DIW BERLIN

1681

Discussion
Papers

Inference of Consumer Consideration Sets

Anna Lu

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

Inference of Consumer Consideration Sets

Anna Lu[*]

Abstract

When consumers face a large number of alternatives, they tend to simplify the decision problem by reducing the number of available alternatives to a subset of relevant alternatives, i.e. a consideration set. Since consideration sets are typically unobserved, most studies in the demand literature have to assume a consideration model. If these consideration models are misspecified, the demand estimates can be biased. In this paper, we develop an approach to formally test any two competing models of consideration against one another in order to determine which model fits the data best. Our test follows the intuition of a menu approach and uses supplemental data on marginal cost-shifters to construct overidentifying restrictions. We show an application to German retailing of coffee and milk. We find that consideration sets are fundamentally different for coffee and milk, and relate our findings to differences in demand and supply conditions of the two product categories.

JEL: D12, L13, M31

Keywords: Demand Estimation, Consideration Set, Retailing

1 Introduction

In the modern market place, consumers face a large variety of products. While people generally value variety, the proliferation of alternatives may pose a complicated decision problem: Consumers need to engage in costly search in order to evaluate and compare alternatives all the while being constrained by cognitive limitations. To simplify the decision problem, consumers have been found to reduce the global set of objectively available alternatives to a subset of "relevant" alternatives. In the marketing and psychology literature, this concept is well established, and the smaller subset of relevant alternatives is typically referred to as the "consideration set" (Howard and Sheth, 1969; Bettman, 1979; Hauser and Wernerfelt, 1990; Roberts and Lattin, 1997; Malhotra, Peterson and Kleiser, 1999).

Due to their cognitive nature, consideration sets are typically unobserved. Consequently, most studies in the demand estimation literature have to assume a model of consideration. For example, the economics literature usually assumes that consumers consider the global set of alternatives (Berry, 1994; Berry, Levinsohn and Pakes, 1995; Nevo, 2000) whereas the marketing literature often uses a two-staged consideration set approach.[1] Importantly, both literatures generally motivate their choice of the consideration model with intuition but rarely support it with statistical evidence.

It is important to choose a consideration model that closely matches actual consumer behavior. This is because misspecified models of consideration can lead to biases in the demand estimates (shown for example by Bronnenberg

---

[1] See for example Allenby and Ginter (1995), Bronnenberg and Vanhonacker (1996), Chiang, Chib and Narasimhan (1998), Draganska and Klapper (2011), and Barroso and Llobet (2012).

and Vanhonacker, 1996; Sovinsky, 2008; Draganska and Klapper, 2011; Conlon and Mortimer, 2013), and this bias will carry over to supply side estimates and policy evaluations because they require demand estimates as an input.

In this paper, we propose a framework which is able to formally test competing models of consideration against one another. Our test follows the intuition of the so-called "menu approach" which is used to infer unobserved firm conduct and compares the equilibrium outcome in an industry to theoretical predictions of a finite set of alternative models . Our test has relatively modest data requirements. In addition to sales data, it requires only data on marginal cost-shifters. At least on an aggregate level, such data is widely available for many industries.

We illustrate our approach in an application to the grocery retailing industry. Specifically, we test the model of global consideration sets against a two-stage model of consideration. We apply our test to the categories of milk and coffee, both of which have been extensively studied in the literature.[2] Our results show that the consideration process fundamentally differs across product categories: While the assumption of global consideration sets performs well in the market for coffee, it performs poorly in the market for milk. Instead, buyers of milk seem to consider milk only at the store at which they are currently shopping. We explain this discrepancy between the two markets with differences in demand and supply conditions, for example in terms of consumer perception of the product category, the level of product differentiation, retailer pricing, and advertising. Our results suggest that the assumption of global consideration sets is better suited for hedonic goods like

---

[2] See for example Guadagni and Little (1983), Krishnamurthi and Raj (1988), Draganska, Klapper and Villas-Boas (2010), Draganska and Klapper (2011), and Bonnet and Bouamra-Mechemache (2015).

coffee or wine, i.e. goods that provide emotional responses like excitement or pleasure. In contrast, the two-stage model is better suited for utilitarian goods, i.e. primarily functional goods like milk, sugar, or flour.[3]

The remainder of this paper is organized as follows: First, we give a brief overview of the related literature in Section 2. We develop our model in Section 3 and describe the data and patterns of consumer behavior in Section 4. In Section 5, we describe the identification strategy, our estimation procedure, and how we allow for household heterogeneity in our estimation. We present and discuss the estimation results in Section 6. Finally, we conclude in Section 7.

2   Related Literature

We contribute to a large literature in economics and marketing that aims to infer consumer preferences from revealed choices. In this literature, discrete-choice methods have gained wide-spread use (for a review see Train (2009)). The central premise of discrete-choice models is that consumers are utility-maximizers, i.e. when faced with a finite number of alternatives, they choose the alternative that gives them the largest utility. A discrete-choice model has to specify two things: Firstly, it needs to specify the utility function in the form of parametric and distributional assumptions. Secondly, it needs to specify the set of alternatives from which the consumer makes her choice.

Economics and marketing have traditionally made different assumptions on the set of products that a consumer considers. In economics, it is typically assumed that consumers consider the global set of products (e.g. Berry, 1994;

---

[3] The difference between hedonic and utilitarian goods is well-established in the marketing literature. For a discussion see Holbrook and Hirschman (1982) or Dhar and Wertenbroch (2000).

Berry, Levinsohn and Pakes, 1995; Nevo, 2001). This modeling assumption is, to a large part, driven by the limited availability of individual-level data in many economics applications. In contrast, a large share of the marketing literature studies consumer packaged goods for which detailed individual-level is often available; marketing researchers thus tend to be able to use richer models of consumer choice. A dominant belief in marketing is that consumers seek to simplify their decision problem by reducing the set of objectively available options to a subset of "relevant" options. The actual choice is then made only from this subset, i.e. the so-called "consideration set".[4] Marketers use two-staged models of consideration to study the determinants of consideration sets, such as advertising (Allenby and Ginter, 1995; Mitra, 1995; Sovinsky, 2008; Draganska and Klapper, 2011; Honka, Hortaçsu and Vitorino, 2017), promotions (Siddarth, Bucklin and Morrison, 1995), or search costs (Mehta, Rajiv and Srinivasan, 2003; De los Santos, Hortaçsu and Wildenbeest, 2012; Seiler, 2013).

What complicates demand estimation is the fact that consideration sets are rarely observed and therefore prone to misspecification. This in turn can bias demand estimates (shown by, for example, Bronnenberg and Vanhonacker, 1996; Sovinsky, 2008; Draganska and Klapper, 2011; Conlon and Mortimer, 2013). A small literature circumvents this problem by collecting direct information on consideration sets. This is typically done via questionnaires in which participants state which products they considered, for example for hypothetical purchases in a virtual supermarket (Van Nierop et al., 2010), or coupled with actual purchase decisions, e.g. in the German coffee market (Draganska and Klapper, 2011) or in the U.S. auto insurance industry (Honka,

---

[4] For a review of the literature on consideration sets see Roberts and Lattin (1997) or Malhotra, Peterson and Kleiser (1999).

