studies*, 82(4):1612–1656, 2015.

M. Smith, D. Yagan, O. Zidar, and E. Zwick. Capitalists in the Twenty-first Century. *The Quarterly Journal of Economics*, 134(4):1675–1745, 2019.

A Appendix

A.1 Model Derivations

A.1.1 Profits

Revenues for a firm f from UPC i in segment n are given by:

$$L_n p_{ifn} q_{ifn} = z_g L_n w_n P_n^{\sigma-1} z_f^{\sigma-1} P_{fn}^{\eta-\sigma} z_{ifn}^{\eta-1} p_{if}^{1-\eta} = z_g L_n w_n \left(\frac{\frac{P_{fn}}{z_f}}{P_n} \right)^{1-\sigma} \left(\frac{\frac{p_{if}}{z_{ifn}}}{P_{fn}} \right)^{1-\eta},$$

where z_g denotes the demand shifter for good g, but for convenience we omit the subscript g on all other variables and parameters. Summing over the varieties sold to segment n yields the following segment-specific revenues:

$$\sum_{i \in I_f} L_n p_{ifn} q_{ifn} = z_g L_n w_n \left(\frac{\frac{P_{fn}}{z_f}}{P_n} \right)^{1-\sigma} \sum_{i \in I_f} \left(\frac{\frac{p_{if}}{z_{ifn}}}{P_{fn}} \right)^{1-\eta} = z_g L_n w_n \left(\frac{\frac{P_{fn}}{z_f}}{P_n} \right)^{1-\sigma}.$$

Let $Lw^N L_n w_n$ denote total income of all consumers, and $\alpha_n = L_n w_n / \sum_{n=1}^N L_n w_n$ denote the share of income for consumers in segment n. Firm f total revenues R_f are given by:

$$R_f = \sum_{n=1}^N \sum_{i \in I_f} L_n p_{ifn} q_{ifn} = z_g L w \sum_{n=1}^N \alpha_n z_f^{\sigma-1} \left(\frac{P_{fn}}{P_n} \right)^{1-\sigma}. \tag{24}$$

We also keep track of the total variable costs of a firm, which are:

$$TC_f = \sum_{n=1}^N \sum_{i \in I_f} L_n c_{if} q_{ifn} = z_g z_f^{\sigma-1} L w \sum_{n=1}^N \alpha_n P_n^{\sigma-1} P_{fn}^{\eta-\sigma} \sum_{i \in I_f} z_{ifn}^{\eta-1} c_{if} p_{if}^{-\eta}.$$

A.1.2 First Order Condition

To simplify the derivations of the first-order condition, let us consider the derivative of revenues with respect to prices p_{if}. In Bertrand competition, firms take into account the effects of changing prices both on the firm and segment-specific price index P_{fn}, and on the product group and segment-specific price index P_n:

$$\frac{\partial R_f}{\partial p_{if}} = z_g L w \sum_{n=1}^N \alpha_n z_f^{\sigma-1} \left(\frac{P_{fn}}{P_n} \right)^{1-\sigma} (1-\sigma) \left[\frac{\partial P_{fn}}{\partial p_{if}} \frac{1}{P_{fn}} - \frac{\partial P_n}{\partial P_{fn}} \frac{P_{fn}}{P_n} \frac{\partial P_{fn}}{\partial p_{if}} \frac{1}{P_{fn}} \right].$$

Using the definitions of market shares in (7) and (8), we obtain:

$$\frac{\partial R_f}{\partial p_{if}} = z_g L w \sum_{n=1}^N \alpha_n s_{fn} (1-\sigma) \left[\frac{s_{ifn}}{p_{if}} - \frac{s_{ifn}}{p_{if}} s_{fn} \right] = (1-\sigma) \frac{z_g L w}{p_{if}} \sum_{n=1}^N \alpha_n s_{fn} s_{ifn} (1 - s_{fn}).$$

Taking the derivative of total costs with respect to p_{if} yields:

$$
\begin{aligned}
\frac{\partial TC_f}{\partial p_{if}} &= z_g z_f^{\sigma-1} Lw \sum_{n=1}^{N} \alpha_n P_n^{\sigma-1} P_{fn}^{\eta-\sigma} \left[(\sigma-1)\frac{s_{ifn}}{p_{if}} s_{fn} + (\eta-\sigma)\frac{s_{ifn}}{p_{if}} \right] \sum_{i \in I_f} z_{ifn}^{\eta-1} c_{if} p_{if}^{-\eta} \\
&+ z_g z_f^{\sigma-1} Lw \sum_{n=1}^{N} \alpha_n P_n^{\sigma-1} P_{fn}^{\eta-\sigma} z_{ifn}^{\eta-1} (-\eta) c_{if} p_{if}^{-\eta-1} \\
&= z_g Lw \sum_{n=1}^{N} \alpha_n \left(\frac{\frac{P_{fn}}{z_f}}{P_n} \right)^{1-\sigma} P_{fn}^{\eta-1} \left[(\sigma-1)\frac{s_{ifn}}{p_{if}} s_{fn} + (\eta-\sigma)\frac{s_{ifn}}{p_{if}} \right] \sum_{i \in I_f} z_{ifn}^{\eta-1} c_{if} p_{if}^{-\eta} \\
&+ z_g Lw \sum_{n=1}^{N} \alpha_n \left(\frac{\frac{P_{fn}}{z_f}}{P_n} \right)^{1-\sigma} P_{fn}^{\eta-1} z_{ifn}^{\eta-1} (-\eta) c_{if} p_{if}^{-\eta-1}.
\end{aligned}
$$

Using the definitions of market shares, we obtain:

$$
\begin{aligned}
\frac{\partial TC_f}{\partial p_{if}} = \frac{z_g Lw}{p_{if}} \sum_{n=1}^{N} \alpha_n s_{fn} \left[(\sigma-1) s_{ifn} s_{fn} + (\eta-\sigma) s_{ifn} \right] \frac{\sum_{i \in I_f} z_{ifn}^{\eta-1} c_{if} p_{if}^{-\eta}}{\sum_{i \in I_f} z_{ifn}^{\eta-1} p_{if}^{1-\eta}} + \\
- \eta \frac{z_g Lw}{p_{if}} \sum_{n=1}^{N} \alpha_n s_{fn} \frac{s_{ifn}}{p_{if}} c_{if}.
\end{aligned}
$$

