The Informational Content of Consumer Choice in Differentiated Product Markets*

Johannes Kandelhardt† André Romahn ‡ Christine Zulehner §

Abstract

We study the impact of consumer inattention on market outcomes for the US ready-to-eat cereal market by estimating a discrete-type mixed logit model with heterogeneous consideration sets within and between consumer types. The full information benchmark model is statistically rejected against all limited consumer attention specifications. Under the full information assumption own-price elasticities are inflated and cross-price elasticities are an order of magnitude smaller than those of our most preferred limited consumer attention specification. Product-level markups are higher under limited attention and are estimated by all models to increase over the period from 2006 to 2020. The consideration proxy that best fits the observable data has on average six products, while there are on average 153 products in the market. While consumer behavior is best explained by limited attention, our model selection tests indicate that firms on average expect consumers to be fully informed when setting prices.

*Disclaimer: Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

We would like to thank Nicolas de Roos and participants at the Brown Bag Seminar and PhD Student Research Workshops at HHU Düsseldorf (DICE), the PARETO workshop in Paris, the Cresse conference 2025 and the 2025 DFG CPCMFR Workshop in Düsseldorf for helpful comments and suggestions. All remaining errors are our own.

We gratefully acknowledge financial support provided by the DFG under the umbrella of the research group FOR 5392 (project number 462020252) and the FWF (project number I6136-G). Computational infrastructure and support were provided by the Centre for Information and Media Technology (ZIM) at Heinrich Heine University Düsseldorf.

†kandelhardt@dice.hhu.de, Düsseldorf Institute for Competition Economics (DICE) at Heinrich-Heine-Universität Düsseldorf

‡romahn@hhu.de, Düsseldorf Institute for Competition Economics (DICE) at Heinrich-Heine-Universität Düsseldorf

§christine.zulehner@univie.ac.at, University of Vienna, Wifo, CEPR

1 Introduction

Understanding and quantifying the effects of mergers, regulation or taxation requires credible models of demand and supply, where the outcomes depend on the underlying firms' cost and consumers' demand. In this paper, we focus on modelling the demand side in markets where products are differentiated. Examples are beer, cars or coffee. The bulk of the existing literature in economics on differentiated product demand imposes a full information assumption. Consumers facing a potentially very large set of products to choose from, are assumed to be fully aware of all these options. There is, however, considerable empirical evidence that this assumption is not realistic.

Evidence on consumers' limited attention to product offerings goes back at least to Hauser and Wernerfelt (1990) who report average consideration set sizes for several differentiated product categories based on survey data. The categories covered include automobiles, analgesics, beer, coffee, shampoo and many others. These consideration set sizes range on average between at least two and at most eight options. Also, recent empirical evidence finds that consumers do not always make fully informed purchasing decisions and choose from a reduced set of alternatives. For example, Honka et al. (2017) provide evidence that the size of households' consideration sets is on average 2.2 banks for financial services. Draganska and Klapper (2011), who investigate the German coffee market, find that households consider 2-4 products on average. For a systematic survey of the literature, see, for example, Crawford et al. (2021).

If limited attention in differentiated product markets is pervasive, the standard models for demand estimation assuming full information (e.g., McFadden (1978), Berry et al. (1995)) yield biased preference estimates. The erroneous information assumption in combination with biased preferences may imply misleading estimates of substitution patterns between the available products in a given market. These demand diversion patterns feed into the model's implied product markups and thereby also give biased estimates of firms' profitability and pricing decisions. It is therefore highly desirable in applied empirical work to account for both preference heterogeneity and unobserved heterogeneous consideration sets among consumers.

Estimating a structural demand model that ticks off both these requirements still faces substantial challenges. Sovinsky (2008) incorporates consideration set heterogeneity into the mixed logit demand model. To distinguish between preference and consideration (or attention) heterogeneity, an exclusion restriction is imposed. Advertising expenditure shifts consumer attention but does not enter consumer utility. Price is thereby not allowed to affect consumer awareness. Imposing such an exclusion restriction can therefore be restrictive in how consumer attention is modelled. Abaluck and Adams-Prassl (2021) identify both preference and consideration set heterogeneity based on observable deviations from Slutsky symmetry.[1] Such deviations can only be rationalized with limited consumer attention. Estimating deviations from Slutsky symmetry suffers from the curse of dimensionality. Making the approach highly flexible, but impractical for grocery categories, which typically have large choice menus. Dardanoni et al. (2020) identifies

[1]Random utility models, such as the logit, nested logit and random coefficient logit impose Slutsky symmetry under full information.

attention heterogeneity from aggregate market share data imposing uniform consumer tastes.[2] The approach requires enumerating all possible consideration sets and is thereby also impractical for markets with large choice menus.

We adopt the approach of Crawford et al. (2021) to estimate demand for ready-to-eat (RTE) cereals. The approach rests on the result of McFadden (1978) that consumer preferences can be consistently estimated based on subsets of the available choice menu, as long as the subsets belong to consumers' true consideration sets. According to this work, preference estimates are biased when we include too many products and not too few. Based on household-level data, we proxy for unobserved consideration sets using the observed purchase histories. The assumption is that if households have purchased products in the past they are highly likely still being considered for purchase in the present. We condition on these consideration set proxies and thereby avoid having to formulate an explicit model of consideration set formation. This allows us to sidestep the identification challenges of allowing for heterogeneity in both preferences and consideration sets and does not re-introduce the curse of dimensionality into demand estimation. The drawback is that we have no microfoundation for how consumers form their consideration sets.

Estimating the model at the individual household-level as in Crawford et al. (2021) is a highly flexible approach, but leads to few observations for each household in our data and thereby relatively imprecise parameter estimates. By using k-means clustering, we discretize consumers into types. This allows us to pool many households within each type and obtain substantially narrower confidence bands. Estimating demand for each type separately, satisfies the requirements of the approach, while obtaining flexible product-level demand functions that break independence of irrelevant alternatives.[3] We use maximum likelihood to estimate our demand model. To address price endogeneity, we exploit the high frequency nature of our data on consumer purchases.[4] Price spells in the US last for four months (Bils and Klenow, 2004, Nakamura and Steinsson, 2008, Karadi et al., 2023). Observing purchases at a daily frequency thereby reduces the correlation of demand shocks, the model's residual, and prices. Any remaining correlation is tackled using product, time controls and observable coupon use.

We focus on the US ready-to-eat (RTE) cereal industry for the years $2006-2020$ and estimate our model using NielsenIQ household level purchase data combined with NielsenIQ supermarket scanner data. Our estimates show substantial differences in substitution patterns between the demand model assuming full information on the consumer side and the comparison model that proxies for heterogeneous consideration sets. In the latter, substitution effects quantitatively

[2]The authors derive a mapping between the consideration probability model, where a given consumer considers any available product with the same probability, and the consideration capacity model, where each consumer considers at most $k = 0, 1, \ldots$ products. Again, each product has the same probability of entering the consideration set.

[3]We estimate a discrete-type mixed logit model with preference heterogeneity as in Berry and Jia (2010) and Conlon et al. (2024) and extend it by allowing for consideration set heterogeneity within and between consumer types.

[4]As we perform model selection tests on the supply side of the market, we do not use a control function approach to address price endogeneity. As discussed in Berry and Haile (2021), control function methods are valid only under strong functional form restrictions (Blundell and Matzkin (2014)), which are violated even in standard parametric specifications of demand for differentiated products.

play a much larger role and are concentrated on relatively small clusters of products. With the full information assumption, substitution between products is much more uniform and on average substantially lower than with the limited information assumption. Moreover, with the full information assumption, consumers are estimated to be relatively more price sensitive, which deflates markups relative to the incomplete consideration specifications. For both informational assumptions, consumers become less price sensitive over time. For the limited attention model, markups increase by roughly 10 percentage points over the sample. Our estimated own-price elasticities are quantitatively close to the existing literature on RTE cereals demand that uses instrumental variable GMM estimation (Backus et al., 2021, Atalay et al., 2023, Nevo, 2001) or covariance restrictions (Döpper et al., 2025). In the full information model, product dummies tend to be large and negative, while these product tastes turn positive under limited attention. If consumers actually exhibit limited attention, these negative utility shifters under full information are needed to rationalize the relatively small observed market shares. This finding is in line with Draganska and Klapper (2011).

Following Vuong (1989), we run demand model selection tests and can reject the full information model against all limited attention specifications (complete purchase history, twelve-month purchase history and eleven-month purchase history).[5] According to all pairwise tests, the consideration set proxies based on consumers' twelve-month purchase history fit the data best. We further contribute to the literature by empirically testing models of consumer demand and firms' expectations about which products consumers consider. Assuming that firms face constant marginal costs and compete Nash-in-prices, we can model firms' price setting using the profit maximization conditions under the different informational assumptions regarding consumer purchases. In total there are four combinations with full information and limited attention being taken into account on both the demand and supply sides. While consumer behavior is best approximated using limited attention, our tests indicate that firms expect consumers to be fully informed.

Our combination of observable consideration set proxies, discretizing consumers into types and not having to rely on instrumental variables, would scale well to multi-category estimation as in Döpper et al. (2025) and Atalay et al. (2023). In future research, we plan to apply our model to many categories to examine broader trends in consumer attention in grocery markets.

The empirical work closest in spirit to our application is Draganska and Klapper (2011). The authors use auxiliary data sources, such as survey data on brand recall and advertising data, to estimate consumer demand and compare models of consumer information. DellaVigna and Gentzkow (2019) analyze firm price setting using the same data and provide estimated bounds for managerial adjustment costs that rationalize uniform price setting in large US grocery retail chains.

The remainder of the paper is organized as follows: Section 2 provides detail on our data, Sections 3 and 4 present our demand and supply models. Identification and our demand estimation and model selection results are discussed in Section 5. We conclude in Section 6.

[5] With maximum likelihood estimation, the testing procedure simplifies to pairwise likelihood ratio tests.

2 Household Data

Our data covers the period from 2006 to 2020 and is provided by the Kilts Center for Marketing at Chicago Booth. The main data set is the NielsenIQ consumer panel, which is complemented by the NielsenIQ scanner data. The panel includes more than 170,000 unique households and more than 1 billion items bought. Detailed consumer demographics along with statistical weights for obtaining a representative sample of the entire population, are also provided.[6] From the panel, we observe consumers' past purchase histories, which we use to obtain three proxies for consumers' consideration sets.[7] We construct a complete purchase history which includes all products bought, and twelve- and eleven-month purchase histories, which include the products bought in the last twelve or eleven months on a rolling basis. The store-level assortments are constructed from the scanner data and we use these to closely track the choice setting under the full information assumption.[8]

2.1 Ready-to-Eat Cereal Purchases

The NielsenIQ consumer panel provides granular information on consumer purchases in the RTE cereals category. A purchase occasion is defined as a panel member household-store-date combination. Each shopping occasion is recorded (date, location, retailer) and includes all items bought (product identifier code, quantity, total price paid per item, coupon usage, coupon value). The product characteristics package size, amount per item, and brand affiliation are matched on using the unique product identifiers.[9] To account for quantity discounts and differences in packaging between products, we measure prices relative to package content in ounces (OZ). All prices are deflated and measured in 2017 US dollars. The data also allows us to directly observe coupon usage at each purchase occasion, so that we obtain transaction-level prices.[10] We match on store assortments at the purchase occasion's date to reconstruct the choice situation under full information. We also observe how many items of each available product a household buys. Multi-unit purchases are not surprising in our context, because many households include more than one member. More than one unit bought of a given product is reflected in a modified weight for the household, which is based on the nationally representative sampling weights provided by NielsenIQ.

Market Definition We adopt NielsenIQ's definition of the RTE cereals category as the relevant set of products. A market is defined as the smallest geographic delineation for which representative sampling weights are provided by NielsenIQ. These are the so-called Scantrack markets, which are large urban areas, such as Chicago, Denver, or Boston.[11]

[6] We use the NielsenIQ household weights throughout. The smallest geographical area for which such representative sampling weights are available is the so-called Scantrack Market. We obtain nationally representative figures by aggregating over all Scantrack Markets.

[7] Crawford, Griffith and Iara (2021) label these as sufficient sets.

[8] This holds with the assumption of non-zero unit sales. We do not keep track of products that have zero sales.

[9] A sixpack of beer has six bottles and each bottle has a given pack size.

[10] NielsenIQ scanner data prices are on the weekly frequency, whereas the prices of household purchases are entered by households directly or drawn from the scanner data. If a household does not enter a price, it is matched on by NielsenIQ from the scanner data.