2014). In related work, Bruno and Vilcassim (2008) and Conlon and Mortimer (2013) combine sales data with information on the physical availability of products in supermarkets and vending machines, respectively.

A recent strand of the literature proposes methods for demand estimation under unobserved consideration sets that does not require the specification of the consideration process. Lu (2016) develops a moment inequalities approach which returns bounds on preference estimates. The basic idea of his approach is that if a consumer chooses product j, the true consideration set must be bounded by the largest and the smallest possible consideration set that contains j. Crawford, Griffith and Iaria (2017) develop an estimation approach for panel data which returns point estimates. It requires that past choices carry a sufficient amount of information about present-day consideration sets. The authors propose multiple scenarios of sufficient intertemporal correlation of consideration sets, e.g. once a consumer chooses a product, it remains in her consideration set for all subsequent periods.

We develop a novel approach that complements the existing methods. For given choice data, our method is able to pick the best-fitting consideration model from a finite set of competing consideration models. Our method follows the intuition of a so-called menu approach. This approach was developed to test for unobserved competitive conduct in an industry. Its basic idea is that it compares the equilibrium outcome in an industry to the theoretical predictions of a finite set (a "menu") of alternative models of competition, and uses a model selection test to identify the model which matches the data best. Our approach is most closely related to Villas-Boas (2007) who tests for different models of vertical relationships. Other notable examples include

Bresnahan (1987), Gasmi, Laffont and Vuong (1992), Kadiyali, Vilcassim and Chintagunta (1996), and Nevo (2001).

Compared to the approach of Crawford, Griffith and Iaria (2017), our approach has less restrictive assumptions on the intertemporal correlation between consideration sets. Since it is able to test any two consideration models against one another, it is able to accommodate models with any intertemporal structure. Compared to the bounds approach of Lu (2016), our approach allows for point identification of demand parameters. However, the performance of our method critically relies on the quality of the menu of consideration models. This is because we cannot identify the true model but only the *best model within the menu.* Generally, our method is best for markets in which we can make reasonably good guesses of consideration models. When there is evidence on how consideration sets are linked across time, the approach of Crawford, Griffith and Iaria (2017) performs better. When a market is generally not well understood, the method of Lu (2016) is preferable because it makes the weakest assumptions.

3 Model

In this paper, we test two prominent demand models against each other: a single-stage mixed logit and a more flexible two-stage mixed logit. The former is the standard model in empirical industrial organization, the latter is popular in marketing research. In the following, we develop both models and derive the corresponding choice probabilities, and then describe how we test two consideration set models against one another.

3.1 Single-Stage Decision Process

We study a market in which each of R competing firms sells at least one of J substitute products. In every period t, consumer i chooses one product j from J partially differentiated competing products with respective prices $p_{1t},...,p_{Jt}$. The consumer obtains a utility equal to

$$U_{ijt} = \alpha_i p_{jt} + x_{jt}\beta + \varepsilon_{ijt}, \quad i = 0,\ldots,I, \quad t = 1,\ldots,T, \tag{1}$$

where x_{jt} is a K-dimensional vector of observed product characteristics and ε_{ijt} is a zero-mean, i.i.d. extreme-value I distributed individual-specific random shock.[5] The coefficient α_i is consumer i's marginal disutility of price and β is a K-dimensional vector of marginal utilities with respect to the K observed product characteristics. Consumers can choose not to buy any of the J products. Since the mean utility from the outside good is not identified, we normalize it to zero. The utility from this outside option is then

$$U_{i0t} = \alpha_i p_{0t} + x_{0t}\beta + \varepsilon_{i0t}. \tag{2}$$

In our estimation, we take household heterogeneity into account in multiple ways. Firstly, we let store choice depend on household travel distance. More specifically, travel distance enters the product characteristics x_{jt} because it varies across choice options, depending on which of the R sellers offers it. There is rich evidence from the marketing literature that the probability of choosing a store is inversely related to distance from the consumer's home. In fact, travel distance has been found to be a major driver of store choice (e.g.

---

[5] This is a distributional assumption that since McFadden (1978) has become extremely popular in demand estimation because it provides closed-form solutions of the probabilities.

Arnold, Oum and Tigert, 1983; Smith, 2004; Briesch, Chintagunta and Fox, 2009). In contrast, the literature finds that other household characteristics like household income or household size do not significantly affect store choice (Leszczyc, Sinha and Timmermans, 2000; Cleeren et al., 2010).

Secondly, we allow for heterogeneity in price sensitivity. Price sensitivity is modeled to contain a mean coefficient and a varying component which depends on observed household characteristics. The individual price coefficient α_i is distributed with density $f(\alpha|\phi)$, where ϕ collectively refers to the parameters of this distribution. We assume that

$$\alpha_i = \alpha_0 + d_i\xi + \sigma^\alpha\nu, \quad \nu \sim N(0,1), \tag{3}$$

where α_0 denotes the mean price response across all consumers, σ^α is the parameter of the random consumer-specific taste variation ν, and d_i is a vector of household characteristics. In particular, we allow d_i to include household income because both economic theory and the empirical literature suggest that it is a major – if not the most important – determinant of price sensitivity (e.g. Berry, Levinsohn and Pakes, 1995; Hoch et al., 1995; Nevo, 2001; Wakefield and Inman, 2003). ξ captures how strongly observed household characteristics enter price sensitivity.

From the logit structure it follows that the probability L_{ijt} of consumer i choosing product j at time t conditional on the consumer-specific taste variation ν is

$$L_{ijt}(\nu) = \frac{\exp(V_{ijt}(\nu))}{1 + \sum_{k=1}^{J} \exp(V_{ikt}(\nu))}, \tag{4}$$

where $V_{ijt} = \alpha_i p_{jt} + x_{jt}\beta$.

When we have panel data, we observe a sequence of household decisions. The probability of a consumer making this sequence of decisions is the product of the probabilities across the T periods

$$L_i(\nu) = \prod_{t=1}^{T} \frac{\exp(V_{ijt}^{chosen}(\nu))}{1 + \sum_{k=1}^{J} \exp(V_{ikt}(\nu))}, \tag{5}$$

where V_{ijt}^{chosen} denotes the indirect utility from the alternative that was chosen by individual i in period t. The unconditional probability of observing the sequence of T choices corresponds to the integral over all possible values of ν:

$$P_i = \int \left(\prod_{t=1}^{T} \frac{\exp(V_{ijt}^{chosen}(\nu))}{1 + \sum_{k=1}^{J} \exp(V_{ikt}(\nu))} \right) \phi(\nu) d\nu. \tag{6}$$

Lastly, we maximize the log-likelihood $\sum_{i=1}^{N} \ln(P_i)$ with respect to the coefficients α_0, β, and σ^α.

3.2 Two-Stage Decision Process

The two-stage model is very similar to the single-stage model. However, each consumer i makes two consecutive choices per period. In the first stage, she chooses a time- and individual-specific consideration set θ from Θ possible consideration sets. Each consideration set θ contains a different subset of the J products in the market. Not all combinations of products have to be available. The utility from choosing consideration set θ is given by

$$U_{i\theta t} = X_{\theta t}\gamma + \eta_{i\theta t}, \quad i = 1, \ldots, I, \quad t = 1, \ldots, T, \quad \theta = 1, \ldots, \Theta, \tag{7}$$

9

where $X_{\theta t}$ is a vector of observed consideration set characteristics and $\eta_{i\theta t}$ is a vector of i.i.d. extreme-value I distributed shocks. We include travel distance in $X_{\theta t}$ because it varies across stores.