Using the definition of market shares of UPC within a firm, we can rewrite a term above as:

$$
\frac{\sum_{i \in I_f} z_{ifn}^{\eta-1} c_{if} p_{if}^{-\eta}}{\sum_{i \in I_f} z_{ifn}^{\eta-1} p_{if}^{1-\eta}} = \frac{\sum_{i \in I_f} s_{ifn} \frac{c_{if}}{p_{if}}}{\sum_{i \in I_f} s_{ifn}} = \sum_{i \in I_f} s_{ifn} \frac{c_{if}}{p_{if}}.
$$

It follow that the derivative of firm total costs is equal to:

$$
\frac{\partial TC_f}{\partial p_{if}} = \frac{z_g Lw}{p_{if}} \left[\sum_{n=1}^{N} \alpha_n s_{fn} s_{ifn} \left[(\sigma-1) s_{fn} + (\eta-\sigma) \right] \sum_{i \in I_f} s_{ifn} \frac{c_{if}}{p_{if}} - \eta \sum_{n=1}^{N} \alpha_n s_{fn} s_{ifn} \frac{c_{if}}{p_{if}} \right].
$$

We therefore obtain the following first-order condition:

$$
\begin{aligned}
&\frac{\partial R_f}{\partial p_{if}} = \frac{\partial TC_f}{\partial p_{if}} \implies \\
(\sigma-1) \sum_{n=1}^{N} \alpha_n s_{fn} s_{ifn} (1 - s_{fn}) &= \eta \sum_{n=1}^{N} \alpha_n s_{fn} s_{ifn} \frac{c_{if}}{p_{if}} \\
&- \sum_{n=1}^{N} \alpha_n s_{fn} s_{ifn} \left[(\sigma-1) s_{fn} + (\eta-\sigma) \right] \sum_{i \in I_f} s_{ifn} \frac{c_{if}}{p_{if}}.
\end{aligned}
$$

It is convenient to rewrite the equation as a function of markups $\mu_{if} = p_{if}/c_{if}$:

$$(\sigma-1)\sum_{n=1}^{N}\alpha_n s_{fn}s_{ifn}(1-s_{fn}) = \frac{\eta}{\mu_{if}}\sum_{n=1}^{N}\alpha_n s_{fn}s_{ifn} - \sum_{n=1}^{N}\alpha_n s_{fn}s_{ifn}\left[(\sigma-1)s_{fn}+(\eta-\sigma)\right]\sum_{i\in I_f}\frac{s_{ifn}}{\mu_{if}}$$

Notice that $\sum_{n=1}^{N}\alpha_n s_{fn}s_{ifn} = s_{if}$ equals the market share of UPC i of firm f within the product group. We divide both sides of the equation by $\sum_{n=1}^{N}\alpha_n s_{fn}s_{ifn}$, to obtain the expression (10) shown in the main text.

A.1.3 Proof of Proposition 1

Proposition 1 is stated for all product groups $g = 1, ..., G$, but in this proof we drop the subscript g. Jensen's inequality states that $\sum_{n=1}^{N}\alpha_n s_{fn}^2 \geq \left(\sum_{n=1}^{N}\alpha_n s_{fn}\right)^2$, and this inequality holds as an equality if and only if $s_{fn} = s_{fm}\ \forall\ m, n$. So the inequality is strict if $s_{fn} \neq s_{fm}$ for some firm f and segments $m \neq n$. Let us consider the ratio s_{fn}/s_{fm} from (5) and (7), using the fact that the firm shifter z_f (appearing as z_{fg} in (5)) cancels because s_{fn} and s_{fm} are for the same firm f:

$$\frac{s_{fn}}{s_{fm}} = \left(\frac{P_{fn}/P_n}{P_{fm}/P_m}\right)^{1-\sigma} = \left(\frac{z_{ifm}}{z_{ifn}}\frac{P_m}{P_n}\right)^{1-\sigma}, \tag{25}$$

where the second equality follows from (6) and using the assumption of a single product firm in the *HetC* model, so that $I_{fn} = \{i\}$. We need to show that $s_{fn} \neq s_{fm}$ for two consumer segments m and n and some firm f to ensure that Jensen's inequality holds strictly.

Suppose to the contrary that $s_{fn} = s_{fm}$ in (25), which implies that $P_m/P_n = z_{ifn}/z_{ifm}$. Then we make use $z_{jf'm}/z_{jf'n} \neq z_{ifm}/z_{ifn}$ from Definition 1, to conclude that

$$1 \neq \left(\frac{z_{jf'm}}{z_{jf'n}}\frac{z_{ifn}}{z_{ifm}}\right)^{1-\sigma} = \left(\frac{z_{jf'm}}{z_{jf'n}}\frac{P_m}{P_n}\right)^{1-\sigma} = \frac{s_{f'n}}{s_{f'm}},$$

where the second equality follows from our assumption that $s_{fn} = s_{fm}$, which implies that $P_m/P_n = z_{ifn}/z_{ifm}$, and the third equality follows by computing $s_{f'n}/s_{f'm}$ just like in (25). So we find that $s_{f'n} \neq s_{f'm}$, which guarantees that Jensen's inequality holds strictly.

A.1.4 Cannibalization Rate

In a slight abuse of notation (because we omit the subscript f), let $\bar{I} \equiv |I_f|$ denote the number of items in the set I_f, and suppose that we rank these items by decreasing demand. Following Hottman et al. (2016), we can compute the cannibalization rate defined as the elasticity of demand for the last UPC, $q_{\bar{I}fn}$, with respect to the number of UPCs $\bar{I}$. Such an elasticity can be derived analytically, assuming that the number of products is large enough to be approximated by a continuum, so that:

$$\frac{\partial P_{fn}}{\partial \bar{I}} \approx \frac{P_{fn} s_{\bar{I}fn}}{1-\eta}.$$