[11] From 2021, which lies outside of our sample, NielsenIQ has replaced Scantrack Markets by Syndicated Major Markets.

Table 1: Suggested Optimal Number of Clusters by Score

Statistic	2018	2019	2020	2021	2022	Pooled
Calinski-Harabasz Score	10	9	10	7	9	9
Silhouette Score	16	18	15	17	16	18
Davies-Bouldin Score	14	10	10	18	11	11
Distortion Score	9	8	8	9	8	8

We determine the optimal number of household clusters based on household characteristics using different metrics. We let each method decide on the optimal number of clusters out of a possible range of 2 to 20 possible clusters. Increasing the upper bound yields similar results. We scale the data up to the population level using NielsenIQ sampling weights. Variables used: household size, income, household composition (married, living alone, living with relatives/non-family), dummy variables for single, widowed, separated, living with underage child in household and ethnicity, as well as age, education and a dummy for elderly.

At the local market level, we modestly trim the long tail of product availability. Products with less than one percent of sales volume are dropped (see Figure B.1). Post-trimming, markets offer on average a choice menu of 153 products.[12]

Outside Option The outside option is part of every consumer's consideration set and is simply not purchasing any RTE cereals. To tie the definition of the outside good to the observable data, we treat consumer shopping trips without any of the cereals options bought as an occasion where the household chooses the outside option.

2.2 Discretizing Households

For our demand estimation, we treat households as consumers and use the two terms interchangeably. In an ideal world, we would have a long time series of purchase occasions for each of the households in the NielsenIQ consumer panel.[13] The actual data, however, contains short time series for most households. Estimating the model at the individual household level is therefore infeasible. We tackle this obstacle by discretizing the sample of consumers into different types as in Conlon et al. (2024). Thereby, each type contains many households with a large number of purchase occasions during the sample period.

To discretize households in the panel data, we use k-means clustering based on the observable household demographics as inputs. In total, there are 19 observable demographics covering household size, age, income, education, ethnicity and additional information on household type and composition. We perform the clustering excercise with the four most commonly used criterion functions to determine the optimal number of consumer types. Table 1 presents the optimal number of consumer types by score and year along with an optimal number for the pooled years. We plot each of the score functions in the Appendix (see Figure B.2). This yields two results for the optimal number of consumer types: eight and eighteen. Our most preferred specification uses the discretization based on eight types, but outcomes are qualitatively identical and quantitatively close under the alternative discretization with eighteen different types. For each

[12]As a comparison, in their multi-category estimation, Döpper et al. (2025) retain the top 20 products by sales in each category.

[13]This would make estimation of the demand models feasible at the individual household level.

Table 2: Summary Statistics - Consumer Panelists

	1	2	3	4	5	6	7	8
size	1.01	1.01	1.63	4.30	2.37	2.94	2.99	2.15
married	0	0	0	1	1	0	0.62	0
single	0.59	0.48	0	0	0	0.34	0.20	0.49
widowed	0.01	0	1	0	0	0.02	0.04	0.08
separated	0.36	0.49	0	0	0	0.48	0.09	0.24
child	0	0	0.09	0.99	0	0.47	0.46	0.02
age	52	51	62	40	54	48	44	48
edu	14.99	14.59	13.18	14.04	13.38	14.45	15.76	14.40
inc	48	39	31	31	38	30	43	36
eld	0.22	0.24	0.71	0.01	0.22	0.10	0.08	0.12
white	0.82	0.76	0.87	0.77	0.84	0.71	0	0.79
black	0.12	0.18	0.10	0.10	0.08	0.20	0	0.10
asian	0	0	0	0	0	0	1	0.04
hisp	0.06	0.07	0.04	0.21	0.13	0.11	0.10	0.11
f_hh	0	1	0.46	0	0	0	0.08	0
m_hh	1	0	0.17	0	0	0	0.10	0
rel	0	0	0.37	0	0	1	0.20	0
n_rel	0	0	0	0	0	0	0	1
fraction	0.12	0.10	0.09	0.21	0.22	0.19	0.04	0.03

Clustered household summary statistics for 8 consumer types. The variables are: household size (size, 1 indicating a single household, 2 indicating two members, and so forth), age by household head, income in thousands of equivalized 2017 US dollars (inc), education (edu, in years of education), indicator variables for married, single, widowed or separated living households, the presence of underage children (child), elderly households (eld), ethnicity, households comprising only a female (f_hh) or male (m_hh) or households living with other related (rel) or other unrelated persons (n_rel). Individual households are weighted by NielsenIQ household weights. Aggregating over households gives the corresponding representative population weights (last row). Data used spans the period from 2006 to 2020. All values are rounded to two decimal places or to the nearest integer, as appropriate.

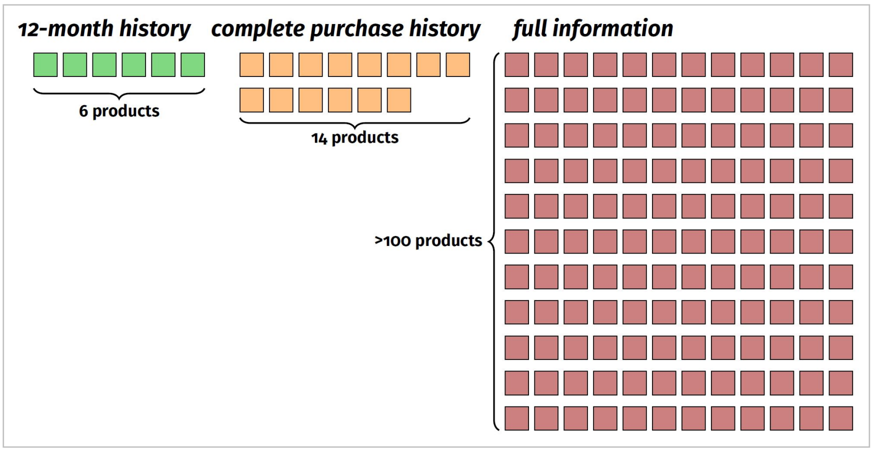
year and type of consumer, we observe $16,541$ purchases on average. Table 2 provides summary statistics for the demographic variables for each of the eight types. A corresponding graphical visualization is Figure B.3 in the Appendix.

Some broad patterns emerge: types 1 and 2 are respectively male and female single households with similar levels of education. Type 4 is a household with children that also has the largest size with on average 4.3 members. This household is also relatively young and highly educated. Type 7 is a highly educated, high income, Asian household. Thus, overall the clustering based on demographics gives intuitive types that make up the relevant consumer population. Each of these types has its own preference parameter vector to be estimated and each member of a given type has an idiosyncratic consideration set proxy.

2.3 Proxy Consideration Sets

The RTE cereal market has on average 153 differentiated products on offer. In some markets, we observe more than 200 products. We obtain proxy consideration sets as subsets from all the products in each consumer's purchase history. Each proxy set contains the outside option and,

Figure 1: Full Information and Past Purchase History Number of Products



Comparison of the average number of unique products by specification of the consideration set for the ready-to-eat cereal market. Specification "12-month history" refers to using households last twelve months unique products bought as consideration set. Specification "complete purchase history" refers to using households unique products ever bought as consideration set. The full information specification uses all products available in the market as consideration set.

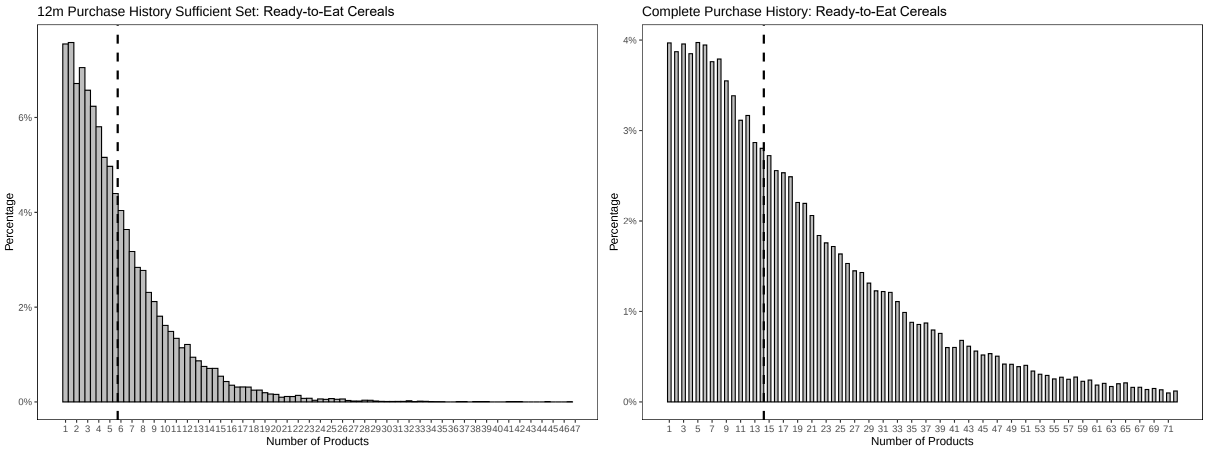
if not yet the case, the current purchase of the household.[14] Apart from using the complete purchase history as a proxy, we also define rolling purchase history windows covering the last eleven and twelve months of purchases to define the proxies. For the rolling window specifications, these product sets change over time.[15] It is important to note that within consumer types, each househould that belongs to a consumer type has an individual-specific purchase history. Each consumer type therefore has heterogeneous consideration over all its members.

Figure 1 shows the average number of products that enter consumers' consideration sets for three cases. Each square represents a unique product the consumer is assumed to evaluate on a given purchase occasion. For the twelve-month purchase history, the average consideration set size is 6. Taking the complete purchase history as a proxy instead yields on average 14 products. Under full information, consumers' consideration sets include on average 153 products, instead. Quantitatively, there is thus a large difference between our consideration set proxies under limited attention and the corresponding choice menu under full information.

Figure 2 shows the size distribution of our consideration set proxies. The left graph shows the number of unique products consumers bought on average for the twelve-month purchase history. Here, consumers bought 6 products on average. In the right graph, we show the corresponding number of unique products if we use the complete purchasing histories, instead. On average, consumers consider 14 products for this proxy specification and their are many cases for which consideration set sizes exceed 30 products. In the Appendix, we show more detailed sufficient set statistics broken down by consumer-type in Figure B.4 and sufficient sets for the eleven months window in Figure B.5.

[14]In our proxies, first-time purchased products therefore enter on the date of the first purchase and not before.
[15]Implicitly, we are assuming that households either forget products they have purchased outside the window or that these products yielded relatively low utility and therefore are no longer part of the considered set of products.

Figure 2: Household Sufficient Set Statistics



Sufficient set statistics using the purchase history to calculate sufficient sets. Each panel shows the percentage of consumers buying 1, 2, 3, etc. unique product of the specified category, using the complete purchase history, the twelve-and eleven-month purchase histories. Dashed vertical lines indicate the weighted average number of product bought conditional on purchasing a category product. NielsenIQ Panel data (years 2006 up to and including 2020). All graphs weighted by NielsenIQ household weights.

3 Demand

The approach of McFadden (1978) has the advantages of breaking demand estimation's curse of dimensionality by using the functional form of the multinomial logit model and not requiring an explicit model of consideration set formation as in for example Sovinsky (2008). We use past purchases to obtain a proxy for consumers' true consideration sets and condition on these directly observable proxy sets. In this way, consumer-specific consideration sets are equivalent to a fixed effect, which can be differenced out. This allows for the consistent estimation of consumer preference parameters even in the presence of unobserved consideration sets. A key point to recall is that including too many products in the consideration set proxy yields biased preference estimates, while including too few products does not break consistency as long as the products included are part of consumers' true consideration sets.[16]

These advantages, however, come at a price. First, consistency of the preference parameter estimates holds for the simple multinomial logit model, which imposes independence of irrelevant alternatives (IIA) on the observed consumer choices. It is well known that IIA leads to unnecessarily restrictive substitution patterns (see e.g. Nevo (2001)). Second, we need to make the assumption that our specific proxy for consumers' unoberserved consideration sets is appropriate.