In the second stage, the consumer chooses a product j from her consideration set θ. The corresponding utility is

$$U_{ijt} = \alpha_i p_{jt} + x_{jt}\beta + \varepsilon_{ijt}, \quad j \in \theta, \quad i = 0,\ldots,I, \quad t = 1,\ldots,T, \tag{8}$$

where x_{jt} is a vector of K observed product characteristics and p_{jt} denotes the price of product j at time t. ε_{ijt} is a zero-mean, i.i.d. extreme-value I distributed individual-specific random shock. α_i is again distributed as specified in Equation 3.

Let L_{ijt} be the probability of consumer i choosing product j conditional on the random consumer-specific taste variation ν. Using Bayes' rule, L_{ijt} can be computed as $\sum_\theta L_{ijt|\theta} L_{i\theta t}$, where $L_{ijt|\theta}$ denotes the probability of choosing product j conditional on having consideration set θ and $L_{i\theta t}$ denotes the probability of choosing consideration set θ. The two probabilities are given by

$$L_{i\theta t} = \frac{\exp(V_{i\theta t})}{1 + \sum_{l=1}^{\Theta} \exp(V_{ilt})} \tag{9}$$

and

$$L_{ijt|\theta}(\nu) = \frac{\exp(V_{ijt}(\nu))}{1 + \sum_{k=1}^{J_\theta} \exp(V_{ikt}(\nu))}, \tag{10}$$

where J_θ is the set of products included in θ. $L_{i\theta t}$ is equal to one under the standard assumption of consumers choosing from all products in the market. $L_{ijt|\theta}$ is zero if product j is not included in consideration set θ.

The unconditional probability of a consumer making the sequence of observed choices of considerations sets and products is then

$$P_i = \int \left(\prod_{t=1}^{T} \frac{\exp(V_{i\theta t})}{1 + \sum_{l=1}^{\Theta} \exp(V_{ilt})} \cdot \frac{\exp(V_{ijt}(\nu))}{1 + \sum_{k=1}^{J_\theta} \exp(V_{ikt}(\nu))} \right) \phi(\nu) d\nu. \quad (11)$$

Again, we maximize the log-likelihood $\sum_{i=1}^{N} \ln(P_i)$ with respect to the coefficients $\alpha_0, \sigma^\alpha, \beta$ and γ.

3.3 Testing

In this section, we describe how we test the two consideration set models against each other. Our test follows the idea of a so-called menu approach which is used to estimate the typically unobserved competitive conduct in an industry. The idea is to compare the equilibrium outcome in an industry to the theoretical predictions of a finite set (a "menu") of different models of competition, and then use a model selection test to identify the model which provides the best match with the observed market outcomes. This approach has been used for example to test for collusive vs. competitive behavior (Bresnahan, 1987; Gasmi, Laffont and Vuong, 1992; Nevo, 2001), or for Stackelberg vs. Cournot competition (Kadiyali, Vilcassim and Chintagunta, 1996).

In particular, our testing approach is closely related to Villas-Boas (2007) who tests for different models of vertical relationships between grocery retailers and yogurt manufacturers. For each vertical model, she uses consumer

11

demand estimates to retrieve the corresponding set of marginal costs. She then regresses each set of implied marginal costs on input prices collected from supplemental data, and uses a non-nested selection test to identify the model with the best fit.

For any two competing consideration models, we estimate consumer demand. We then assume a model of seller price-setting behavior and recover a set of marginal costs implied by each demand model's estimates. Finally, we regress each set of marginal costs on marginal cost-shifters and use a non-nested model selection test à la Vuong (1989) to test the null hypothesis that two models perform equally well. The identification of the best model comes from the overidentifying restriction that the marginal costs have to be well-explained by supplemental data on input prices.

The main caveat of our testing procedure is that we have to make assumptions on the model of competition. This is not a problem in markets in which competitive conduct is well known, for example from previous research, reports by competition authorities, etc. However, when we know little about how firms compete in a market, we can only jointly identify the model of competition and consideration: If we have a set of candidate models of competition $A = \{1, 2, 3, ..., n_A\}$ and a set of candidate models of consideration $B = \{1, 2, 3, ..., n_B\}$, we have to test $n_A \cdot n_B$ model combinations to identify the best-fitting pair (a^{best}, b^{best}), where $a^{best} \in A$ and $b^{best} \in B$.

We now formally describe the testing procedure and detail how we recover marginal cost estimates. Each consideration model $z \in \{1, ..., Z\}$ returns a different $J \times 1$ vector of marginal cost estimates c^z. To compute these marginal costs, we need to assume a model of seller competition. In our application, we assume Bertrand-Nash competition (for a discussion of this assumption see

12

Section 5.1). In the following, we set up and solve the maximization problem of the seller. We omit time subscripts t and model subscripts z because the problem is invariant across time and model.

Each seller r sets prices for all products in her assortment S_r which is a non-empty set of products. The seller obtains profits

$$\Pi_r(p) = \sum_{j \in S_r} (p_j - c_j) s_j(p), \tag{12}$$

where c_j is the marginal cost of selling product j, p is a vector of prices $(p_1, p_2, ..., p_J)$, and s_j is the market share of product j. The seller sets her prices such that she maximizes profit Π_r. The corresponding first-order condition is

$$s_j + \sum_{m \in S_r} (p_m - c_m) \frac{\partial s_m}{\partial p_j} = 0. \tag{13}$$

For notational simplicity, we switch to matrix notation in the following. Let T denote the $J \times J$ seller ownership matrix where element $T(j, k)$ is equal to 1 if products j and k are sold by the same firm and 0 otherwise. Let Δ be a $J \times J$-matrix of first derivatives of all market shares with respect to all prices, i.e. element $\Delta(j, k)$ is defined as $\partial s_k / \partial p_j$. Stacking up the first-order conditions for all products and rearranging terms, we obtain the $J \times 1$-vector of marginal costs

$$c = p + (T * \Delta) s(p), \tag{14}$$

where c is a $j \times 1$-vector of marginal costs, p is a $J \times 1$-vector of prices, $s(p)$ is a $J \times 1$-vector of market shares, and $*$ denotes element-wise matrix

multiplication. The marginal cost is identified from the market shares $s(p)$ and the ownership matrix T which we observe in the data, and by the matrix Δ which we obtain from our demand estimates $(\alpha, \sigma^{\alpha}, \beta, \gamma)$. We repeat this procedure to recover c^z for each model $z = 1, ..., Z$.

The most accurate consideration set model z^{best} will return the most accurate set of marginal cost estimates c^{best}. To evaluate the "goodness" of a marginal cost vector and to identify the best model, we use external data on cost-shifters. Specifically, we regress each marginal cost vector c^z on a set of marginal cost-shifters

$$c^z = \xi \delta + \mu, \tag{15}$$

where ξ is a $J \times L$-matrix of cost-shifters, L is the number of different cost-shifters, δ is a $L \times 1$-vector of cost-shifter weights, and μ is a $J \times 1$-vector of mean-zero i.i.d. errors. The regression returns the $L \times 1$-vector of estimated parameters $\hat{\delta}$.