As a result, we can compute the cannibalization rate as follows:

$$\begin{aligned}\frac{\partial q_{\bar{I}fn}}{\partial \bar{I}} &= (\sigma-1)\frac{q_{\bar{I}fn}}{P_n}\frac{\partial P_n}{\partial P_{fn}}\frac{\partial P_{fn}}{\partial \bar{I}} + (\eta-\sigma)\frac{q_{\bar{I}fn}}{P_n}\frac{\partial P_{fn}}{\partial \bar{I}} \\ &= (\sigma-1)\frac{q_{\bar{I}fn}}{P_f n}\frac{\partial P_n}{\partial P_{fn}}\frac{P_{fn} s_{\bar{I}fn}}{1-\eta} + (\eta-\sigma)\frac{q_{\bar{I}fn}}{P_n}\frac{P_{fn} s_{\bar{I}fn}}{1-\eta} \\ &= (\sigma-1)\frac{q_{\bar{I}fn} s_{fn} s_{\bar{I}fn}}{1-\eta} + (\eta-\sigma)\frac{q_{\bar{I}fn} s_{\bar{I}fn}}{1-\eta},\end{aligned}$$

with which we can obtain the cannibalization rate by segment:

$$\frac{\partial q_{\bar{I}fn}}{\partial \bar{I}}\frac{\bar{I}}{q_{\bar{I}fn}} = -\left(\frac{\sigma-1}{\eta-1}s_{fn} + \frac{\sigma-\eta}{\eta-1}\right) s_{\bar{I}fn}\bar{I}. \tag{26}$$

Thus, expanding the scope reduces the demand for an individual UPC from a segment and such a reduction is proportional to the market share of the firm in the niche (s_{fn}), the market share of the UPC in the segment-firm ($s_{\bar{I}fn}$), and the number of UPCs ($\bar{I}$). Following Hottman et al. (2016), let us define the cannibalization rate as (26) evaluated at a particular product that has a market share in the segment equal to the average market share, so that $s_{\bar{I}fn}\bar{I} = 1$ and $s_{\bar{I}fn} = 1/\bar{I}$.

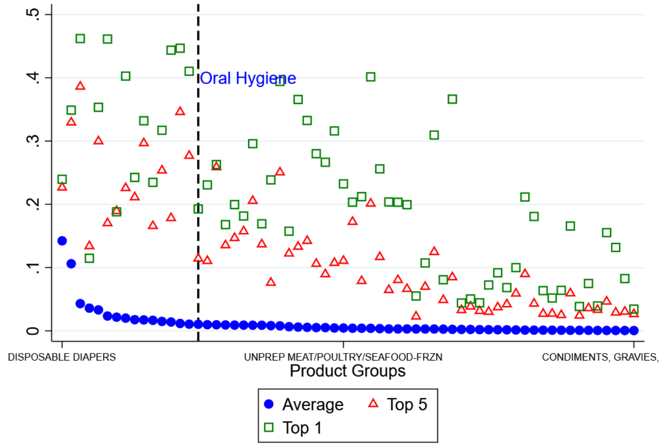
To account for the fact that introducing a new UPC cannibalizes demand across other segments, we compute an average cannibalization rate across segments where we use segment size (α_n) as weights:

$$\sum_{n=1}^{N} \alpha_n \frac{\partial q_{\bar{I}fn}}{\partial \bar{I}}\frac{\bar{I}}{q_{\bar{I}fn}} = -\left(\frac{\sigma-1}{\eta-1}\sum_{n=1}^{N}\alpha_n s_{fn} + \frac{\sigma-\eta}{\eta-1}\right), \tag{27}$$

which has the convenient property of being identical to the formula in Hottman et al. (2016), since $\sum_{n=1}^{N} \alpha_n s_{fn}$ is the market share of the firm. Although the average cannibalization rate is independent of consumer heterogeneity, the rate is heterogeneous across segments, which could be a crucial factor for firms deciding to introduce new products.

We compute the cannibalization rate by segment and the average cannibalization rate for firms defined in equations (26) and (27). Then, for each firm, we compute the difference between the maximum cannibalization rate and its minimum, across segments. In Figure A.1, we report the average difference by product group and the difference for the top one and five firms. For instance, in the oral hygiene case, the difference between the maximum and minimum cannibalization rate is 0.01, which suggests minimal differences across segments. However, when we consider the top five firms and the top one firm, the rate jumps to 0.11 and 0.19. Since the average cannibalization rate is about 0.5 (ranging between 0.23 and 0.75), the heterogeneity across segments for the largest firms is sizable

Figure A.1: Max-Min Difference of Cannibalization Rates across Segments



A.2 Consumer Segments

Table A.1 reports the number of consumer types by product group using our clustering methodology of section 4.1. Table A.2 reports the percentage point difference in markups between the *HetC* and the *RepC* models across product groups.

A.3 Firm Strategy to Target Segments

While in the main text we document how the strategies firms follow to deal with consumer heterogeneity over time, in this section, we consider the year 2012 and examine the heterogeneity across product groups and across firms. In Figure A.2, we plot the results from the variance decomposition discussed in section 6.1 across product groups. In particular, for each product group, we report how much of the variance of the demand shifter is explained by the segment-specific component.

Panel (a) of Figure A.3 reports the coefficients obtained by regressing UPC log sales on the estimated UPC taste parameter in logs common across segments. Panel (b) of Figure A.3 displays the slope of a regression of $\ln \bar{\phi}_{ifg}$ on the variance of $\ln \phi_{ifgn}$ across segments for all product groups. Both figures report the point estimate of the coefficient and the 95% confidence interval, where we control for firm fixed effects in each regression and cluster standard errors at the firm level. The results indicate that products that have larger common component of the demand shifters have larger sales and have lower variance of the segment-specific component.

Figure A.4 reports the coefficient of a regression of IP_{fg} on firm sales (panel (a)) and firm scope (panel (b)). The coefficients are generally close to zero and often statistically insignificant. Figure A.10 reports the coefficient of a regression of EP_{fg} on firm sales (panel (a)) and firm scope (panel (b)). Contrary to the results on the intensiveness of personalization, here we find a positive and statistically significant coefficient for all product groups. The robust pattern indicates that larger firms target more segments than smaller firms.