We overcome the first restriction by modelling the population of consumers as several discrete types. Preferences within types are uniform and therefore satisfy IIA, while (proxied) consideration sets are specific to each household of the same type. Thus, each type is characterized by uniform preferences for product attributes and heterogeneous consideration sets. Preferences

[16] A potential disadvantage of specifying very small proxy sets in finite samples is the reduction in the observable data used.

between consumer types are estimated independently of each other and vary over types. Each of the type-specific preference parameter vectors is estimated consistently. For details on our construction of consumer types see Section 2.2. Product-level demand functions are obtained by aggregating choice probabilities over all consumer types and all individual households within each type. This aggregation at the product level breaks IIA and allows for flexible substitution patterns between products. We obtain a discrete type mixed logit model as in (Conlon et al., 2024) and Berry and Jia (2010) and extend it by allowing for heterogeneous consideration sets.

Regarding the appropriateness or usefulness of our consideration set proxies, we consider several variants of the mapping between unobserved consideration sets and the observed past purchase history. Following Crawford et al. (2021), our limited attention demand model uses the following proxies for consumers' unobserved consideration sets: (1) the complete past purchase history of the household, (2) the purchase history over the past twelve months, and (3) the purchase history over the past eleven months. We always include the current month in each definition of the consideration proxies. If a product is purchased for the first time, it is therefore included at the time of purchase.[17]

3.1 Utility Specification

Each market offers $j = 0, 1, \ldots, J$ products to choose from with $j = 0$ denoting the outside option of not purchasing one of the products. As is standard in the literature, we assume that the outside option is a member of every consideration set. The set of all available products is denoted by $\mathcal{J}$. A consumer type τ is composed of individual households indexed by i with household-specific consideration sets $\mathcal{C}_i \subseteq \mathcal{J}$. The consideration sets vary with the informational assumption that we place on households. With the full information assumption that is common in much empirical work, all $\mathcal{C}_i$ are identical to $\mathcal{J}$. With limited attention, however, the $\mathcal{C}_i$ will be subsets of the directly observed complete purchase histories of the households. To satisfy the requirements of McFadden (1978), we impose uniform preference parameters θ_τ for each consumer type. Thus, a household member of type τ derives utility $u^\tau_{i,j}|\mathcal{C}_i$ from choosing option j.

$$\begin{aligned} u^\tau_{i,j}|\mathcal{C}_i &= x_j\beta_\tau - \alpha_\tau p_j + \xi_{\tau,j} + \epsilon^\tau_{i,j} \\ &= \delta_j + \epsilon^\tau_{i,j} \end{aligned} \tag{1}$$

Consumer types have preferences over the directly observable product attributes gathered in x_j, denoted by β_τ and α_τ denotes the consumer type's price sensitivity. Thus, type τ's preference parameters are $\theta_\tau = (\beta_\tau, \alpha_\tau)'$. Assuming that the $\epsilon^\tau_{i,j}$ follow a Gumbel distribution, yields the familiar multinomial logit choice probabilities.

$$Prob^\tau_{i,j}(\theta_\tau; \mathbf{x}, \mathbf{p})|\mathcal{C}_i = \begin{cases} \dfrac{e^{\delta_j(\theta_\tau)}}{1 + \sum_{k \in \mathcal{C}_i} e^{\delta_k(\theta_\tau)}}, & j \in \mathcal{C}_i \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

[17] Getting at products, which have previously not been bought, but are considered, would require us to specify a model of consideration set formation. While this is a very interesting question, it exceeds the scope of this paper.

Here, $\mathbf{x}$ and $\mathbf{p}$ are a matrix and vector, respectively, that gather all product attributes and prices in the market. Compared with the full information assumption, incomplete consideration among consumers introduces sparsity into individual consumer's choice probabilities.

3.2 Product-level demand

To arrive at product-level demand functions, we aggregate over all consumer types and their household members. Each consumer type is associated with a population weight w_τ with all type weights summing to one.[18] We accommodate the fact that some households of a given type buy multiple units of a given product j by calculating household i's units bought share for option j in the market, $w_i^\tau = n_{i,j}/\sum_{i'} n_{i',j}^\tau$. Over all members of a consumer type, these weights sum to one. Market-level demand for product j is then obtained by aggregating over all consumer types.

$$\begin{aligned} s_j(\theta; \mathbf{x}, \mathbf{p}, \mathbf{w}, \mathcal{C}) &= \sum_\tau w_\tau \sum_{i\in\tau} w_i^\tau Prob_{i,j}^\tau(\theta_\tau; \mathbf{x}, \mathbf{p})|\mathcal{C}_i \\ &= \sum_\tau w_\tau \sum_{i\in\tau} w_i^\tau \frac{\mathbf{1}(j \in \mathcal{C}_i) e^{\delta_j(\theta_\tau)}}{1 + \sum_{k\in\mathcal{C}_i} e^{\delta_k(\theta_\tau)}} \end{aligned} \tag{3}$$

All relevant consumer weights are collected in the market-level vector $\mathbf{w}$. In contrast to consumer-type-level choice probabilities, this product-level demand function does not satisfy the IIA property, but instead yields a flexible functional form to obtain reasonable substitution patterns between the available products. The product-level demand function is characterized by both preference and consideration heterogeneity.

4 Supply

The set of available products in a given market, $\mathcal{J}$, is produced and sold by $f = 1, \ldots, F$ firms. For simplicity, we assume that either retailers offer products competitively or that manufacturers and retailers jointly maximize profits. We therefore abstract from the vertical relationship between manufacturers and retailers. Our model-implied markups should therefore be interpreted as the total markup over the supply chain and not only as manufacturer markups. We denote firm f's set of products as $\mathcal{F}_f$.

4.1 Profit Maximization

To characterize firm behavior, we assume that firms compete Nash-in-prices and face constant marginal costs, c: each firm f maximizes its profit by optimally setting prices for all products in $\mathcal{F}_f$. Let the market's size be measured by Q. The total operating profit of f follows.

$$\pi_f = Q \sum_{j\in\mathcal{F}_f} s_j(\theta; \mathbf{x}, \mathbf{p}, \mathbf{w}, \mathcal{C})(p_j - c_j) \tag{4}$$

[18] This is labelled "fraction" in the bottom row of Table 4.

At a profit maximum, the derivatives of the product-level demand functions that firm f faces are zero with respect to the prices set by the firm.

$$\begin{aligned}\frac{\partial \pi_f}{\partial p_j} =& \frac{\partial s_j(\theta;\mathbf{x},\mathbf{p},\mathbf{w},\mathcal{C})}{\partial p_j}(p_j - c_j) + \\ & \sum_{k\in\mathcal{F}_f, k\neq j} \frac{\partial s_k(\mathbf{x},\mathbf{p},\mathbf{w},\mathcal{C})}{\partial p_j}(p_k - c_k) = -s_j(\theta;\mathbf{x},\mathbf{p},\mathbf{w},\mathcal{C}), \text{ for all } j \in \mathcal{F}_f\end{aligned} \tag{5}$$

These product-level optimality conditions can be collected for all firms in the market and concisely expressed in vector-matrix notation.

$$\Omega(\theta;\mathbf{x},\mathbf{p},\mathbf{w},\mathcal{C})(\boldsymbol{p} - \boldsymbol{c}) = -\boldsymbol{s}(\theta;\mathbf{x},\mathbf{p},\mathbf{w},\mathcal{C}) \tag{6}$$

Here, $\boldsymbol{p}$, $\boldsymbol{c}$ and $\boldsymbol{s}$ are vectors collecting all prices, marginal costs and product-level shares in the market. Ω is a matrix that contains all the relevant own- and cross-price effects and its entry in the j-th row and k-th column is defined as follows.

$$\Omega_{j,k} = \begin{cases} \dfrac{\partial s_j(\theta;\mathbf{x},\mathbf{p},\mathbf{w},\mathcal{C})}{\partial p_k}, & \exists j,k : j,k \in \mathcal{F}_f \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

4.2 Implied Marginal Costs

With (6) in hand, we can determine each product's model-implied marginal cost. Once we have obtained the estimate $\widehat{\theta}$, note that the consumer type and household weights w, the product-level market shares $\boldsymbol{s}$ and prices $\boldsymbol{p}$ are directly observed in the data. All entries in $\Omega(\widehat{\theta}; w)$ are determined by the estimated parameter vector and the model-implied market share functions with the proxies for consideration sets also being directly observed. The only unobservable remaining is the vector of product-level marginal costs, which can therefore be backed out. The specific informational assumption regarding consumers' consideration sets affects the own- and cross-price effects and thereby also shifts the backed out marginal costs. We express this fact by treating the set of all consideration sets $\mathcal{C}$ as a parameter.

$$\boldsymbol{c}(\widehat{\theta};\mathbf{x},\mathbf{p},\mathbf{w},\mathcal{C}) = \boldsymbol{p} + \Omega(\widehat{\theta};\mathbf{x},\mathbf{p},\mathbf{w},\mathcal{C}) \cdot \mathcal{S} \tag{8}$$

We therefore obtain different backed out marginal costs for the full information and limited attention assumptions. The observed market-level shares for all products are gathered in the vector $\mathcal{S}$. For backing out the model-implied costs, prices are also directly observed.

In our model, there are two drivers for the sparsity in Ω, which determines how demand is diverted from one product to its rivals in response to an adverse change in the product's utility, such as a price increase or a drop in consumers' willingness to pay for its non-price attributes. First, firms' ownership pattern of the available products pins down, which cross-price derivatives have a direct effect on firm price setting and which do not. If two products are owned by the same firm, the corresponding entries in Ω are non-zero. If on the other hand these products are owned by rival firms, these entries are exactly zero.[19] Changes in prices or non-price attributes for such product pairs only work through indirect equilibrium effects.

The second determinant is the informational assumption we place on consumers. Imposing full information means all consumers consider all available products. Consideration sets are identical to $\mathcal{J}$ and thereby cannot contribute to the sparsity in demand diversion patterns. With limited attention, however, each consumer is only considering strict subsets of $\mathcal{J}$. Even if products j and k are owned by the same firm, it is possible that for a given consumer i only one (or none) of the products is considered. A change in the price of a product that is not considered has no effect on consumer i's demand. If this scenario applies to all consumers in the market, then the corresponding entry in Ω is also exactly zero. Thus, incomplete consideration can introduce additional zeros into the demand diversion patterns. It can also strengthen the non-zeros in demand diversion, because each consumer with limited attention considers fewer alternatives than the fully informed counterpart. Diverted demand is thereby not stretched over as many rival products. Limited attention's total effect on demand diversion patterns depends on the directly observed consideration set proxies, ownership patterns and the estimated preference parameters. How limited consumer attention affects market outcomes is therefore an empirical question.

5 Empirical Analysis

Estimation of the demand models is implemented using standard maximum likelihood. To estimate the effect of preference and attention changes over time, we split the sample into rolling three-year windows with one year of overlap between the individual windows. The estimation is run for each window separately.

5.1 Identification and Estimation

For the limited attention models, let $\mathcal{C} = \{\mathcal{C}_i\}_{i=1,\dots,I}$ denote the collection of all consideration set proxies and let $m = 1, \dots, M$ index individual markets in our sample of data. As we directly observe the proxies, $\mathcal{C}$ is treated as a parameter and choice probabilities are conditional on the realizations of the proxies. All prices and non-price characteristics are collected in $\mathbf{p}$ and $\mathbf{x}$. With the market-level shares defined by (3), the log-likelihood function follows as:

$$\mathcal{L}(\theta; \mathbf{x}, \mathbf{p}, \mathbf{w}, \mathcal{C}) = \sum_{m} \sum_{j \in m} \log \left(s_j(\theta; \mathbf{x}, \mathbf{p}, \mathbf{w}, \mathcal{C}) \right). \tag{9}$$

Our estimation approach assumes that all quantitatively relevant unobservables that cause a systematic correlation between the models' residuals and observed prices are pulled out of the error terms. Achieving this for our setting requires i) avoiding the use of sales-weighted product prices and ii) exploiting the fact that prices are rigid. First, NielsenIQ defines products narrowly, but different packages for one and the same product can be pooled. If we do not account for this package pooling, product-package-level demand shocks can cause systematic correlation between

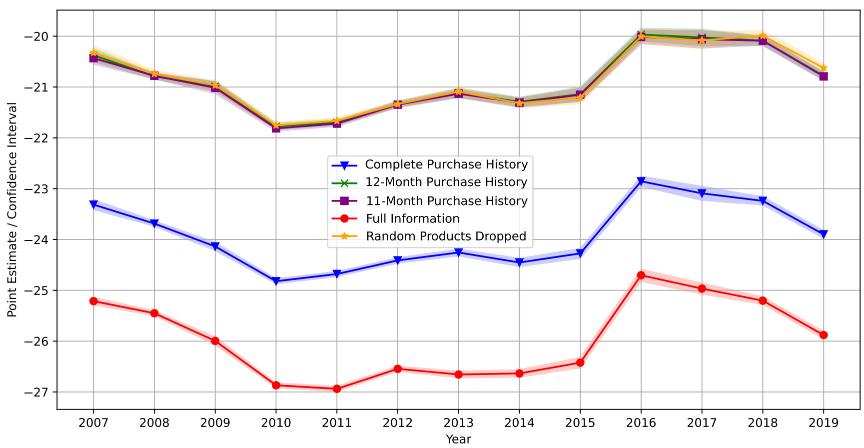
[19] We assume Nash-Bertrand competition and do not allow for deviations in conduct from this benchmark.

sales-weighted prices and the models' residuals. To tackle this issue, we use the average price over all package variants within products.[20] Second, in the presence of rigid prices, the high frequency nature of the purchase data, the direct observability of coupon use and household weights leads to a low correlation between the models' error terms and prices by construction. Additional dummies for products and time, specifically retailer dummies, year-month-week and day-of-week dummies, as well as dummies for major holidays and events, contribute further to absorbing any remaining unobservables.[21]

5.2 Estimation Results

Preference parameter estimates Figure 3 shows the time series of the estimated price coefficients averaged over consumer types. We use colors and markers to distinguish between specifications and show 95% confidence intervals as shaded areas around the point estimates. All point estimates are tightly estimated.