Lastly, we use a model selection test to identify the model with the best fit for the estimation of Equation (15). Specifically, we use the closeness test proposed by Vuong (1989). This test does not require any of the competing models to be correctly specified. Instead, it indicates which model is closest to the true data generation process. The Vuong test states that under the null hypothesis that two non-nested models 1 and 2 fit the true data generation process equally well, the log-likelihood ratio statistic LR asymptotically follows a normal distribution. The Vuong closeness test statistic for two

competing models 1 and 2 is computed as

$$V(1,2) = \frac{LR_N(\hat{\delta}_1, \hat{\delta}_2)}{\sqrt{N}\omega_N} \longrightarrow N(0,1), \tag{16}$$

where

$$LR = L1_N(\hat{\delta}_1) - L2_N(\hat{\delta}_2) - \frac{K1 - K2}{2} \cdot log(N). \tag{17}$$

In equation (16), ω_N denotes the variance of LR and N denotes the sample size. $L1_N$ and $L2_N$ denote the likelihoods of the two models, and $K1$ and $K2$ are the numbers of estimated coefficients in model 1 and 2, respectively. In the final step, we compare the sample value of $V(1,2)$ with critical values of the standard normal distribution.

4 Data

We use German household scanner panel data provided by the market research company GfK. Our data cover all milk and coffee purchases of 1,251 German households in 2010. All households in our sample live in North Rhine-Westphalia, the most populous state of Germany. Each observation in our sample corresponds to one purchase of milk or coffee. We observe the date of the purchase, the retail chain, the paid price, the brand, the characteristics of the product and the sociodemographic characteristics of the household. In total, we observe 31,387 milk purchases and 4,240 coffee purchases.

The German supermarket landscape is characterized by a highly concentrated market structure. In the following, we focus on the seven largest chains which together capture almost 90% of the market (Bundeskartellamt, 2013). Four of the seven biggest chains are full-line retailers, the rest are discounters.

Discounters are popular, with a smaller store size and a narrower assortment, typically with bare-bones store designs and a large share of private labels. Table 1 shows that the discounters in our sample have systematically smaller category assortments: On average, a discounter (full-line retailer) carries 11 (50) milk and 22 (55) coffee varieties (column 4). In general, discounters carry more private labels than national brands.

<div align="center">

Table 1: Summary Statistics: Retail Chains

</div>

Retailer	Format	Market share	#Products	#National Brands / #Private Label
		MILK		
Retail Chain 1	Discount	32.1%	6	0
Retail Chain 2	Full-Line	9.0%	59	4.36
Retail Chain 3	Full-Line	6.5%	38	3.75
Retail Chain 4	Discount	17.4%	8	0.6
Retail Chain 5	Discount	6,1%	19	0.27
Retail Chain 6	Full-Line	9.4%	50	1.94
Retail Chain 7	Full-Line	19.5%	51	2.4
		COFFEE		
Retail Chain 1	Discount	39.0%	16	0
Retail Chain 2	Full-Line	4.9%	57	not defined
Retail Chain 3	Full-Line	3.6%	40	not defined
Retail Chain 4	Discount	25.6%	27	1.75
Retail Chain 5	Discount	3.0%	24	2
Retail Chain 6	Full-Line	7.3%	54	not defined
Retail Chain 7	Full-Line	17.0%	67	13

Source: GfK

The five columns show retail chains, their formats, market shares, assortment size, and brand penetration, i.e. the number of national brands over the number of private labels. For four retail chains, brand penetration is not defined because the store does not carry private labels. For confidentiality reasons we cannot disclose the identity of the chains.

We define a product as a unique combination of characteristics. In the milk market, we define a product as a combination of retail chain, brand, a private

label dummy, fat content, a UHT dummy, and an organic dummy. In the coffee market, we define it as a combination of retail chain, brand, a private label dummy, an organic dummy, a fair trade dummy, and a dummy for decaffeinated coffee. Table 2 shows descriptive statistics for the 50 best-selling products in the milk category and the 30 best-selling products in the coffee category.

Milk (see Panel I in Table 2) is typically sold in cardboard cartons of 1 liter and is almost always pasteurized, i.e. subjected to heating for a short time in order to increase its shelf-life. Different pasteurization procedures yield either fresh milk (with a market share of 64.8%) or ultra-high temperature (UHT) processed milk, the two of which differ in shelf life and taste.[6] Milk usually comes in two different fat levels: 1.5% (semi-skimmed) and 3.5% (full-fat), with roughly equal market shares. Organic milk is a niche market and makes up less than 3% of the total sales. The milk market is largely dominated by private label products: 95% of all milk is sold under a private label, national brands capture only a small share of the market. In particular, discounters sell none or very few national brands (see column 1 in Table 1). Promotions are rarely offered for milk: Only about 1% of all milk sales have promotional prices. In general, milk is a relatively cheap product with an average price of 53.6 euro cents/liter.

Panel II of Table 2 displays the summary statistics for the coffee market. Ground coffee is typically sold in vacuum-sealed packs of 500 grams. It is a storable[7] good with a relatively strong presence of national brands: About 45% of all sold products are branded, and three out of four full-line retailers

---

[6] Heating milk for about 15 seconds up to 75°C produces what is termed regular fresh milk. Heating milk for 1-4 seconds up to 135-150°C yields so-called UHT milk.

[7] We are aware that consumers may stock products and that there could be an upward bias in the price coefficient. A dynamic stockpiling model (see for example Erdem, Imai

17

do not carry any private label coffee (see column 5 in Table 1). Promotions are frequent and popular in the coffee category. In more than 30% of all cases, consumers purchased coffee that was on sale. Mild coffee varieties, i.e. varieties with a lighter roast, have a market share of 30.8%. Decaffeinated (5.7%), organic (1.9%), and fair trade (1.5%) varieties have small market shares.

For the households in the panel we observe two key characteristics. Firstly, we observe the ZIP code of the household's home. Since Germany is divided into 28,683 post code areas, five-digit ZIP codes are a relatively precise measure of location. Secondly, we observe net monthly household income in brackets. We divide households into groups of low income (less than 1,751 euros per month), medium income (more than 1,750 euros but less than 2,751 euros per month), and high income (more than 2,750 euros per month) such that the groups are of roughly the same size (see Table A.1 in Appendix).

We construct a variable to capture how accessible a retail chain is to a household. To do so, we collect all supermarket locations from the German Yellow Pages (2010 edition). We then compute chain r's accessibility to household i as the number of r's outlets divided by the total number of retail outlets in a 10 km radius around household i's home.[8] Values of accessibility

---

and Keane, 2003; Hendel and Nevo, 2006; Lu, 2017) is currently beyond the scope of this paper.

[8] This radius is an approximation of how far consumers are willing to travel to do their shopping. We are aware that it neglects cases in which households do their shopping far from home, for example during travel or next to their work place. However, those cases are difficult to observe because linked store-consumer data is often not available. Consequently, the radius assumption has become widely popular and is used both in research (e.g. Villas-Boas, 2007) and by antitrust authorities like the FTC (Ellickson, Grieco and Khvastunov, 2016) and the German Cartel Office, e.g. `www.bundeskartellamt.de/SharedDocs/Entscheidung/DE/Entscheidungen/Fusionskontrolle/2010/B2-52-10.pdf` or `www.bundeskartellamt.de/SharedDocs/Entscheidung/DE/Entscheidungen/Fusionskontrolle/2015/B2-96-14.pdf` (both last accessed on 21 March 2017).