Table A.1: Number of Customer Types by Product Group

Product Group	Number of Consumer Types
Baby Food	7
Baking Supplies	3
Batteries And Flashlights	7
Beer	7
Bread And Baked Goods	2
Breakfast Food	7
Butter And Margarine	7
Candy	3
Cereal	5
Cheese	2
Coffee	7
Condiments, Gravies, And Sauces	2
Cookies	3
Cosmetics	3
Cot Cheese, Sour Cream, Toppings	7
Cough And Cold Remedies	7
Crackers	5
Deodorant	7
Detergents	2
Disposable Diapers	6
Dressings/Salads/Prep Foods-Deli	3
Eggs	5
Electronics, Records, Tapes	3
First Aid	3
Fresh Meat	2
Fresh Produce	2
Hair Care	2
Household Cleaners	2
Household Supplies	2
Ice Cream, Novelties	2
Jams, Jellies, Spreads	4
Juice, Drinks - Canned, Bottled	2
Kitchen Gadgets	4
Laundry Supplies	3
Liquor	4
Medications/Remedies/Health Aids	2
Milk	2
Nuts	2
Oral Hygiene	5
Packaged Meats-Deli	2
Packaged Milk And Modifiers	4
Paper Products	7
Pet Care	3
Pet Food	3
Pizza/Snacks/Hors D'oeuvres-Frozen	5
Prepared Food-Dry Mixes	2
Prepared Food-Ready-To-Serve	5
Prepared Foods-Frozen	2
Salad Dressings, Mayo, Toppings	2
Shaving Needs	3
Shortening, Oil	7
Snacks	2
Soft Drinks-Non-Carbonated	3
Soup	3
Spices, Seasoning, Extracts	2
Tea	5
Tobacco & Accessories	5
Unprepared Meat/Poultry/Seafood-Frozen	4
Vegetables - Canned	2
Vegetables-Frozen	5
Vitamins	2
Wine	2
Wrapping Materials And Bags	6
Yogurt	2

Table A.2: Percentage Point Difference between *HetC* and *RepC* markup formula

Product Group	Average	Top 5 Firms	Top 1 Firm
Eggs	130.7	205.3	227.1
Butter And Margarine	50.1	74.0	64.4
Shaving Needs	48.1	53.3	80.1
Packaged Milk And Modifiers	40.0	46.2	59.2
Cot Cheese, Sour Cream, Toppings	31.7	55.8	67.7
Disposable Diapers	29.1	29.3	36.2
Tobacco & Accessories	29.1	55.0	94.9
Crackers	24.6	27.4	35.0
Vegetables-Frozen	22.3	30.9	40.3
Coffee	21.8	32.5	39.4
Cereal	21.5	23.9	27.6
Shortening, Oil	21.1	34.5	39.3
Jams, Jellies, Spreads	20.0	26.8	41.5
Baby Food	19.3	21.5	21.8
Wrapping Materials And Bags	19.1	20.5	22.2
Batteries And Flashlights	19.0	20.2	25.9
Soup	15.8	20.5	26.0
Paper Products	14.9	19.8	16.5
Breakfast Food	14.8	18.9	23.3
Tea	11.3	19.9	28.0
Yogurt	11.2	13.1	28.9
Cough And Cold Remedies	10.9	16.9	26.2
Oral Hygiene	10.9	15.7	24.8
Unprep Meat/Poultry/Seafood-Frzn	8.8	22.8	41.0
Salad Dressings, Mayo, Toppings	8.7	11.7	22.2
Prepared Food-Ready-To-Serve	8.3	17.3	18.4
Pizza/Snacks/Hors D'oeuvres-Frozen	8.2	12.7	14.7
Electronics, Records, Tapes	7.8	15.6	26.2
Beer	7.7	14.3	22.1
Pet Food	7.1	9.8	13.0
Deodorant	7.0	8.1	5.1
First Aid	6.1	12.3	26.5
Nuts	5.8	8.4	16.1
Detergents	5.3	9.1	18.1
Cookies	5.2	7.3	11.3
Laundry Supplies	4.0	7.0	17.7
Cheese	3.9	6.6	12.6
Pet Care	3.5	6.6	13.7
Spices, Seasoning, Extracts	3.4	5.8	8.1
Kitchen Gadgets	2.9	7.8	14.9
Baking Supplies	2.8	5.7	4.8
Packaged Meats-Deli	2.8	5.6	11.8
Fresh Produce	2.7	4.6	6.4
Dressings/Salads/Prep Foods-Deli	2.2	5.1	7.6
Candy	1.9	3.1	4.3
Cosmetics	1.4	2.0	4.6
Fresh Meat	1.2	2.0	3.7
Milk	1.2	2.9	3.1
Liquor	0.9	3.0	8.6
Vegetables - Canned	0.8	1.5	3.5
Prepared Food-Dry Mixes	0.8	1.4	2.2
Prepared Foods-Frozen	0.7	1.5	2.9
Household Cleaners	0.7	1.4	2.6
Household Supplies	0.6	1.1	2.5
Ice Cream, Novelties	0.4	1.0	1.5
Soft Drinks-Non-Carbonated	0.4	0.4	0.5
Juice, Drinks - Canned, Bottled	0.3	0.5	1.1
Snacks	0.3	0.8	1.8
Vitamins	0.3	0.7	1.2
Condiments, Gravies, And Sauces	0.3	0.7	0.5
Hair Care	0.2	0.4	0.5
Wine	0.2	0.8	1.8
Medications/Remedies/Health Aids	0.1	0.5	1.1
Bread And Baked Goods	0.1	0.4	0.4