Figure 3: Estimated Price Coefficients for Ready-to-Eat Cereals



Estimated price coefficients for RTE Cereals over time. Estimates are based on three-year rolling averages calculated for each cluster, spanning from 2006 through 2020. We report the mid point estimates.

The red line shows the estimated price coefficients for the full information benchmark model. Consumers are estimated to have become increasingly price sensitive from 2007 to 2010. This holds for all specifications and coincides with the US economic downturn. During that period consumers may track prices closely, seek bargains and choose relatively more affordable products. During the remainder of the sample period, which stretches from 2011 to 2020, this pattern reverses. Consumers thereby become less price sensitive.

The purple, yellow and green lines are visually indistinguishable and show the estimated price coefficients for the eleven-month purchase histories with and without random product drops and the coefficients based on the specification using twelve-month purchase histories. The estimates are statistically indistinguishable. We randomly drop products from the specification

[20]The data on product characteristics allow us to calculate prices in units of measurement. All prices are expressed in ounces.

[21]We include holiday dummies for the following periods: Easter, Summer Holidays, Super Bowl, Thanksgiving/Black Friday, and Christmas/New Year's Eve.

Table 3: Limited Attention Consumer Type Price Coefficient Estimates

Year	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7	Type 8	Average
2007	−21.92	−18.05	−19.27	−20.79	−20.82	−20.15	−18.91	−21.25	−20.38
	(0.10)	(0.10)	(0.22)	(0.12)	(0.20)	(0.06)	(0.35)	(0.12)	
2008	−22.96	−18.63	−20.10	−20.82	−21.63	−20.20	−19.16	−21.13	−20.83
	(0.09)	(0.10)	(0.07)	(0.05)	(0.07)	(0.06)	(0.27)	(0.59)	
2009	−22.35	−18.71	−20.30	−21.65	−21.27	−21.02	−19.81	−21.02	−21.06
	(0.09)	(0.11)	(0.15)	(0.05)	(0.22)	(0.06)	(0.32)	(0.09)	
2010	−23.27	−19.62	−21.06	−22.56	−21.46	−21.18	−22.84	−21.42	−21.69
	(0.10)	(0.10)	(0.08)	(0.05)	(0.08)	(0.05)	(0.09)	(0.09)	
2011	−22.88	−20.11	−22.24	−22.53	−21.78	−20.66	−22.37	−21.21	−21.76
	(0.09)	(0.11)	(0.08)	(0.04)	(0.08)	(0.05)	(0.08)	(0.09)	
2012	−22.87	−19.93	−21.88	−21.99	−21.80	−19.80	−22.20	−20.97	−21.45
	(0.10)	(0.11)	(0.08)	(0.04)	(0.08)	(0.06)	(0.08)	(0.10)	
2013	−21.71	−19.33	−21.55	−21.54	−22.49	−20.41	−22.12	−19.65	−21.32
	(0.09)	(0.11)	(0.14)	(0.06)	(0.08)	(0.11)	(0.11)	(0.12)	
2014	−20.26	−19.76	−21.15	−22.36	−22.14	−20.86	−22.06	−19.98	−21.34
	(0.09)	(0.12)	(0.09)	(0.15)	(0.09)	(0.16)	(0.19)	(0.11)	
2015	−20.46	−18.59	−19.23	−21.99	−21.92	−21.88	−21.10	−19.82	−21.12
	(0.27)	(0.26)	(0.27)	(0.07)	(0.14)	(0.10)	(0.13)	(0.13)	
2016	−21.02	−17.59	−18.22	−20.55	−20.34	−20.26	−20.08	−19.01	−19.96
	(0.11)	(0.13)	(0.30)	(0.12)	(0.14)	(0.18)	(0.10)	(0.11)	
2017	−19.58	−17.06	−17.56	−20.41	−21.50	−20.55	−20.15	−19.02	−19.99
	(0.27)	(0.34)	(0.30)	(0.17)	(0.20)	(0.18)	(0.13)	(0.12)	
2018	−20.06	−16.83	−19.09	−21.84	−20.23	−19.78	−19.76	−19.09	−19.95
	(0.11)	(0.12)	(0.17)	(0.09)	(0.10)	(0.16)	(0.10)	(0.13)	
2019	−20.28	−18.47	−19.70	−21.85	−21.78	−20.74	−19.88	−20.04	−20.78
	(0.11)	(0.12)	(0.13)	(0.11)	(0.10)	(0.07)	(0.16)	(0.13)	

We report estimated price coefficients for each consumer type and year combination for twelve-month purchase history consideration sets. Each estimate is based on a three year block of data, e.g. 2007 estimates are based on 2006, 2007 and 2008 data. Standard errors are reported in parenthesis.

with eleven-month purchase histories to informally check if further reductions in the size of our consideration set proxies have a measurable effect on the coefficient estimates. This is based on the result by McFadden (1978) that having products in the consideration set that are actually not considered leads to estimation bias. The reverse, having proxies that are smaller than the true consideration sets, does not yield biased parameter estimates. Thus, esimation on consideration set subsamples is consistent. If we randomly drop single products from the eleven-month proxies and there is no effect on the preference estimates, we take this as evidence that our proxies have the correct size. If too many products were in the proxies, parameter estimates should change.

We formalize these model comparisons using pairwise coefficient tests. The results are shown in Table A.1 in the Appendix. The coefficients based on the twelve- and eleven-month proxy sets cannot be distinguished from each other statistically. This holds for the entire series of coefficient estimates. The same applies to the comparison between the eleven-month proxy set coefficients without and with random product drops. Therefore, the three proxy set specifications cannot be distinguished. Using the complete purchase histories as proxies, however, leads to pairwise test statistics that clearly reject the null that both sets of estimates are statistically identical. The

Table 4: Full Information Consumer Type Price Coefficient Estimates

Year	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6	Type 7	Type 8	Average
2007	−26.79	−23.11	−24.40	−25.03	−26.33	−24.32	−24.72	−24.40	−25.22
	(0.10)	(0.09)	(0.15)	(0.05)	(0.13)	(0.05)	(0.35)	(0.15)	
2008	−26.78	−23.56	−25.09	−25.01	−26.20	−24.78	−24.75	−25.09	−25.49
	(0.08)	(0.08)	(0.07)	(0.05)	(0.06)	(0.04)	(0.28)	(0.07)	
2009	−27.00	−24.07	−25.31	−26.34	−26.56	−25.45	−25.86	−25.31	−26.09
	(0.08)	(0.09)	(0.16)	(0.04)	(0.16)	(0.05)	(0.20)	(0.16)	
2010	−27.22	−24.67	−26.19	−27.80	−26.60	−28.41	−29.08	−26.19	−26.82
	(0.08)	(0.09)	(0.06)	(0.04)	(0.07)	(0.07)	(0.09)	(0.06)	
2011	−26.75	−25.70	−27.40	−27.57	−26.93	−27.99	−29.04	−27.40	−26.98
	(0.08)	(0.08)	(0.07)	(0.04)	(0.07)	(0.08)	(0.08)	(0.07)	
2012	−26.33	−25.30	−27.00	−26.92	−27.44	−28.47	−29.10	−27.00	−26.75
	(0.08)	(0.09)	(0.07)	(0.06)	(0.07)	(0.10)	(0.08)	(0.07)	
2013	−25.28	−25.21	−27.23	−26.44	−27.89	−27.67	−28.61	−27.23	−26.79
	(0.09)	(0.10)	(0.08)	(0.05)	(0.07)	(0.07)	(0.09)	(0.08)	
2014	−25.83	−25.52	−27.14	−27.29	−27.63	−25.65	−28.55	−27.14	−26.74
	(0.10)	(0.10)	(0.08)	(0.08)	(0.08)	(0.07)	(0.11)	(0.08)	
2015	−25.57	−24.72	−25.23	−26.18	−27.41	−25.71	−27.20	−25.23	−26.27
	(0.11)	(0.19)	(0.18)	(0.06)	(0.08)	(0.06)	(0.13)	(0.18)	
2016	−24.47	−23.37	−24.10	−24.11	−25.23	−24.68	−25.90	−24.10	−24.67
	(0.10)	(0.11)	(0.19)	(0.16)	(0.10)	(0.14)	(0.12)	(0.19)	
2017	−24.50	−23.14	−23.61	−24.39	−26.50	−24.52	−25.94	−23.61	−24.94
	(0.10)	(0.20)	(0.19)	(0.20)	(0.15)	(0.15)	(0.13)	(0.19)	
2018	−24.65	−23.29	−25.45	−25.90	−25.27	−25.83	−26.18	−25.45	−25.12
	(0.10)	(0.11)	(0.14)	(0.09)	(0.08)	(0.11)	(0.17)	(0.14)	
2019	−25.41	−23.74	−26.15	−25.94	−27.13	−26.23	−25.98	−26.15	−25.87
	(0.11)	(0.11)	(0.16)	(0.12)	(0.08)	(0.09)	(0.12)	(0.16)	

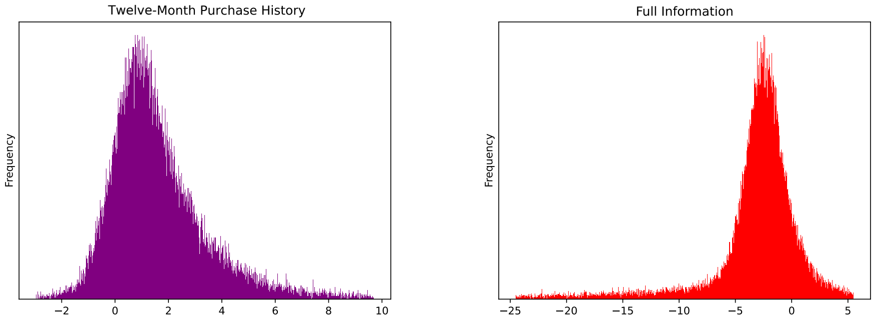
We report estimated price coefficients for each consumer type and year combination for full information consideration sets. Each estimate is based on a three year block of data, e.g. 2007 estimates are based on 2006, 2007 and 2008 data. Standard errors are reported in parenthesis.

full information coefficient estimates are also statistically different from their complete purchase histories counterparts. This leads directly to the question, which proxy specification is better able to explain the observable data. We turn to model selection on the demand (and supply) side further below in Section 5.3.

Tables 3 and 4 compare the coefficient estimates across specifications in detail. All estimated price coefficients are negative and therefore have the economically expected sign. While the changes over time in price sensitivity are highly correlated across specifications, price sensitivity estimates are clearly ranked. Larger consideration set proxies yield larger price coefficients and thereby also more elastic demand. Consumer types show heterogeneity in their price sensitivities. To look at this more closely, we take the year 2007 estimates. Type 1 households are households of higher education and income, male and relatively older than other household types and have relatively large consideration sets (see Table 2 and Table A.2). We estimate type 1 to be more price sensitive than other households in the year 2007. The least price sensitive household type, type 2, is statistically significantly less price sensitive than household type 1 and also less price sensitive than the second least price sensitive consumer type, type 7 (see Table 3).