Table 2: Summary Statistics: Product Characteristics and Cost-Shifters

Variable	Mean	Std. Dev.	Min.	Max.
I. MILK				
Price (euro cents)	53.6	8.590	25	109
Private Label (0=brand, 1=private label)	0.950	0.217	0	1
Organic (0=conventional, 1=organic)	0.028	0.165	0	1
Fresh (0=UHT, 1=fresh)	0.648	0.478	0	1
Fat (%)	2.315	1.021	0.1	3.8
Promotion (1= yes, 0=no)	.010	.098	0	1
Retailer 1 (Discounter)	.349	.477	0	1
Retailer 2 (Full-Line)	.076	.265	0	1
Retailer 3 (Full-Line)	.061	.241	0	1
Retailer 4 (Discounter)	.189	.392	0	1
Retailer 5 (Discounter)	.061	.240	0	1
Retailer 6 (Discounter)	.080	.272	0	1
Retailer 7 (Full-Line)	.182	.386	0	1
Local Market Share	.160	.084	0	.40625
Number of Observations	31,387			
II. COFFEE				
Price (euro cents)	346.92	102.312	119	999
Private Label (1=PL, 0=NB)	.552	.497	0	1
Fair (1=fair, 0=conventional)	0.015	0.123	0	1
Organic (1=organic, 0=conventional)	0.019	0.135	0	1
Decaf (1=decaffeinated, 0=caffeinated)	.057	.232	0	1
Mild (1=mild, 0=not mild)	.308	.462	0	1
Promotion (1= yes, 0=no)	.309	.462	0	1
Retailer 1 (Discounter)	.369	.482	0	1
Retailer 2 (Full-Line)	.079	.271	0	1
Retailer 3 (Full-Line)	.048	.214	0	1
Retailer 4 (Discounter)	.220	.414	0	1
Retailer 5 (Discounter)	.050	.218	0	1
Retailer 6 (Discounter)	.070	.256	0	1
Retailer 7 (Full-Line)	.164	.370	0	1
Local Market Share	.385	.092	.15	.65
Number of Observations	4240			
III. MARGINAL COST-SHIFTERS				
German Raw Milk Price (euro cents/liter)	31.004	2.568	27.95	34.65
Arabica Coffee (USD/kg, world market price)	4.320	.687	3.480	5.471
Robusta Coffee (USD/kg, world market price)	1.736	.207	1.483	2.074
Paper (index)	102.655	3.201	98.2	107
Diesel (index)	100.043	3.307	92.7	106.6
Electricity (index)	99.998	0.689	98.7	100.8
Labor Costs (index)	102.173	5.01	94.085	112.493

vary across households from 0, i.e. a chain not being in a household's shopping radius at all, to 0.406. No retailer is a local monopolist by being the only one to have outlets in the shopping radius of some households.

Finally, we add industry-wide data on marginal cost-shifters. They are provided by the German Federal Statistical Office. We use monthly price indices for the inputs raw milk, coffee beans of the two most popular species (*Coffea arabica* and *Coffea robusta*), paper, diesel, electricity, and labor (see Panel III of Table 2). Marginal cost data of a higher (lower) collection frequency result in a more (less) powerful model selection test.

5 Estimation

5.1 Identification

In the following, we informally discuss identification. The main contribution of this paper is to identify the best-fitting model of consideration among a set of competing models. Each consideration model comes with a different set of estimated marginal costs. Identification of the best model comes from the fact that the marginal costs have to be well-explained by externally collected marginal cost-shifters, i.e. we use the cost-shifter data to construct an overidentifying restriction.

We have to make assumptions on the supply side in order to identify marginal costs. Our first assumption is that retailers compete in Bertrand-Nash fashion.[9] Indeed, German retail chains compete fiercely in prices; the press regularly refers to retail competition as a "price war".[10] Farmers

---

[9] We do not model vertical relationships between retailers and suppliers. This is not a limitation to our estimation because wholesale prices – the result of vertical relationships – will be included in the marginal costs that we back out.

[10] 2010. "Preiskampf der Discounter geht weiter." *Frankfurter Allgemeine Zeitung*, 14 January. `www.faz.net/aktuell/wirtschaft/unternehmen/lebensmittel-einzelh`

frequently protest against downward pressure on wholesale prices[11] and brand manufacturers express concerns that low retail prices may harm brand reputation.[12] Prices are set simultaneously, documented by the fact that price changes typically occur on Mondays. The rare exception are special promotions, e.g. promotions valid only on the weekend.

Our second assumption is that retail chains know true consumer consideration sets. This is supported by large retailer investments in understanding consumer behavior. In Germany for example, more than 2.5 billion euros were spent in 2015 on market research alone.[13] In applications to markets in which retailers are less invested in market research, our assumption can be relaxed. For example, one could model retailers to observe true consideration sets with a measurement error.

Taste parameters are identified by variation in product characteristics (see Table 2). Store-fixed effects explain why consumers may choose a store which offers products at worse conditions than its competitor. The error term is individual-, time- and alternative-specific. It rationalizes why, on two different shopping trips, a consumer may choose differently even when all conditions remain exactly the same. The error term captures, among others,

---

andel-preiskampf-der-discounter-geht-weiter-1596161.html. Last accessed on 16 March 2017.

[11] 2016. "Preiskampf zwischen Aldi und Lidl bedroht Bauern." *Focus*, 5 May. www.focus.de/finanzen/news/milchpreis-im-freien-fall-billige-milchpr odukte-gefaehrden-existenz-von-bauern_id_5503602.html. Last accessed on 16 March 2017.

[12] 2015. "Unilever kritisiert Aldi, Lidl und Co." *Handelsblatt*, 20 July. www.handelsblatt.com/unternehmen/handel-konsumgueter/preiskampf-im-ein zelhandel-unilever-kritisiert-aldi-lidl-und-co-/12079182.html. Last accessed on 16 March 2017.

[13] Statista: de.statista.com/statistik/daten/studie/161551/umfrage/umsatz-der -marktforschungsinstitute-in-deutschland. Last accessed on 16 March 2017.

the momentary mood of the consumer, advertising exposure, and end-of-aisle displays.[14]

5.2 Estimation Technique

We estimate demand using a simulated maximum likelihood estimator (see appendix 8.4). We draw the price coefficient from a lognormal distribution. We do not specify an outside option; instead, demand is estimated conditional on purchase. We do so because milk and coffee are both important staple goods, and their consumption remains remarkably stable despite price variations (see Figures A.1 and A.2).

For both product markets, we test two models (A) and (B) against one another. Model (A) corresponds to the single-stage model of global consideration sets which is described in Section 3.1. Model (B) corresponds to a two-stage approach in which consumers first choose a supermarket and afterwards select a product from the chosen supermarket; it is described in Section 3.2.