Figure A.2: Contribution of Segments $\ln \phi_{ifgn}$ to Explain Product Appeal

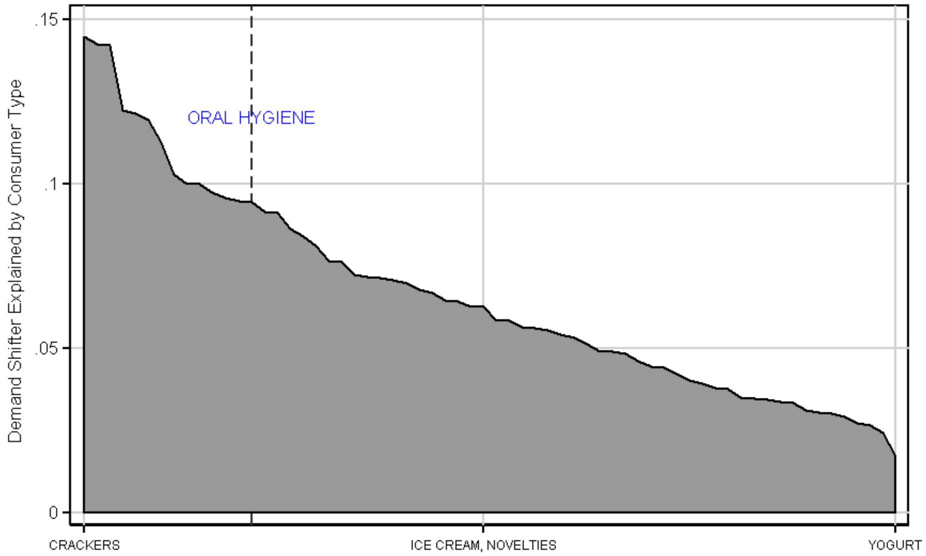
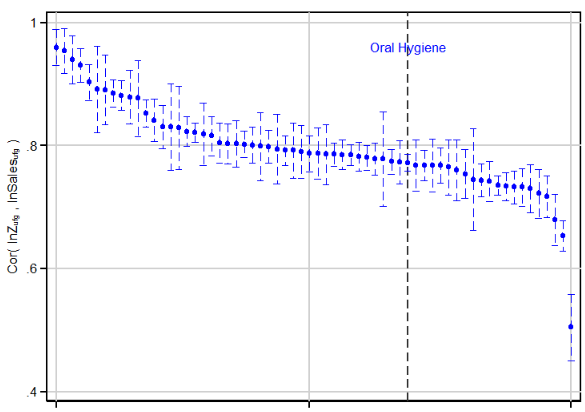


Figure A.3: Demand shifters and Segments (All UPC) - Type (Consumer Panel)

(a) $\bar{\phi}_{ifg}$ and UPC sales



(b) $\bar{\phi}_{ifg}$ and variance of ϕ_{ifgn}

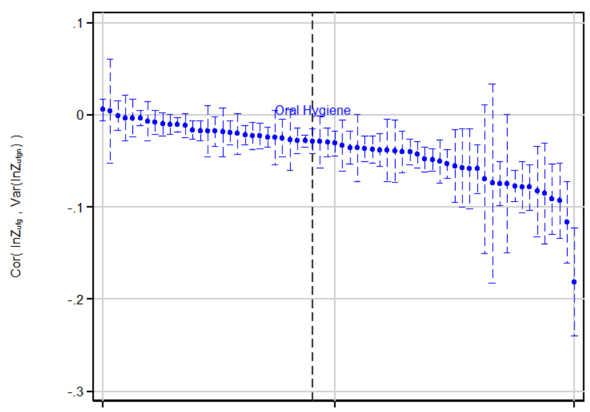


Figure A.4: Intensiveness of Personalization and Firm Characteristics (All UPC) - Type (Consumer Panel Data)

(a) Firm Sales (b) Firm Scope

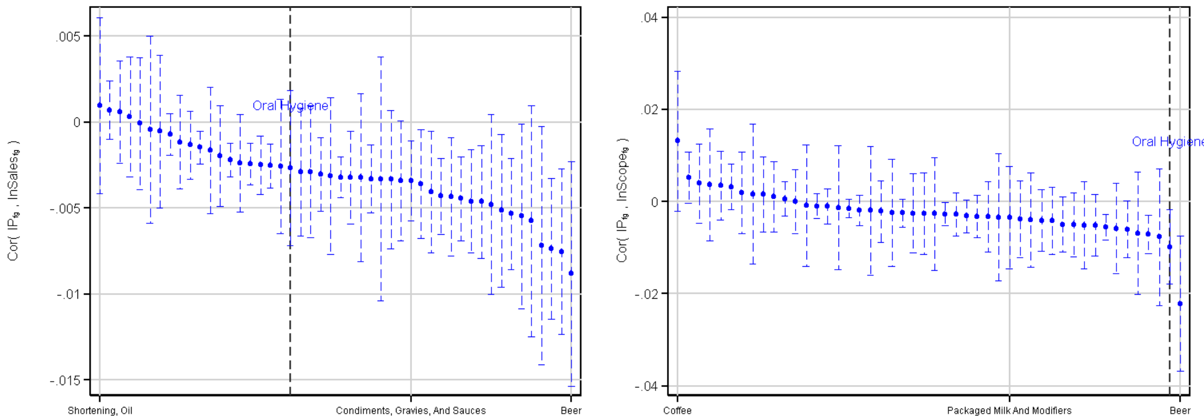
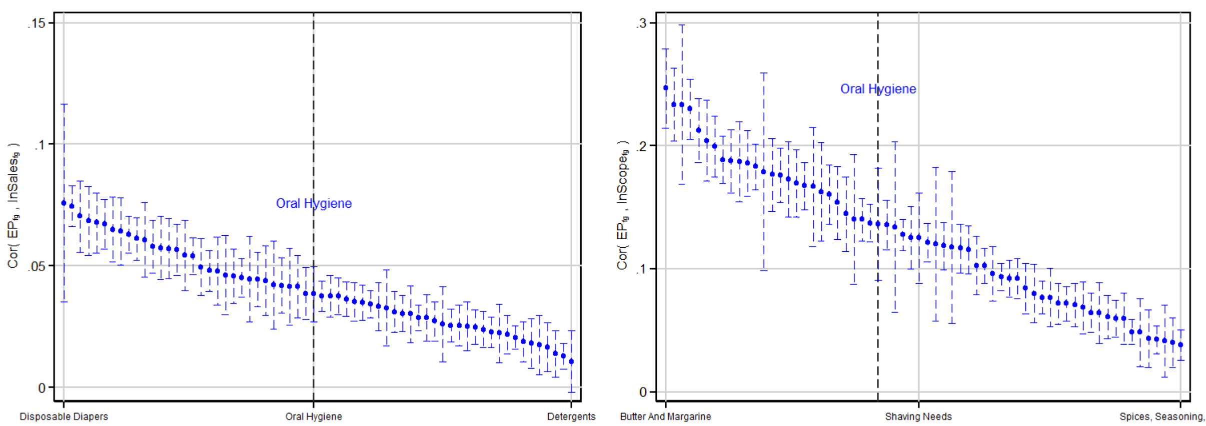


Figure A.5: Extensiveness of Personalization and Firm Characteristics (All UPC) - Type (Consumer Panel Data)

(a) Firm Sales (b) Firm Scope



A.3.1 Firm Personalization Strategies

In this section, we determine the segment characteristics are mostly correlated with firm personalization strategies. Consider the share of best-ranked products of a firm f in segment n:

$$B_{fgn} = \frac{\text{\# best-ranked products of firm f in segment n}}{\text{Total \# of products of firm f}} \tag{28}$$