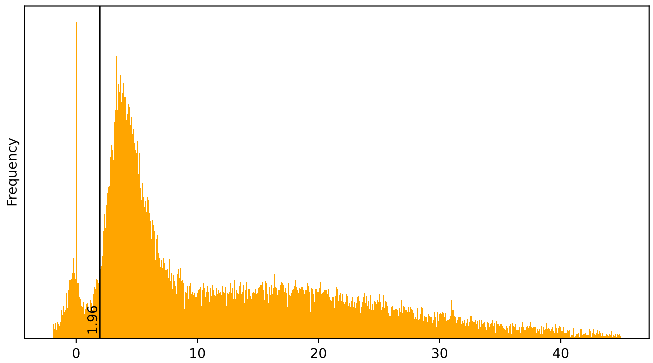
Turning to the estimated product dummies across specifications, Figure 4 shows the distribution

Figure 4: Estimated Product Coefficients for Ready-to-Eat Cereals



Estimated product coefficients, pooled for the years 2006 through 2020 and all RTE cereals products, truncated at the 5% and 99%-quantile to prevent very large negative values in the full information case. The reference product is the composite private label product, which has a normalized utility of zero. We estimate the mean utility of the outside option, which is simply the constant term in consumers' utility specifications.

Figure 5: Testing Differences in Product Preferences Between Models

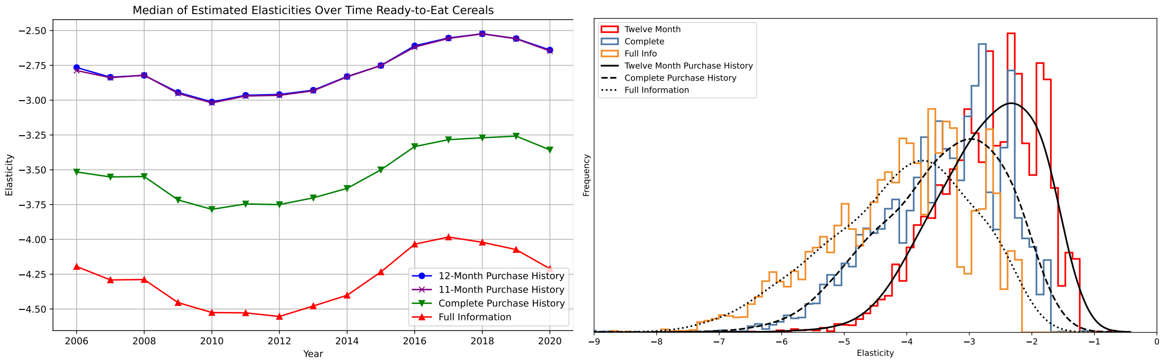


We test for differences in product preferences for the years 2006 through 2020 in the ready-to-eat cereals category. For all estimated consumer product preference values, we conduct pairwise z-score tests of the product preference estimates between the twelve-month past purchase history and the full information consideration set. We plot the resulting distribution of z-scores. The vertical line highlights the threshold for a statistically significant difference in estimates at the 5% significance level, with 92.28% statistically significantly different estimates.

of product coefficients by specification, pooled across consumer-types and years. The left panel presents the product dummy distribution for the twelve-month purchase histories. The bulk of the probability mass lies between –2 and 10, with an average of 1.72. Consumers therefore largely prefer branded products over the comparison group of private label products. The right panel presents the corresponding distribution for the full information specification. In contrast to the limited attention model, most estimated product coefficients are negative and between -25 and 5, with an average of -3.18. These coefficient estimates imply that private label products are preferred to nationally branded products by the bulk of consumers.[22] These results are in line with the findings of Draganska and Klapper (2011). Imposing the full information assumption yields a downward shift in product preferences. Intuitively, if actual consumers have limited attention, their choice probability mass is distributed over relatively narrow sets of

[22]Over time, the share of positive brand coefficients is largely constant. We show details in Appendix Figure B.6.

Figure 6: Estimated Own-Price Elasticities



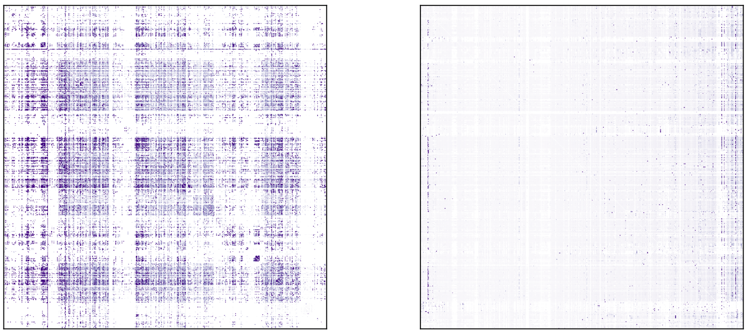
Left: Estimated medians of own-price elasticities for ready-to-eat cereals over time. Estimates are based on three-year rolling averages calculated for each cluster, spanning from 2006 through 2020. Weighted by NielsenIQ sampling weights. **Right:** Estimated product level elasticity distribution by specification and pooled across the years 2006 through 2020 for the ready-to-eat cereal category. The median own-price elasticity is -2.71 (-2.63) for the twelve-month purchase history specificication, -3.34 (-3.32) for the complete purchase history specification and -4.14 (-3.98) for the full information specification. We drop products with less than 1% of sales. Weighted by sales.

products, which leaves very small market shares for many products. With the full information assumption, this can only be rationalized by large differences in observable product attributes between products that have high and low market shares. Consequently, the full information specification yields many product dummy estimates with large negative values.

To assess whether the estimated product preference parameters differ statistically across models, we perform pairwise tests. For each estimated product parameter in each year and cluster, we compare the estimated product preference estimate between the full information and twelve-month purchase histories models. We plot the resulting distribution of z-scores in Figure 5 and highlight the critical value of 1.96 for statistical significance at the 95%-level. We find 92.3% of estimates to be statistically significantly different from each other. The reported differences in product valuations in Figure 4 are therefore also statistically meaningful.

Own-price elasticities Next, we present median own-price elasticities for the eleven-month, twelve-month, complete purchase histories models and the full information benchmark model. Estimated averages are quantitatively similar while being more elastic (+0.25 on average, see Figure B.7 in the appendix). The left panel in Figure 6 shows the development over time and the differences between specifications. Due to the functional form of choice probabilities, the plot correlates strongly with the estimated price coefficients. Thus, the full information specification yields the most elastic demand compared with the limited attention models. For the full information model, the median own-price elasticity is roughly -4.2, while the limited attention model with twelve-month purchasing history proxies delivers a median own-price elasticity of roughly -2.7. Our limited attention elasticities are quantitatively close to those in the existing literature. Döpper et al. (2025) estimate a median elasticity of -2.03 for the year 2007 and -2.97 for the year 2019 (Döpper et al. (2025), Online Appendix, Table E2), Backus et al. (2021)

Figure 7: Comparison of Cross-Elasticies By Consideration Set Specification



(a) Twelve-Month Purchase History (b) Full Information

Comparison of cross-elasticies by Consideration set specification for ready-to-eat cereals and the year 2020. Each point and color per row and column represents a product and its consumer based substitution pattern to other products in the market. The magnitude of cross-substitution is color coded, with white indicating zero values, and darker colors indicating higher values and stronger substitution.

estimate a median elasticity of -2.67 for the years 2007-2017, and the estimated single-product firm markups in Nevo (2001) (p.333, Table 8) imply a median own-price elasticity of -2.79. These studies, however, assume that consumers are fully informed, while we allow for limited attention. Döpper et al. (2025) retain the top 20 products in each category they cover, so that a direct comparison should be taken with a grain of salt. We can show that our full information model estimates align quantitatively closely with those of Nevo (2001) and Backus et al. (2021), if we measure product prices using package unit sales within product definitions instead of taking the simple average over different package prices, and do not use the directly observable household weights and do not account for household-specific coupon use. This is illustrated in Figure B.8. Using the NielsenIQ consumer panel gives us additional information on consumers relative to the scanner data that when controlled for in the estimation yields more elastic consumer demand. Overall, our own-price elasticity estimates are quantitatively close to those in the existing literature.

The right panel in Figure 6 shows the full distribution of own-price elasticities for the limited attention and full information models. Reducing consideration set size from the full information benchmark shifts the entire elasticitity distribution to the right, while all estimated own-price elasticities remain strictly negative.

Cross-price elasticities Markets such as the RTE cereals industry are characterized by a large number of products that is produced by in comparison few and large multiproduct firms. Substitution within and between firm-level product portfolios can therefore be expected to have quantitatively important effects on market outcomes and policy interventions, such as for example merger review or targeted taxation. Figure 7 shows our estimated cross-price elasticities for the twelve-month and full information models. Each panel shows the full matrix and each entry is plotted as a single dot. To visualize the magnitude of the many entries, we apply the same color scale to both panels. The darker the dot, the bigger the magnitude of the corresponding cross-price elasticity. The quantitatively large differences between the full information

Table 5: Elasticities By Specification of Consideration Set

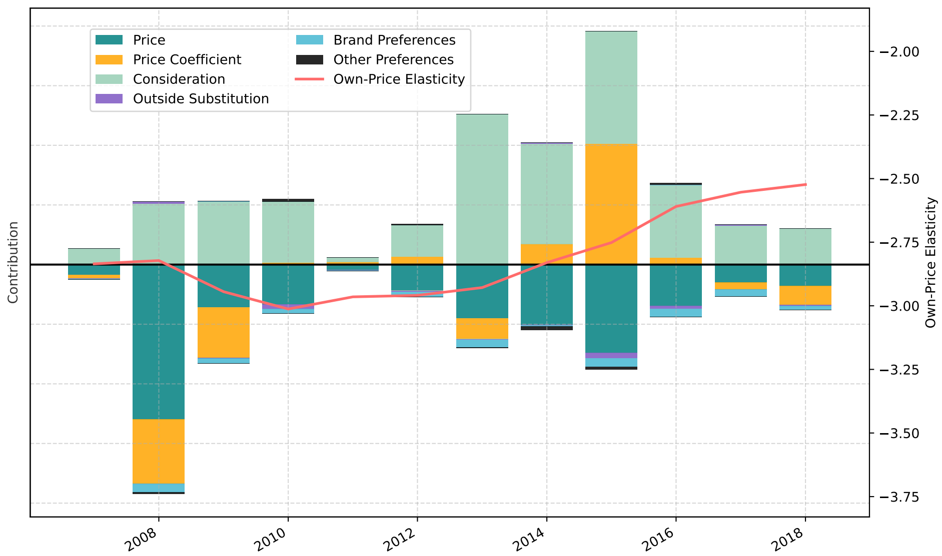
Twelve Month Past Purchase History						
	Brand 1	Brand 2	Brand 3	Brand 4	Brand 5	Brand 6
Brand 1	-3.403	0.410	0.403	0.447	0.264	0.500
Brand 2	0.317	-2.267	0.361	0.163	0.709	0.394
Brand 3	0.348	0.549	-3.517	0.481	0.282	0.553
Brand 4	0.416	0.292	0.443	-2.923	0.367	0.598
Brand 5	0.161	1.100	0.298	0.328	-4.205	0.271
Brand 6	0.376	0.278	0.546	0.632	0.304	-2.786
Complete Purchase History						
	Brand 1	Brand 2	Brand 3	Brand 4	Brand 5	Brand 6
Brand 1	-3.678	0.174	0.202	0.177	0.148	0.238
Brand 2	0.139	-2.348	0.136	0.123	0.150	0.168
Brand 3	0.186	0.163	-3.725	0.206	0.176	0.249
Brand 4	0.192	0.129	0.207	-3.312	0.166	0.262
Brand 5	0.122	0.111	0.167	0.140	-4.294	0.173
Brand 6	0.202	0.140	0.240	0.260	0.187	-3.276
Full Information						
	Brand 1	Brand 2	Brand 3	Brand 4	Brand 5	Brand 6
Brand 1	-4.545	0.013	0.023	0.019	0.026	0.014
Brand 2	0.020	-3.506	0.022	0.018	0.025	0.012
Brand 3	0.019	0.013	-4.244	0.018	0.026	0.014
Brand 4	0.020	0.013	0.023	-3.783	0.026	0.014
Brand 5	0.018	0.012	0.022	0.018	-4.878	0.013
Brand 6	0.020	0.013	0.023	0.019	0.026	-4.198

Comparison of randomly drawn and anonymized estimated elasticities by consideration set specification for the year 2020 in the ready-to-eat cereals market. The values represent the percentage change in row's product market share resulting from a 1% price increase in a column product.

and limited attention substitution patterns is immediately apparent. With limited attention, substitution within markets is quantitatively much more important and concentrated on several clusters of products. In comparison, the full information substitution patterns are much more uniform and it is hard to make out clusters of products that appear to be close substitutes. To allow for closer inspection and provide a direct numerical comparison, we randomly select 6 products and show the corresponding entries in the cross-price elasticity matrices for different informational assumptions in Table 5. The cross-price elasticities based on the twelve-month purchase histories estimates are an order of magnitude larger than those for the full information model. Differences in own-price elasticities are also substantial, but quantitatively more muted. Both illustrations highlight that the informational assumption that we place on consumers has potentially large implications for the implied substitution patterns in the market.