In both models it is implicit that, if consumers consider a supermarket, they are aware of all products sold at that supermarket. This is not a necessary assumption. The model can easily be extended by an additional stage in which consumers choose a within-store consideration set. These consideration sets could be modeled as a function of marketing instruments (Sovinsky, 2008; Draganska and Klapper, 2011) or search costs (Seiler, 2013). Importantly, our

---

[14] We are aware that the error term may be correlated with the price. For example, marketing instruments such as advertising can increase both the price and the demand of a product. We run robustness checks in which we tackle endogeneity using the control function approach proposed by Petrin and Train (2010) (see Appendix 8.2) and find that model selection is not affected by it.

assumption does not affect the mechanism of our test; instead, it is able to test any two models against each other, regardless of their level of complexity.[15]

6 Results

6.1 Results in the Milk Category

Tables 3(a) and 3(b) present the estimation results in the milk category for the single-stage and the two-stage model. As expected, we find a negative price coefficient in both models. However, it is in absolute terms smaller in the single-stage model. The reason lies in the inflexibility of the single-stage model: When consideration sets are assumed to be global and consumer i does not react to a price change in product j, this is rationalized by consumer i having a low price sensitivity. Once we allow for a two-stage consideration process, a non-reaction can also be attributed to consumer i not including product j in her consideration set.

Our estimates show a substantial level of household heterogeneity: The standard deviation of the price coefficient is significantly different from zero, thus indicating that price sensitivity varies significantly across households. We find that, in both models, households with low or medium incomes are more price-sensitive than households with a high income. Also, distance plays a crucial role: Households are more likely to select a chain if they live close to its outlets.

Plugging the demand estimates into Equation (14), we recover marginal costs and retailer margins for both models (see Panel I of Table 4). The estimated median margin is 5.3 euro cents in the single-stage model and 4.2 euro cents in the two-stage model. The single-stage model yields a higher

---

[15] In particular, the two models do not have to be nested in each other.

margin because it estimates a lower price sensitivity than the two-stage model, which in turn implies that retailers are more able to raise prices. Our estimates are of the same order of magnitude as those from industry reports.[16]

Table 3: Estimation Results: Milk

(a) Single-Stage Model			(b) Two-Stage Model		
Mean			*Mean*		
Retailer 1	0.9224***	(0.0207)	Retailer 1	0.4255***	(0.0180)
Retailer 2	-0.0066***	(0.0345)	Retailer 2	-0.7432***	(0.0250)
Retailer 3	-0.2666***	(0.0373)	Retailer 3	-0.4873***	(0.0344)
Retailer 4	1.0070***	(0.0240)	Retailer 4	0.2214***	(0.0200)
Retailer 5	-1.0372***	(0.0321)	Retailer 5	-0.7201***	(0.0299)
Retailer 6	0.2396	(0.0382)	Retailer 6	-0.2537***	(0.0319)
Local Market Share	3.8725***	(0.1385)	Local Market Share	3.7837***	(0.1355)
Freshness	0.4802***	(0.0124)	Freshness	0.4306***	(0.0129)
Fat	2.4263***	(0.0799)	Fat	1.9468***	(0.2300)
Private Label	-3.0270***	(0.1786)	Private Label	-0.5588	(0.5263)
Organic	4.7643***	(0.2944)	Organic	-0.4028	(0.8670)
Price	-1.0303***	(0.0435)	Price	-1.7200***	(0.2531)
× Income (Low)	-0.0056	(0.0044)	× Income (Low)	-0.1376***	(0.0037)
× Income (Medium)	-0.0170***	(0.0040)	× Income (Medium)	-0.0865***	(0.0038)
Standard Deviation			*Standard Deviation*		
Price	3.4283	(2.3954)	Price	3.5415***	(0.0595)
No. of Households	1251		No. of Households	1251	
No. of Choice Occasions	31387		No. of Choice Occasions	31387	

Standard errors are in parentheses. The symbols *, ** and *** denote significance at the 1%, 5%, and 10% level, respectively.

Next, we regress estimated marginal costs on observed cost-shifters, i.e. prices of input factors (raw milk, diesel, electricity, labor, and paper), product characteristics (fat content and dummies for private label, organic, and fresh milk), and retailer- and product-dummies. Table A.3 shows the results. While no cost-shifter has a significant coefficient in the single-stage model, several cost-shifters are significant in the two-stage model, e.g. the retailer dummies and the price indices for paper, diesel, and electricity.

Table 4: Estimation Results: Marginal Costs of Milk and Coffee

Model	Marginal Cost in euro cents	Retailer Margin in euro cents	Price in euro cents
I. MILK			
Single-Stage Model	47.476	5.282	55.968
Two-Stage Model	47.861	4.217	55.968
II. COFFEE			
Single-Stage Model	256.850	45.170	299
Two-Stage Model	267.920	36.670	299

We display the median values instead of the means because of outliers in the marginal cost estimates.

We perform a Vuong model selection test and find that the best-performing model is the model in which consumers consider only the products of the retailer they currently shop at: The two-stage model outperforms the single-stage model at a 1%-significance level (see Appendix 8.6). This is economically relevant: If we fail to allow for consideration sets, we obtain margins that are overestimated by 26.2%.

6.2 Results in the Coffee Category

Tables 5(a) and 5(b) show the demand estimates in the coffee category for the single-stage model and the two-stage model. Like in the results for the milk category – and following the same intuition – the price coefficient in the coffee

category is in absolute terms larger for the two-stage model. Again, we find that households with low and medium incomes tend to be more price-sensitive. In both models, the standard deviation of the price is significantly different from zero, indicating substantial household heterogeneity in price sensitivity.

The average consumer prefers, *ceteris paribus*, regular coffee over decaffeinated coffee. The coefficient for mild roasts is insignificant in the single-stage model but significantly positive in the two-stage model. The coefficient for private labels is significantly negative in the single-stage model but insignificant in the two-stage model. Interestingly, unlike in the case of milk, the local availability of a retailer does not seem to affect coffee choice in either of the two models.

Panel II in Table 4 shows the estimated marginal costs and retailer margins. To offer coffee on its shelves, the average supermarket incurs a total cost of around 2.6 euros per pack. The median retailer margin is 45.2 euro cents in the single-stage model and 36.7 euro cents in the two-stage model. These estimated retailer margins are close to those from industry reports.[17]

We regress both sets of marginal cost estimates on the input prices (for Arabica and Robusta beans, diesel, electricity, labor, and paper), product characteristics (dummies for mild, decaffeinated, and private label coffee), and retailer- and product-dummies (see Table A.4 in Appendix). The two models have very similar R-squared statistics and yield similar coefficients, which already suggests that none of the models significantly outperforms the other. Finally, we compute the Vuong test statistic and find that, indeed, the single-stage model and the two-stage model do not perform significantly

---

[17] 2013. "Brennpunkt Kaffee." *Brand Eins*, 27 May. `www.brandeins.de/fileadmin/red aktion/wissen/presse/2013_05_27_focus.pdf`. Last accessed on 9 March 2017.

differently from each other in the market for ground coffee (see Appendix 8.6).