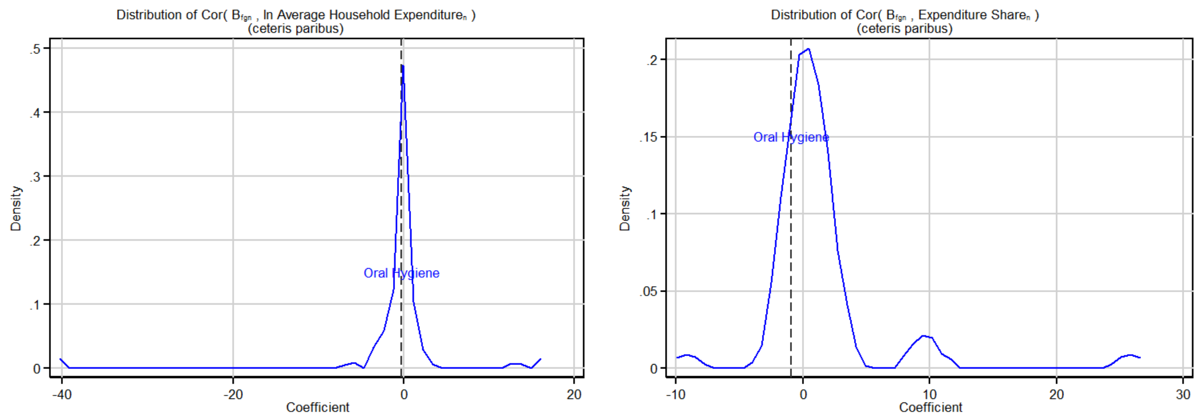
By definition, $B_{fgn} \in [0, 1]$. If a firm has no products targeting a given segment then $B_{fgn} = 0$. In contrast, if all products of a firm target a segment n then $B_{fgn} = 1$. To examine the main drivers of firms sorting into different segments, we run a univariate regression of

B_{fgn} on segment size, defined as the share of total expenditures in the product group, and average expenditure per household.

Figure A.6 shows the distribution of the coefficients of univariate regression of B_{fgn} on the average household expenditures and expenditure share of each segment. On average, the effect of household expenditure is zero. In contrast, the size of a segment measured by its expenditure share tends to have a positive effect on B_{fgn}. This means that firms tend to mainly target the segment with the largest expenditure share.

Figure A.6: Consumer Characteristics and Product Personalization (All UPC) - Type (Consumer Panel Data)

(a) Average Household Expenditure (b) Expenditure Share

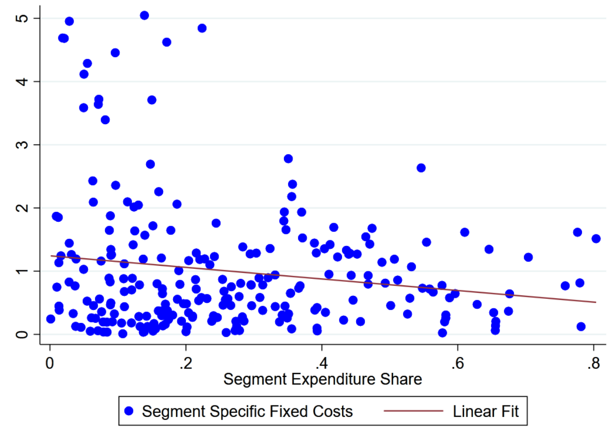


This fact, combined with the positive relationship between extensiveness of personalization and firm size, suggests the presence of a segment ladder: the smaller firms tend to target a smaller number of segments, and these segments tend to be the larger ones. Larger firms are able to expand the number of segments they target, by further targeting the smaller segments. Although examining such a pattern for all product groups can be hard to illustrate, in the Appendix A.6.2 on the oral hygiene product group, we confirm and show such a segment ladder.

A.4 Fixed Costs

Figure A.7 plots the estimated segment-specific fixed cost bounds and the segment-specific expenditure share. We find a negative correlation: the larger segments are those with a smaller fixed cost, which allows a greater number of firms to offer varieties that target these segments.

Figure A.7: Fixed Marketing Costs and Segment Characteristics



A.5 Robustness

For robustness, we also divide consumers into segments based on geography and distribution channel. In the next paragraphs, we describe how we divide consumers into segments according to the two dimensions.

Geography. We consider the division by state, using the home scanner data to divide consumers according to the states in which they reside.

Distribution Channel. We consider the division by retail chain. We assign consumers to a retail chain according to the store belonging to the chain in which they conduct the majority of their purchases. That is, both for the geographical and the distribution channel criteria, the demand from each segment n consists of the average demand of consumers in each geographical area or retail chain when we use the home scanner data.

In Figures A.8, A.9, A.10, and A.11 we use the two criteria described in this section along with consumer type to divide consumers in segments. Hence, one segment is a geography-retail chain-type triplet. Figure A.8 confirms the results of Figure A.3, since we find a positive correlation between sales and the common component of the demand shifter (panel (a)) and a negative correlation between the common component of the demand shifter and the variance of the segment-specific component (panel (b)). In Figure A.9, we plot the coefficient of a regression of IP_{fg} on firm sales (panel (a)) and scope (panel (b)). The results are in line with those shown in Figure A.4, in which consumers are divided in segments only by their type. In Figure A.10, we confirm the results of Figure A.5 as we find a positive correlation between EP_{fg} and firm sales (panel (a)) and scope (panel (b)). The results indicate that larger firms use their products to target different segments, defined at the state-retail chain-type level. Finally, Figure A.11 replicates the results shown in section A.3.1. The figure shows the distribution of the coefficients of a univariate regression of B_{fgn}, which is defined in (28), on the average household expenditures and expenditure share of each segment.