Decomposition over time Next, we decompose the factors determining elasticities over time. We do so by looking at elasticity differences between years. We calculate the decomposition by

Figure 8: Contributing Factors to Elasticity Changes

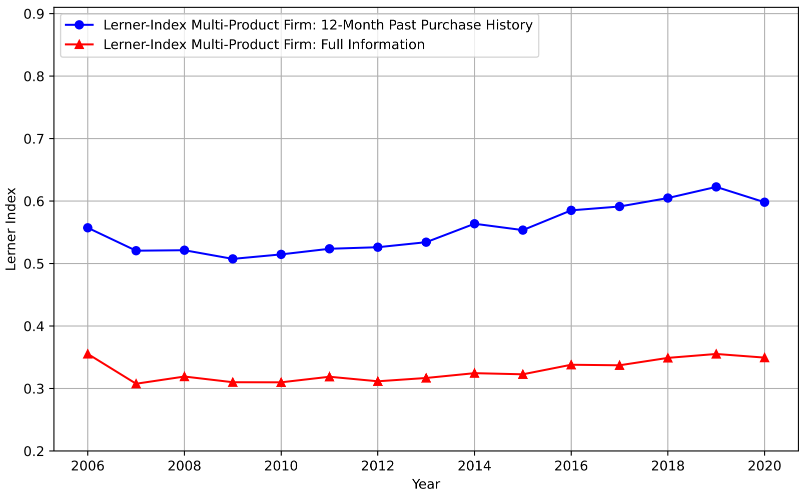


The figure shows a decomposition of elasticity changes in the ready-to-eat cereals industry. The red line indicates the twelve-month past purchase history elasticity estimate (right axis). The bars indicate the contribution to elasticity changes by each given factor. The height of the bars indicate absolute importance of each factor over time, with values below the thick black horizontal line indicating negative values and values above the thick black line indicating positive values.

keeping all factors constant, except one factor which is allowed to change, and observe the impact on the estimated elasticities, and consolidate changes into a percentage contribution of each factor. The factors of interest are price changes, changes in substitution to the outside good, changes in the price sensitivity of consumers, as well as changes in consumer product preferences and additional preferences such as retailer valuations. In Figure 8, we plot the backed out factors scaled on the basis of the absolute change in the elasticity along a line plot for the development of the elasticity over time for the case of the twelve-month purchase history specification. The plot highlights three points: first, price changes, changes in consumer attention and consumer price sensitivity are the most important and driving factors of demand. Second, in the first part of our sample, consumers are becoming more sensitive to changes in prices, driving consumer responses, whereas in the second half of the observation period, consumers are becoming less price sensitive on average. Third, while price changes drive variation in consumer elasticity, consideration remains a stable factor that shifts consumers to become less elastic over time.

Estimated markups Last, we show estimated markups for all products in the ready-to-eat cereal category over time, expressed as the Lerner-Index. We define the Lerner-Index as $L_j = (p_j - c_j)/p_j$, where the marginal costs c_j are calculated according to (8). Figure 9 shows the estimated multi-product level markups over time for the twelve-month purchase history specification and full information. Estimated markups drop from 0.56 (0.36) in the year 2006 to 0.52 (0.31) in 2009, and increase from that point onward steadily up to the year 2019 for the twelve-month purchase history (full information). In 2020, we observe a decline. Overall, markups increase by roughly 10 percentage points for the twelve-month purchase history specification and 5 percentage points for the full information specification. The markups of the

Figure 9: Measuring Markups Over Time: Ready-to-Eat Cereals



We report the evolution of median markups over time in the ready-to-eat cereals industry, as measured by the Lerner-Index, defined as $L_j = (p_j - c_j)/p_j$. Model-implied markups are based on the estimates reported in Tables 3 and 4.

full information specification are roughly 25 percentage points lower than the limited attention specification using twelve-month purchasing history proxies.

5.3 Model Selection

Based on our estimates for demand and supply, we now move on to test demand and supply combinations against each other. We count four models that may describe our data best: i) consumers' demand is best described by twelve-month past purchase history and firms' supply behavior is best described by knowing this demand pattern and optimizing accordingly ii) consumers' demand is best described by twelve-month past purchase history, but firms' supply behavior ignores this demand pattern and optimizes assuming consumers have full information iii) consumers' demand is best described by full information and firms' supply behavior is best described by knowing this demand pattern and optimizing accordingly as well as iv) consumers' demand is best described by full information, but firms' supply behavior ignores this demand pattern and optimizes assuming consumers's demand is best described by twelve-month past purchase history.

To be able to choose among our (nonnested) models, we follow Gasmi et al. (1992) and apply pairwise likelihood ratio tests for nonnested hypotheses introduced by Vuong (1989). These tests let us determine which of the above proposed consumer and firm behavior explains the data best. The null hypothesis is that two competing models adjust the data equally well and the alternative hypothesis is that one model fits better. The tests are symmetric, directional, and do not require either model to be correctly specified. We calculate the likelihood ratio statistic for each pair of candidate models according to

$$n^{-\frac{1}{2}}\frac{LR_n(\hat{\theta}_1,\hat{\theta}_2)}{\hat{w}_n}=n^{-\frac{1}{2}}\frac{\mathcal{L}^1_n(\hat{\theta}_1)-\mathcal{L}^2_n(\hat{\theta}_2)}{\hat{w}_n} \tag{10}$$

with $\mathcal{L}^m_n(\hat{\theta}_m)=\sum w\log f(y|x;\theta_m)$ and $f(\cdot)$ representing the model implied likelihood contributions, with $n=n^d+n^s$, or the total number of observations determined by the sum of demand and supply side observations. We have suppressed parameter arguments for the sake of notational brevity. The normalization term, $\hat{\omega}_n$, is obtained from the respective models' likelihood contributions and the test statistic follows a standard normal distribution under the null of equal model fit for nonnested models and standard regularity conditions (Vuong, 1989). The model specific marginal costs are denoted by c_j and yield likelihood contributions according to $f^s_j=[2\pi\sigma^2]^{-1/2}\exp(-(c_j-x_j\gamma)^2/(2\sigma^2))$, where x_j are the product attributes and additional controls that we also use on the demand side and γ estimated coefficients. Demand-side likelihood contributions are given by consumer purchase choice probabilities, or $f^d=Prob(\cdot)$. We apply equal weights to demand and supply by giving appropriate weights to each term in $L^m_n(\cdot)$, with $w^d=n^s/(n^s+n^d)$ and $w^s=n^d/(n^s+n^d)$.

Before we provide our test results for demand and supply, we check the demand side only. To formalize an evaluation of the different proxies, we test the demand systems against each other on the implied likelihood contributions using above test procedure, but with observations from the demand side only. We select the model based on the twelve-month purchase history as the preferred model. Table 6 shows the results of the pairwise likelihood ratio tests. An absolute value exceeding 1.96 provides evidence that the models are statistically distinguishable at the 5% significance level. Negative values favor the row model, whereas positive values favor the column model. The full information model is always rejected in favor of the complete purchase history model. The complete purchase history model is rejected by the monthly purchase history models, and the twelve-month purchase history model is preferred over the eleven-month purchase history model. Our preferred proxy specification is therefore based on each household's directly observed purchase history over the past twelve months.

Table 6: Testing Models of Demand

	Twelve-Month	Complete Purchase History	Full Information
Eleven-Month	11.95	−1314.35	−1626.02
Twelve-Month		−1313.06	−1625.19
Complete Purchase History			−1370.46

We conduct pairwise tests for for each combination of demand models. The numbers in the table represent model pairwise normalized likelihood ratio statistics based on the Vuong (1989) likelihood ratio testing framework. An absolute value exceeding 1.96 provides evidence that the models are statistically distinguishable at the 5% significance level. Negative values favor the row model, whereas positive values favor the column model.

Then, we conduct pairwise tests for each model combination of demand and supply. In the column headers of Table 7, demand is shortened to "d" and supply is shortened to "s". The numbers in the table represent model pairwise normalized likelihood ratio statistics based on Vuong (1989). An absolute value exceeding 1.96 provides again evidence that the models are statistically distinguishable at the 5% significance level. Negative values favor the row model, whereas positive values favor the column model. Each combination of models stacks demand and supply side likelihood contributions, where the demand side likelihood contributions ensue from respective demand estimation choice probabilities, and supply side likelihood contributions are obtained from an auxiliary regression of marginal costs on explanatory variables. For the entry in the top left cell, we test likelihood contributions of the limited information model for supply and demand (row) against likelihood contributions of demand with supply side likelihood contributions based on the full information demand and supply (column).

The results can be seen in Table 7. We observe that the full information demand specification is always rejected in favor of the limited attention demand specification. For example, the value of -1590.77 strongly rejects the demand side full information model in favor of the limited attention model. If demand is symmetric in the candidate models to be tested, the likelihood ratio test selects the full information price setting on the supply side (17.07 in the top left or bottom right cell).

Table 7: Tests for Model Selection

	d: limited info s: full info	d: full info s: limited info	d: full info s: full info
demand: limited info, supply: limited info	17.07	-1590.77	-345.36
demand: limited info, supply: full info		-380.73	-1590.77
demand: full info, supply: limited info			17.07

Pairwise test results. In the column headers, demand is shortened to "d" and supply is shortened to "s". We report pairwise normalized likelihood ratio statistics based on Vuong (1989). Negative values favor the row model, whereas positive values favor the column model. For the entry in the top left cell, we test likelihood contributions of the limited information model for supply and demand (row) against likelihood contributions of demand with supply side likelihood contributions based on the full information demand and supply model (column).

Given the likelihood ratio test results, we are confident that consumer decision making is well approximated by models of limited attention, and firms' price setting is aligned with a model of full information price setting. Nevertheless, we caution and plan to run robustness exercises in several aspects. First and foremost, we use demand side consumer level data, whereas we use aggregate product level reponses on the supply side in our tests. We balance this by applying equal weights to both demand and supply. This is not optimal. In the future, we plan to aggregate consumer responses to the product level and test directly on the balanced panel.

6 Conclusion

We adopt the approach of Crawford et al. (2021) and estimate demand for ready-to-eat (RTE) cereals using data from the US. The approach builds on McFadden (1978), who shows that consumer preferences can be consistently estimated based on subsets of the available choice menu, as long as the subsets belong to consumers' true consideration sets and we do not include too many products. Based on household-level data, we generate three proxies for unobserved consideration sets using observed purchase histories and discretize consumers into types using k-means clustering. We use maximum likelihood to estimate demand under the full-information assumption and various models with limited attention. To address price endogeneity, we exploit the high frequency of our consumer purchase data by including product and time controls, as well as observable coupon use.

Using household and scanner data for the period 2006–2020, we find that estimates obtained under full information yield more elastic demand and therefore lower markups than those obtained under limited consumer attention. Substitution between products, measured by cross-price elasticities, are an order of magnitude larger in our most preferred limited attention specification compared with the full information benchmark. The benchmark model is statistically rejected against all limited attention models. Our most preferred proxy consideration set specification is based on consumers' purchase history over the past 12 months. These proxies have on average six products in consumers' consideration sets, while the total choice menu has on average 153 products to choose from. Additional model selection tests also indicate that firms on average expect consumers to be fully informed when setting prices.[23]

These findings stress that the informational assumption placed on consumers in differentiated product markets has a large impact on the model-implied observed and counterfactual market outcomes. Reliable predictions of counterfactual changes in these market settings require a careful analysis of consumers' informational awareness. Having ten times larger substitution effects between products can be expected to yield substantially different price effects of a horizontal merger, for example.