Table 5: Estimation Results: Coffee

(a) Single-Stage Model			(b) Two-Stage Model		
Mean			*Mean*		
Retailer 1	1.3245***	(0.0865)	Retailer 1	0.9536***	(0.0524)
Retailer 2	0.0842	(0.0759)	Retailer 2	-0.9905***	(0.0744)
Retailer 3	-0.2004**	(0.1167)	Retailer 3	-1.7386***	(0.1140)
Retailer 4	0.2703***	(0.0676)	Retailer 4	0.3574***	(0.0514)
Retailer 5	-0.3676***	(0.1215)	Retailer 5	-1.9184***	(0.1129)
Retailer 6	-0.3479***	(0.0902)	Retailer 6	-0.8502***	(0.0876)
Local Market Share	0.3666	(0.3866)	Local Market Share	0.2229	(0.3666)
Decaffeinated	-0.5451***	(0.0689)	Decaffeinated	-0.3585***	(0.0728)
Mild Roast	0.0107	(0.0374)	Mild Roast	0.0885**	(0.0391)
Private Label	-0.1248**	(0.0675)	Private Label	0.0696	(0.0685)
Price	-2.1214***	(0.1139)	Price	-3.1748***	(0.1528)
× Low Income	-0.0078***	(0.0009)	× Low Income	-0.0064	(0.0018)
× Medium Income	-0.0001	(0.0009)	× Medium Income	-0.0137	(0.0022)
Standard Deviation			*Standard Deviation*		
Price	2.2507***	0.1202	Price	2.2701***	(0.1167)
No. of Households	318		No. of Households	318	
No. of Choice Occasions	4240		No. of Choice Occasions	4240	

Standard errors are in parentheses. The symbols *, ** and *** denote significance at the 1%, 5%, and 10% level, respectively.

6.3 Discussion

In this section we discuss the discrepancy between the findings for coffee and the findings for milk: While the two-stage model outperforms the single-stage model in the milk category, it does not perform significantly better in the coffee category. This is because the two-stage model imposes a timing structure: Consumers first choose a store and then a product. This structure implicitly assumes that store choice is not affected by product choice. In the following we explain why this assumption is likely to be violated in the coffee market but appropriate in the milk market.

Firstly, coffee is subject to frequent price promotions which tend to be heavily advertised. At the same time, it is a relatively expensive grocery item, i.e. taking advantage of sales promotions can yield large absolute savings. As a result, consumers have an incentive to collect information about which supermarkets offer coffee promotions, and they may want to select supermarkets depending on their promotional coffee prices. Press reports support the notion that German consumers have a strong preference for bargain-hunting.[18] Secondly, coffee is a product category that is heavily differentiated, both vertically and horizontally. Previous literature has found that coffee is linked to strong brand loyalty on the part of consumers (Krishnamurthi and Raj, 1988). This suggests that consumers may choose stores based on whether they carry the preferred coffee brands and varieties.

Unlike coffee, milk is a relatively cheap product with barely any price promotions; in our sample, only 1% of all milk is sold under a promotion.

---

[18] Heidtmann, Jan. 2016. "Den Deutschen können Lebensmittel nicht billig genug sein." *Süddeutsche Zeitung*, 30 May. www.sueddeutsche.de/wirtschaft/ernaehrung-den-deu tschen-koennen-lebensmittel-nicht-billig-genug-sein-1.3012509. Last accessed on 20 March 2017.

Advertising is similarly rare, mostly because of the dominance of private labels. National brands have a weak position in the milk market, as only 5% of all sales are branded products. In fact, the taste of milk sold by different manufacturers is virtually indistinguishable.[19] All of this suggests that the decision to buy milk is unlikely to affect supermarket choice, i.e. consumer behavior in this market is consistent with the two-stage model.

In general, consideration set formation is driven by both demand and supply conditions. On the demand side, consumer tastes determine whether they care enough about a product to make supermarket choice conditional on it. Consumers care about some product categories more than about others. This fact is closely related to the marketing concept of hedonic and utilitarian products: While hedonic products provide emotional responses like excitement and pleasure, utilitarian goods are primarily functional (Dhar and Wertenbroch, 2000). In the food context, coffee, wine or cheese tend to be hedonic goods whereas milk, flour, and salt fall into the utilitarian category. On the supply side, the price level as well as the frequency and advertising of promotions determine whether price-sensitive consumers select stores in order to take advantage of a promotion in the relevant category.[20]

7 Conclusion

Understanding consideration sets is important. For policy-makers, it is valuable in many applications; for example, antitrust authorities may reach very different conclusions about welfare implications depending on their

---

[19] For example, Joubert and Poalses (2012) find that perceived taste differences between milk brands can be explained by brand reputation and disappear in blind tests.

[20] To be precise, the supply conditions are themselves a result of the demand conditions: Firms optimize prices, promotions and advertising conditional on primitives of consumer demand, such as category-specific sensitivity to advertising and promotions.

assumptions about consideration sets (Sovinsky, 2008; Conlon and Mortimer, 2013). Also, a wide range of policies directly targets consideration sets and consequently affects consumer welfare, e.g. advertising bans (Honka, Hortaçsu and Vitorino, 2017), regulation of choice in public services like health (Gaynor, Propper and Seiler, 2016), or improved education of doctors and its effects on which treatment options they discuss with patients (Fiebig et al., 2015).

In order to infer typically unobservable consideration sets, we construct a test that can compare any two models of consideration and identify which model fits the data better. We use external data on marginal cost-shifters to construct overidentifying restrictions. Our approach has limited data requirements: Next to increasingly accessible household-level purchase data we require only widely available cost-shifter data. We illustrate our approach with an application to supermarket shopping and test two models against each other: a single-stage model with global consideration sets, and a two-stage model in which consumers first choose a store and then a product.

Our results suggest that the single-stage model performs better in hedonic product categories with strong product differentiation, high levels of brand loyalty, and frequent and well advertised promotions. On the other hand, the two-stage model tends to perform better in functional product categories with little product differentiation and few promotions. Importantly, our findings show that there is no "one-size-fits-all" model of consideration. Instead, researchers need to carefully tailor their demand models to the product markets they study.

References

Allenby, Greg M., and James L. Ginter. 1995. "The Effects of In-Store Displays and Feature Advertising on Consideration Sets." *International Journal of Research in Marketing*, 12(1): 67–80.

Arnold, Stephen J., Tae H. Oum, and Douglas J. Tigert. 1983. "Determinant Attributes in Retail Patronage: Seasonal, Temporal, Regional, and International Comparisons." *Journal of Marketing Research*, 20(2): 149–157.

Barroso, Alicia, and Gerard Llobet. 2012. "Advertising and Consumer Awareness of New, Differentiated Products." *Journal of Marketing Research*, 49(6): 773–792.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841–890.

Berry, Steven T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics*, 25(2): 242–262.

Bettman, James R. 1979. *Information Processing Theory of Consumer Choice.* Reading, MA: Addison-Wesley Educational Publishers Inc.

Bonnet, Céline, and Zohra Bouamra-Mechemache. 2015. "Organic Label, Bargaining Power, and Profit-sharing in the French Fluid Milk Market." *American Journal of Agricultural Economics*, 98(1): 113–133.

Bresnahan, Timothy F. 1987. "Competition and Collusion in the American Automobile Industry: the 1955 Price War." *Journal of Industrial Economics*, 35(4): 457–482.

Briesch, Richard A., Pradeep K. Chintagunta, and Edward J. Fox. 2009. "How Does Assortment Affect Grocery Store Choice?" *Journal of Marketing Research*, 46(2): 176–189.