Figure A.8: Demand shifters and Segments

(a) $\bar{\phi}_{ifg}$ and UPC sales (b) $\bar{\phi}_{ifg}$ and variance of ϕ_{ifgn}

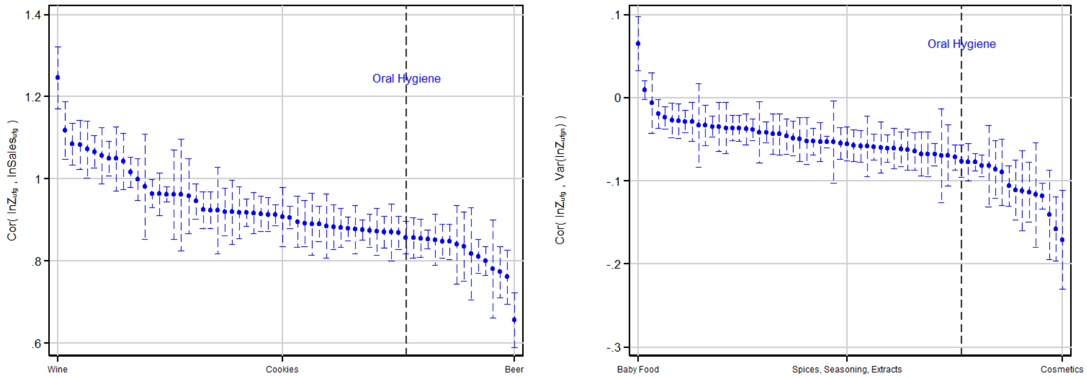


Figure A.9: Intensiveness of Personalization and Firm Characteristics

(a) Firm Sales (b) Firm Scope

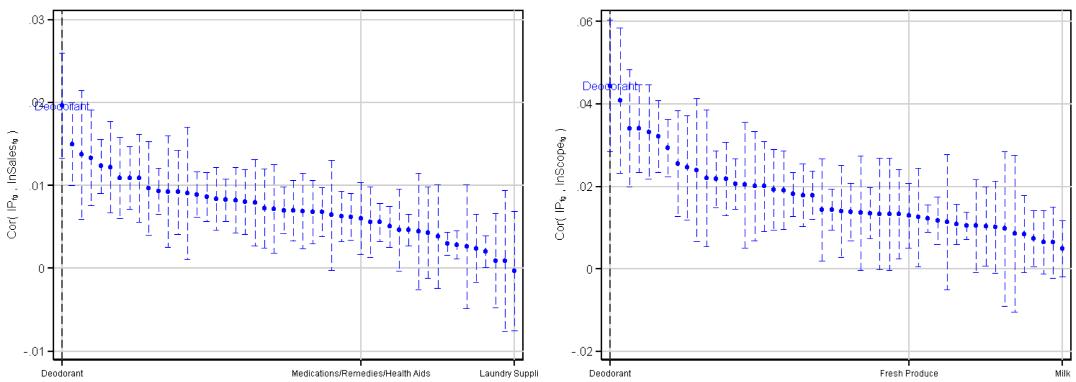


Figure A.10: Extensiveness of Personalization and Firm Characteristics

(a) Firm Sales (b) Firm Scope

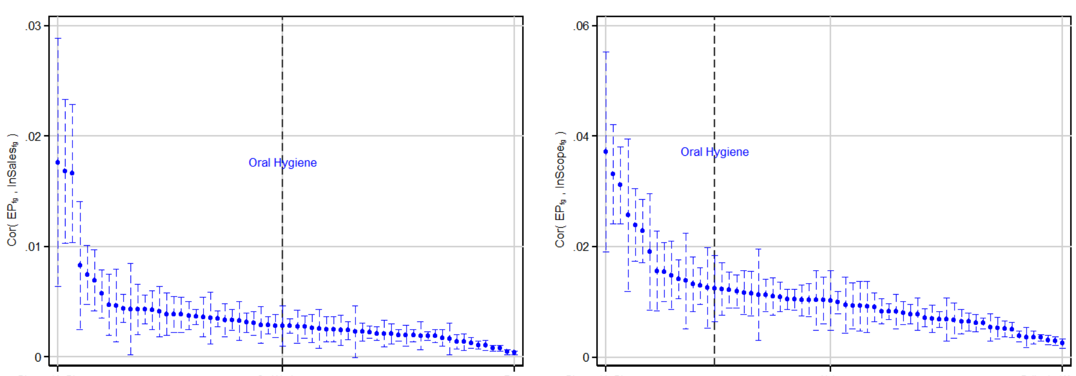
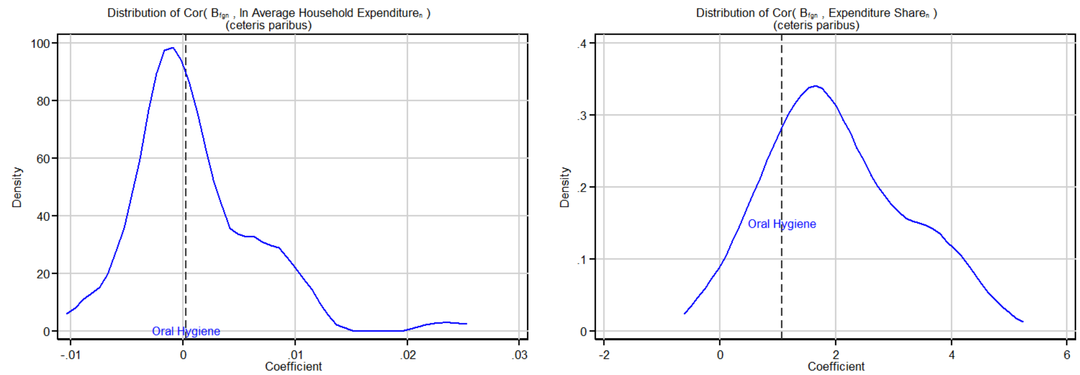


Figure A.11: Consumer Characteristics and Product Personalization

(a) Average Household Expenditure (b) Expenditure Share



A.6 The Oral Hygiene Product Case

A.6.1 The Oral Hygiene Product Case

The tables of this section show the differences in consumption patterns and demographic characteristics of the five segments described in Table 1 for the product group of oral hygiene products. Table A.3 shows the share of expenditures within segment on different brands. For instance, consumers of Segment 1 disproportionally purchase Colgate products, which account for 77% of their expenditure share on oral hygiene products. The second most purchased brand is Crest, which accounts for 10% of expenditures. Table A.4 shows the share of expenditures within segment on different on different product groupings. For instance, consumers of Segment 1, whose favorite brand is Colgate, predominantly purchase Colgate Total (13% of expenditures) and Colgate Optic White (9% of expenditures).