The combination of proxy consideration sets and high frequency demand estimation in the presence of sticky prices is well suited to multi-category estimation as in Atalay et al. (2023) and Döpper et al. (2025). In future research, we aim to explore whether there are systematic changes in consumer attention over time in US consumer packaged goods markets and if so how these changes affect consumer welfare and firm profitability.

References

Abaluck, Jason and Abi Adams-Prassl, "What do consumers consider before they choose? Identification from asymmetric demand responses," *The Quarterly Journal of Economics*, 2021, *136* (3), 1611–1663.

Atalay, Enghin, Erika Frost, Alan T Sorensen, Christopher J Sullivan, and Wan-jia Zhu, "Scalable demand and markups," Technical Report, National Bureau of Economic Research 2023.

Backus, Matthew, Christopher Conlon, and Michael Sinkinson, "Common ownership and competition in the ready-to-eat cereal industry," Technical Report, National Bureau of Economic Research 2021.

Berry, Steven and Panle Jia, "Tracing the woes: An empirical analysis of the airline industry," *American Economic Journal: Microeconomics*, 2010, *2* (3), 1–43.

[23]We did not test for heterogeneity in firm expectations about consumers' attention. This is an interesting question for future research.

___ , **James Levinsohn, and Ariel Pakes**, "Automobile prices in market equilibrium," *Econometrica*, 1995, *63* (4), 841–890.

Berry, Steven T and Philip A Haile, "Foundations of demand estimation," in "Handbook of industrial organization," Vol. 4, Elsevier, 2021, pp. 1–62.

Bils, Mark and Peter J Klenow, "Some evidence on the importance of sticky prices," *Journal of political economy*, 2004, *112* (5), 947–985.

Blundell, Richard and Rosa L Matzkin, "Control functions in nonseparable simultaneous equations models," *Quantitative Economics*, 2014, *5* (2), 271–295.

Conlon, Christopher, Nirupama Rao, and Yinan Wang, "Who pays sin taxes? understanding the overlapping burdens of corrective taxes," *Review of Economics and Statistics*, 2024, *106* (6), 1719–1729.

Crawford, Gregory S., Rachel Griffith, and Alessandro Iaria, "A Survey of Preference Estimation with Unobserved Choice Set Heterogeneity," *Journal of Econometrics*, 2021, *222* (1), 4–43.

Dardanoni, Valentino, Paola Manzini, Marco Mariotti, and Christopher J. Tyson, "Inferring Cognitive Heterogeneity From Aggregate Choices," *Econometrica*, 2020, *88* (3), 1269–1296.

DellaVigna, Stefano and Matthew Gentzkow, "Uniform pricing in us retail chains," *The Quarterly Journal of Economics*, 2019, *134* (4), 2011–2084.

Döpper, Hendrik, Alexander MacKay, Nathan H Miller, and Joel Stiebale, "Rising markups and the role of consumer preferences," *Journal of Political Economy*, 2025, *133* (8), 2462–2505.

Draganska, Michaela and Daniel Klapper, "Choice set heterogeneity and the role of advertising: An analysis with micro and macro data," *Journal of Marketing Research*, 2011, *48* (4), 653–669.

Gasmi, Farid, Jean Jacques Laffont, and Quang Vuong, "Econometric analysis of collusive behaviorin a soft-drink market," *Journal of Economics & Management Strategy*, 1992, *1* (2), 277–311.

Hauser, John R and Birger Wernerfelt, "An evaluation cost model of consideration sets," *Journal of consumer research*, 1990, *16* (4), 393–408.

Honka, Elisabeth, Ali Hortaçsu, and Maria Ana Vitorino, "Advertising, consumer awareness, and choice: Evidence from the US banking industry," *The RAND Journal of Economics*, 2017, *48* (3), 611–646.

Karadi, Peter, Juergen Amann, Javier Sánchez Bachiller, Pascal Seiler, and Jesse Wursten, "Price setting on the two sides of the Atlantic-Evidence from supermarket scanner data," *Journal of Monetary Economics*, 2023, *140*, S1–S17.

McFadden, Daniel, "Modeling the Choice of Residential Location," in "Spatial Interaction Theory and Planning Models," Vol. 1, North-Holland: Karlqvist, A., Lundqvist, L., Snickars, F., Weibull, J., 1978, pp. 75–96.

Nakamura, Emi and Jón Steinsson, "Five facts about prices: A reevaluation of menu cost models," *The Quarterly Journal of Economics*, 2008, *123* (4), 1415–1464.

Nevo, Aviv, "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 2001, *69* (2), 307–342.

Sovinsky, Michelle, "Limited information and advertising in the US personal computer industry," *Econometrica*, 2008, *76* (5), 1017–1074.

Vuong, Quang H, "Likelihood ratio tests for model selection and non-nested hypotheses," *Econometrica: journal of the Econometric Society*, 1989, pp. 307–333.

Appendix

A Tables

A.1 Testing Preference Estimates

We test price coefficient preference estimates by pairwise tests for statistically significant differences for each year and combination of specification and report the results in Table A.1. The reported values are absolute z-scores. A value greater than 1.96 implies statistically significant differences between estimates of respective specifications (5%, two-tailed). The specifications are shorthand "12m" for twelve-month past purchase history, "11m" for eleven-month past purchase history, "CPH" for complete past purchase history, "FI" for full information consideration sets, and "RD" for random products dropped from the eleven-month past purchase history for each consumer and choice occasion.

Table A.1: Testing Preference Estimates: Z-Scores

Year	12m-11m	11m-RD	CPH-12m	CPH-11m	FI-CPH	FI-12m	FI-11m
2007	0.667	0.484	39.271	37.951	31.164	66.937	65.118
2008	0.080	1.141	66.624	74.461	49.481	108.602	121.578
2009	0.458	0.406	44.541	46.360	30.130	71.806	75.177
2010	1.040	1.303	81.391	80.247	61.319	139.545	138.354
2011	0.657	1.242	80.827	79.209	65.274	144.201	141.779
2012	0.358	1.508	69.077	68.780	51.385	130.994	130.711
2013	0.312	1.258	59.437	54.137	50.401	119.070	106.180
2014	0.160	1.065	46.969	51.610	41.232	87.402	98.471
2015	0.078	1.109	40.640	41.518	37.193	68.421	69.926
2016	0.467	1.279	34.101	33.688	23.889	57.303	56.939
2017	0.048	1.173	27.535	26.677	19.847	46.025	44.533
2018	0.036	1.715	50.534	52.600	36.117	83.869	87.447
2019	0.216	1.811	51.164	53.466	32.098	86.215	90.508

A.2 Complete Purchase History Statistics

We report statistics on consumer limited attention for ready-to-eat cereals in Table A.2. The reported statistics are: the average number of unique products for each consumer type which describes the within type average size of consumers' consideration sets for the complete purchase history. We normalize the household specific number of unique products by equivalized household size, reported in the second column. Normalization uses standard equivalization factors to account for household composition. The raw statistic is reported in the right column. We round to two decimals. Standard deviations are reported in parenthesis.

Table A.2: Number of Products of Sufficient Sets

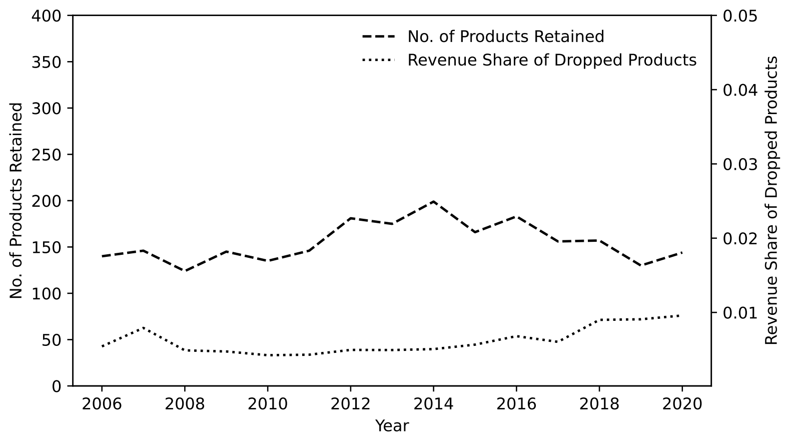
	No. of Products (Normalized)		Average Household Size (Equivalized)		No. of Products (Raw)	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Type 1	11.44	12.13	1.01	0.07	11.55	12.25
Type 2	10.52	9.68	1.01	0.07	10.63	9.78
Type 3	9.48	8.50	1.35	0.55	12.80	11.47
Type 4	9.06	6.98	2.40	0.38	21.75	16.74
Type 5	8.85	7.55	1.95	0.49	17.25	14.72
Type 6	7.89	6.51	1.89	0.49	14.92	12.30
Type 7	7.58	7.20	1.97	0.67	14.94	14.18
Type 8	5.53	5.08	1.79	0.30	9.89	9.10

B Figures

B.1 Dropped Products Revenue Share and Number of Retained Products

We drop products that have very small market shares which results in unreasonable markup and marginal cost estimates and report a summary in Figure B.1. We report the number of retained products and the revenue share of dropped products for the ready-to-eat cereal category on the right x-axis. The revenue share of dropped products is below 1% of market revenue for all years. We report the number of products retained, which is displayed on the left y-axis. We retain 153 products on average.

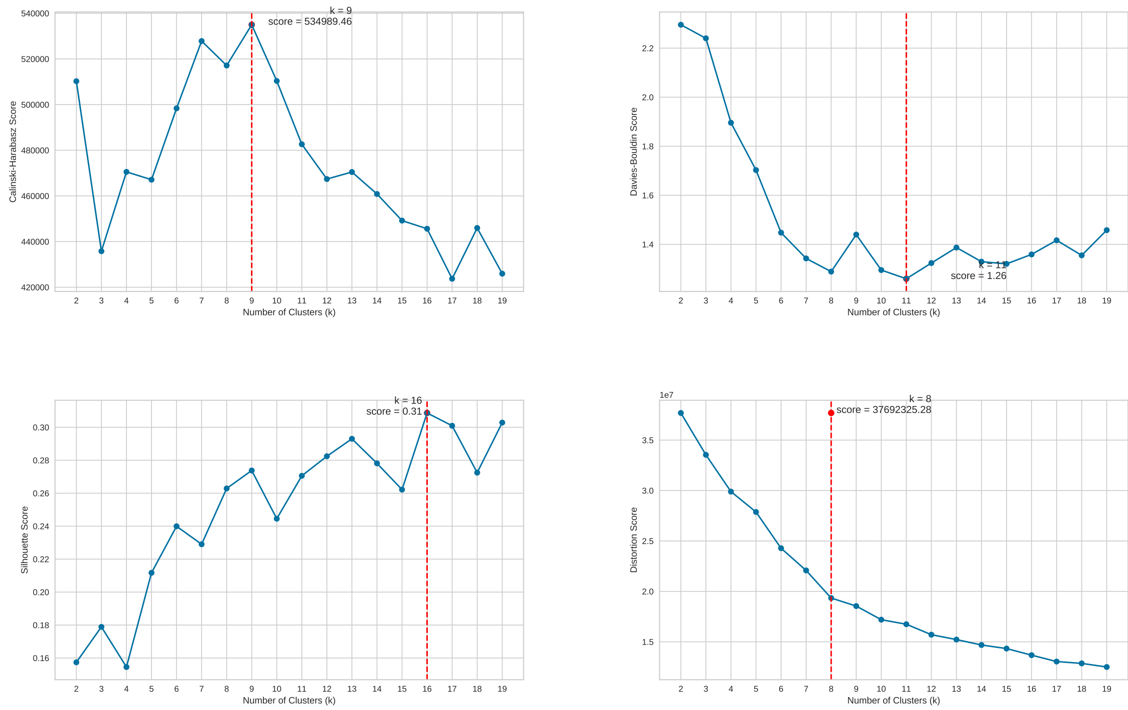
Figure B.1: Dropped Products Revenue Share and Number of Retained Products



B.2 Determining the Optimal Number of Consumer Types

We determine the optimal number of consumer types based on household characteristics using different metrics and show the results in Figure B.2. We let each method decide on the optimal number of consumer types out of a possible range of 2 to 20 possible consumer types. Increasing the upper bound yields similar results. We scale the data up to the population level using NielsenIQ sampling weights. The variables that we use are household size, income, household composition (married, living alone, living with relatives/non-family), dummy variables for single, widowed, separated, living with underage child in household and ethnicity, as well as age, education and a dummy for elderly.