Tables A.5, A.6, and A.7 report the distribution across segments of different consumer demographic characteristics. Table A.5 shows the population shares of different race/ethnicity for each segment. The segmentation across segments does not follow these demographic variables, as the share of each race/ethnicity is very similar across segments. For example, the share of white households ranges between 0.89 and 0.91. In Table A.6, we show the share of the population by income, age, and education. We find that Segment 5 is overly represented by low-income households, since they account for 41% of the segment population, while in the other segments the share of low-income household is 24-27%. Segment 5 also has a higher average age as it has the lowest share of young households (22%) and the highest share of older households (62%). Finally, in Table A.7, we consider the population shares by head of household, marital status, and presence of children. For example, segments 1, 2, and 3 tend to have a larger share of the population with a child than the other segments.

Table A.3: Types characteristics - Top 5 Brands (Within-type share in parenthesis, %)

Rank	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
1	Colgate (77)	Crest (78)	Sensodyne (62)	Aquafresh (42)	Aim (44)
2	Crest (10)	Colgate (10)	Colgate (7)	Arm & Hammer (28)	Ultra Brite (17)
3	Aquafresh (2)	Sensodyne (2)	Crest (7)	Colgate (10)	Pepsodent (14)
4	Sensodyne (2)	Aquafresh (2)	Tom's Of Maine (6)	Crest (7)	Colgate (8)
5	Arm & Hammer (2)	Arm & Hammer (2)	Biotene (3)	Sensodyne (2)	Crest (3)

Table A.4: Segments characteristics - Top 5 Brand Description (Within-segment share in parenthesis, %)

Rank	Segment 1	Segment 2	Segment 3
1	Colgate Total (13)	Crest (30)	Sensodyne (30)
2	Colgate Optic White (9)	Crest Pro-Health (22)	Sensodyne Pro Namel (26)
3	Colgate Max Fresh (6)	Crest 3D White (10)	Tom's Of Maine (5)
4	Colgate Total Advanced Clean (6)	Crest Complete (6)	Sensodyne Full Protection (3)
5	Colgate (5)	Crest Whitening Expressions (3)	Mentadent (3)
Rank	Segment 4	Segment 5	
1	Aquafresh T-P Extreme Clean (16)	Aim (44)	
2	Arm & Hammer Advance White (11)	Pepsodent (14)	
3	Aquafresh T-P (8)	Ultra Brite All In One (13)	
4	Arm & Hammer Peroxi Care (7)	Ultra Brite (4)	
5	Arm & Hammer Complete Care (7)	Close-Up (3)	

Table A.5: Segments characteristics - Race and Ethnicity (Population Share)

Segment #	White	Black	Asian	Hispanic
1	0.89	0.08	0.02	0.06
2	0.91	0.06	0.02	0.05
3	0.91	0.06	0.02	0.05
4	0.90	0.07	0.02	0.06
5	0.90	0.08	0.01	0.05

Table A.6: Segments characteristics - Income, Age, and Education (Population Share)

Segment #	Low-Income	High-Income	Young	Old	High School (or below)	College	Graduate
1	0.26	0.37	0.36	0.48	0.17	0.65	0.18
2	0.24	0.38	0.34	0.49	0.16	0.66	0.18
3	0.26	0.38	0.27	0.57	0.16	0.65	0.19
4	0.27	0.36	0.37	0.47	0.17	0.66	0.17
5	0.41	0.21	0.22	0.62	0.22	0.64	0.14

Table A.7: Segments characteristics - Head of Household, Marital Status, and Children (Population Share)

Segment #	Female Head	Both Head	Married	Child under 12	Teenagers
1	0.25	0.67	0.68	0.18	0.07
2	0.24	0.68	0.68	0.16	0.08
3	0.27	0.65	0.66	0.12	0.05
4	0.24	0.66	0.66	0.18	0.07
5	0.27	0.58	0.59	0.10	0.06

A.6.2 Segment Ladder

Firms tend to mainly target the segments with the largest average expenditures. There is, in fact, a positive relationship between the average expenditure of a segment and the share of firm products targeting such segment. In addition, larger firms tend to target multiple segments, but they predominantly focus their products to cater to the taste of the same segments targeted by the smaller firms.

We investigate a possible segment ladder for consumer type segments. We divide firms into three bins indexed by b according to their product scope. In particular, the first bin contains firms with narrow scope (bottom 33 percentile with respect to number of UPC products), the second firms with middle scope (33 to 66 percentile), and the third firms with wide scope (top 33 percentile). In each bin there are F_b firms. For each bin and segment, we compute the average B_{fgn} previously defined in (28). Table A.8 summarizes the average B_{fn} for firms with narrow, middle, and wide scope.[28]

Table A.8: segment Ladder and Firm Scope

All UPCs			
segment #	Narrow-scope	Middle-scope	Wide-scope
1	0.133	0.117	0.124
2	0.278	0.076	0.066
3	0.500	0.304	0.281
4	0.056	0.070	0.113
5	0.000	0.069	0.038
Refined UPCs			
segment #	Narrow-scope	Middle-scope	Wide-scope
1	0.117	0.069	0.063
2	0.075	0.022	0.041
3	0.208	0.143	0.070
4	0.042	0.074	0.094
5	0.067	0.075	0.042

[28] In the sample consisting of all UPC products, the average numbers of UPC products per brand are 2.83, 5.67, and 87.75 for the narrow-, middle-, and wide-scope firms. In the sample that consists of UPC products sold in all segments, the average numbers of UPC products per brand are 5.33, 12.83, and 162.17 for the narrow-, middle-, and wide-scope firms.

Narrow scope firms tend to disproportionally offer large shares of their product varieties to segment 3 and, to a smaller extent, segment 1 and 2. Focusing on the sample with all UPC, on average, narrow scope firms offer 50% of their scope to segment 3, 27% to segment 2, and 13% to segment 1. Recall that segment 3 has the largest average expenditure on toothpaste. Segment 3 is also mainly targeted by middle- and wide- scope firms, as 30% and 28% of their product varieties rank first in segment 3. These firms second most targeted segment is segment 1 (12% of their scope ranks first in segment 1). However, these firms tend to also target other segments as well. The comparison of targeted segments across firms highlights a ladder in the scope expansion. Firms tend to first target segment 3 and segments 1-2. As firms add more products, these are targeted to segments 4 and 5.