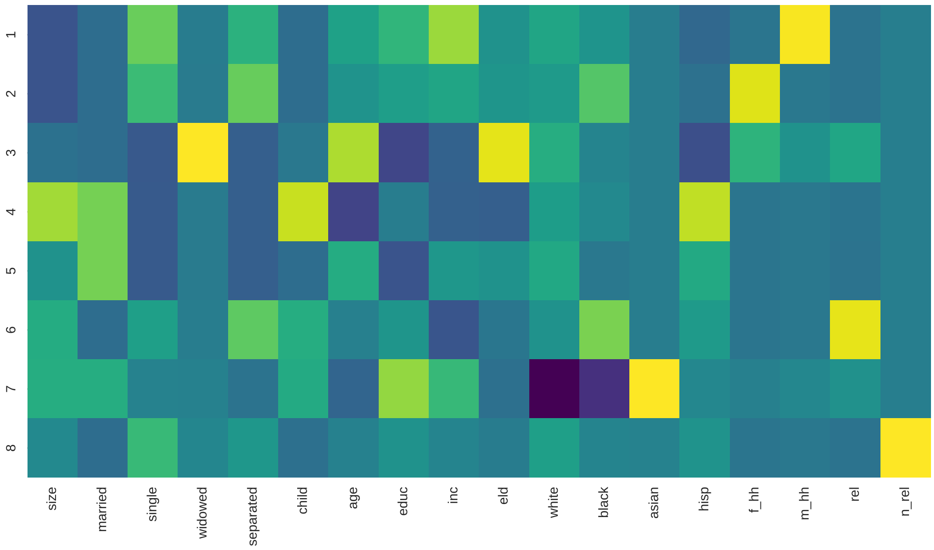
Figure B.2: Optimal Number of Clusters By Statistic



B.3 Heatmap of Consumer Types

We provide a representation of households types in a heatmap with 8 clusters, representing normalized comparisons of cluster variable means. The results can be seen in Figure B.3. The variables are: household size (size), age by household head, income (inc), education (edu), indicator variables for married, single, widowed or separated living households, the presence of underage son/daughter (child), elderly households (eld), ethnicity, households comprising only a female (f_hh) or male (m_hh) or households living with other related (rel) or other unrelated persons (n_rel). Weighted by NielsenIQ household weights. Data used spans the period from 2006 through 2020.

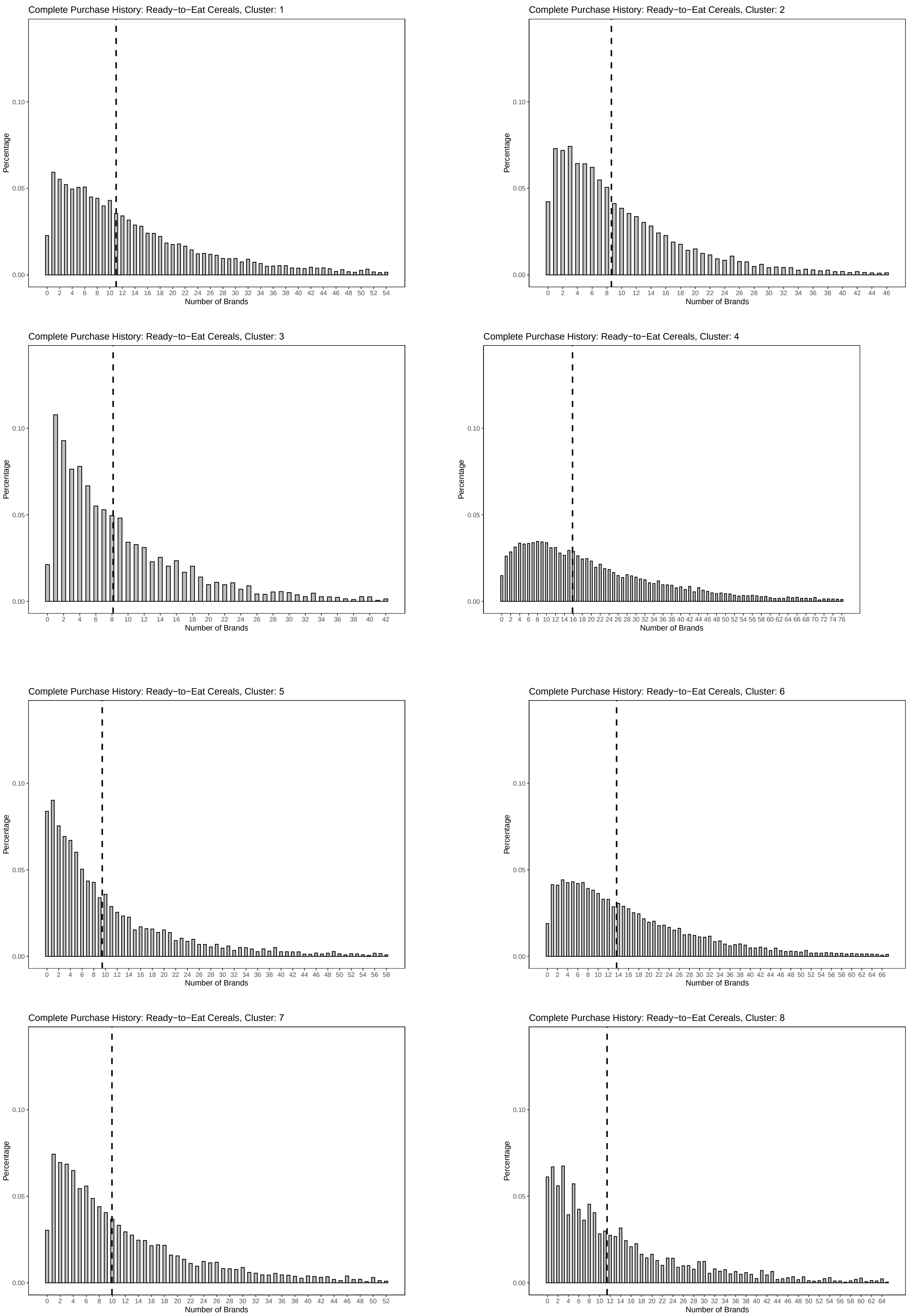
Figure B.3: Heatmap of Household Cluster Means



B.4 Household Sufficient Set Statistics By Cluster

We show sufficient set statistics using the past purchase history to calculate sufficient sets for ready-to-eat cereals in Figure B.4. Each subplot shows for respective cluster the percentage of consumers within the cluster buying 0, 1, 2, etc. unique products of cereals, using the complete purchase history. Dashed vertical lines indicate the weighted average number of products bought conditional on the cluster and a category purchase. NielsenIQ consumer panel data (years 2006 up to and including 2020). Vertical scale is fixed between 0 and 0.15 for all subplots. All figures weighted by NielsenIQ household weights.

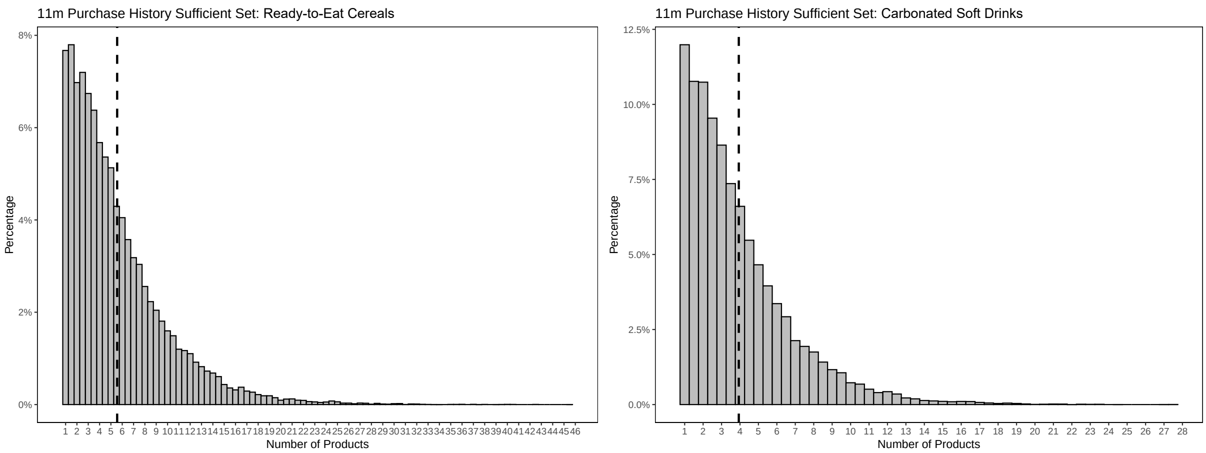
Figure B.4: Household Sufficient Set Statistics By Cluster For Cereals



B.5 Eleven-Month Purchase History Sufficient Sets

We show sufficient set statistics using the purchase history to calculate sufficient sets in Figure B.5. Each subplot shows the percentage of consumers buying 1, 2, 3, etc. unique products of the specified category, using the complete purchase history, the twelve-month purchase history and eleven-month purchase history. Dashed vertical lines indicate the weighted average number of products bought conditional on purchasing a category product. NielsenIQ Panel data (years 2006 up to and including 2020). All figures weighted by NielsenIQ household weights.

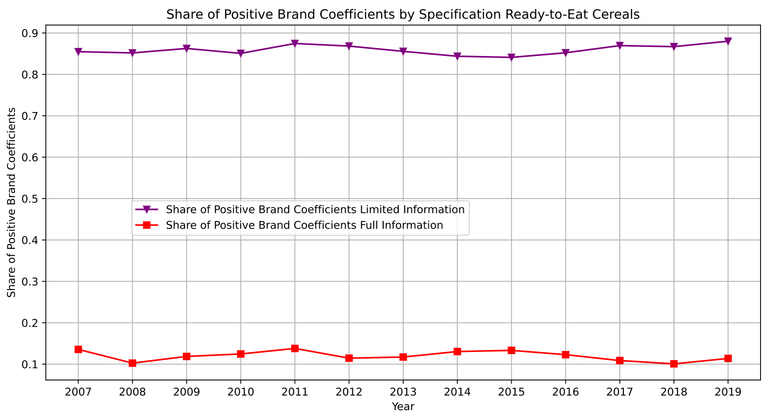
Figure B.5: Household Sufficient Set Statistics



B.6 Share of Positive Product Coefficients Over Time By Specification

For each year, we report the share of positive product coefficients for the ready-to-eat cereals category in Figure B.6.

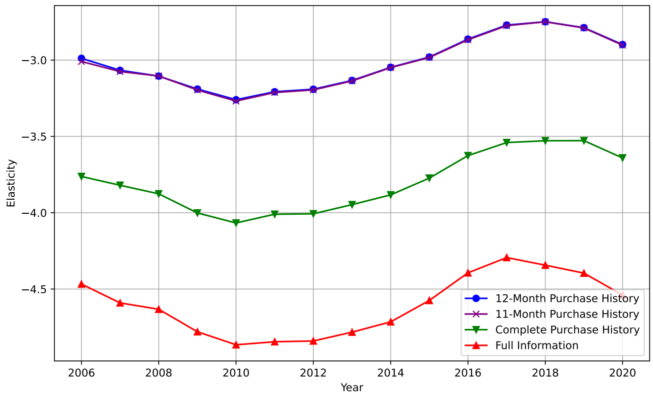
Figure B.6: Share of Positive Product Coefficients Over Time By Specification



B.7 Average Own-Price Elasticities for Ready-to-Eat Cereals

We show estimated average own-price elasticities for ready-to-eat cereals over time in Figure B.7. Estimates are based on three-year rolling averages calculated for each cluster, spanning from 2006 through 2020. Weighted by NielsenIQ sampling weights.

Figure B.7: Estimated Average Own-Price Elasticities



B.8 Estimated Elasticities for Ready-to-Eat Cereals and Full Information by Configuration

We show estimated own-price elasticities for ready-to-eat cereals over time in Figure B.8. Estimates are based on three-year rolling averages for each consumer type, spanning from 2006 through 2020. All estimates are full information estimates. Unweighted ML refers to maximum likelihood estimation that does not weigh by NielsenIQ panelist weights and does not weigh by the product quantity purchased by an household. Simple average prices / sales-weighted prices refers to the weighting scheme when averaging prices at the UPC-level to aggregate to the product level. Elasticities are not corrected with consistent preference estimates based on eleven-month past purchase history. No Coupons and With Coupons indicates the exclusion / inclusion of consumer coupon use.

Figure B.8: Estimated Elasticities for Ready-to-Eat Cereals and Full Information by Specification