Bronnenberg, Bart J., and Wilfried R. Vanhonacker. 1996. "Limited Choice Sets, Local Price Response and Implied Measures of Price Competition." *Journal of Marketing Research*, 33(2): 163–173.

Bruno, Hernán A., and Naufel J. Vilcassim. 2008. "Research Note-Structural Demand Estimation with Varying Product Availability." *Marketing Science*, 27(6): 1126–1131.

Bundeskartellamt. 2013. "Sector Inquiry Into the Food Retail Sector."

Chiang, Jeongwen, Siddhartha Chib, and Chakravarthi Narasimhan. 1998. "Markov Chain Monte Carlo and Models of Consideration Set and Parameter Heterogeneity." *Journal of Econometrics*, 89(1): 223–248.

Cleeren, Kathleen, Frank Verboven, Marnik G. Dekimpe, and Katrijn Gielens. 2010. "Intra- and Interformat Competition Among Discounters and Supermarkets." *Marketing Science*, 29(3): 456–473.

Conlon, Christopher T., and Julie Holland Mortimer. 2013. "Demand Estimation under Incomplete Product Availability." *American Economic Journal: Microeconomics*, 5(4): 1–30.

Crawford, Gregory S., Rachel Griffith, and Alessandro Iaria. 2017. "Estimating Demand Parameters with Choice Set Misspecification." *Mimeo.*

De los Santos, Babur, Ali Hortaçsu, and Matthijs R. Wildenbeest. 2012. "Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior." *American Economic Review*, 102(6): 2955–2980.

Dhar, Ravi, and Klaus Wertenbroch. 2000. "Consumer Choice Between Hedonic and Utilitarian Goods." *Journal of Marketing Research*, 37(1): 60–71.

Draganska, Michaela, and Daniel Klapper. 2011. "Choice Set Heterogeneity and the Role of Advertising: an Analysis with Micro and Macro Data." *Journal of Marketing Research*, 48(4): 653–669.

Draganska, Michaela, Daniel Klapper, and Sofia B. Villas-Boas. 2010. "A Larger Slice or a Larger Pie? an Empirical Investigation of Bargaining Power in the Distribution Channel." *Marketing Science*, 29(1): 57–74.

Ellickson, Paul B., Paul L.E. Grieco, and Oleksii Khvastunov. 2016. "Measuring Competition in Spatial Retail." Working Paper.

Erdem, Tülin, Susumu Imai, and Michael P. Keane. 2003. "Brand and Quantity Choice Dynamics Under Price Uncertainty." *Quantitative Marketing and Economics*, 1(1): 5–64.

Fiebig, Denzil G., Rosalie Viney, Stephanie Knox, Marion Haas, Deborah J. Street, Arne R. Hole, Edith Weisberg, and Deborah Bateson. 2015. "Consideration Sets and Their Role in Modelling Doctor Recommendations About Contraceptives." *Health Economics*, 26: 54–73.

Gasmi, Farid, Jean Jacques Laffont, and Quang Vuong. 1992. "Econometric Analysis of Collusive Behavior in a Soft-Drink Market." *Journal of Economics & Management Strategy*, 1(2): 277–311.

Gaynor, Martin, Carol Propper, and Stephan Seiler. 2016. "Free to Choose? Reform, Choice, and Consideration Sets in the English National Health Service." *American Economic Review*, 106(11): 3521–3557.

Guadagni, Peter M., and John D. C. Little. 1983. "A Logit Model of Brand Choice Calibrated on Scanner Data." *Marketing Science*, 2(3): 203–238.

Hauser, John R., and Birger Wernerfelt. 1990. "An Evaluation Cost Model of Consideration Sets." *Journal of Consumer Research*, 393–408.

Hendel, Igal, and Aviv Nevo. 2006. "Measuring the Implications of Sales and Consumer Inventory Behavior." *Econometrica*, 74(6): 1637–1673.

Hoch, Stephen J., Byung-Do Kim, Alan L. Montgomery, and Peter E. Rossi. 1995. "Determinants of Store-Level Price Elasticity." *Journal of Marketing Research*, 32(1): 17–29.

Holbrook, Morris B., and Elizabeth C. Hirschman. 1982. "The Experiential Aspects of Consumption: Consumer Fantasies, Feelings, and Fun." *Journal of Consumer Research*, 9(2): 132–140.

Honka, Elisabeth. 2014. "Quantifying Search and Switching Costs in the US Auto Insurance Industry." *RAND Journal of Economics*, 45(4): 847–884.

Honka, Elisabeth, Ali Hortaçsu, and Maria Ana Vitorino. 2017. "Advertising, Consumer Awareness, and Choice: Evidence from the US Banking Industry." *RAND Journal of Economics*, forthcoming.

Howard, John A., and Jagdish N. Sheth. 1969. *The Theory of Buyer Behavior.* New York: Wiley.

Joubert, Johan P.R., and Jacolize Poalses. 2012. "What's in a Name? The Effect of a Brand Name on Consumers' Evaluation of Fresh Milk." *International Journal of Consumer Studies*, 36(4): 425–431.

Kadiyali, Vrinda, Nauel J. Vilcassim, and Pradeep K. Chintagunta. 1996. "Empirical Analysis of Competitive Product Line Pricing Decisions: Lead, Follow, or Move Together?" *Journal of Business*, 69(4): 459–487.

Krishnamurthi, Lakshman, and S.P. Raj. 1988. "A Model of Brand Choice and Purchase Quantity Price Sensitivities." *Marketing Science*, 7(1): 1–20.

Leszczyc, Peter T.L., Ashish Sinha, and Harry J.P. Timmermans. 2000. "Consumer Store Choice Dynamics: an Analysis of the Competitive Market Structure for Grocery Stores." *Journal of Retailing*, 76(3): 323–345.

Lu, Anna. 2017. "Consumer Stockpiling and Sales Promotions." Working Paper.

Lu, Zhentong. 2016. "A Moment Inequality Approach to Estimating Multinomial Choice Models with Unobserved Consideration Sets." Working Paper.

Malhotra, Naresh K., Mark Peterson, and Susan B. Kleiser. 1999. "Marketing Research: a State-Of-The-Art Review and Directions for the Twenty-First Century." *Journal of the Academy of Marketing Science*, 27(2): 160–183.

McFadden, Daniel. 1978. *Modelling the Choice of Residential Location.* Institute of Transportation Studies, University of California.

Mehta, Nitin, Surendra Rajiv, and Kannan Srinivasan. 2003. "Price Uncertainty and Consumer Search: a Structural Model of Consideration Set Formation." *Marketing Science*, 22(1): 58–84.

Mitra, Anusree. 1995. "Advertising and the Stability of Consideration Sets over Multiple Purchase Occasions." *International Journal of Research in Marketing*, 12(1): 81–94.

Nevo, Aviv. 2000. "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand." *Journal of Economics & Management Strategy*, 9(4): 513–548.

Nevo, Aviv. 2001. "Measuring Market Power in the Ready-To-Eat Cereal Industry." *Econometrica*, 69(2): 307–342.

Petrin, Amil, and Kenneth Train. 2010. "A Control Function Approach to Endogeneity in Consumer Choice Models." *Journal of Marketing Research*, 47(1): 3–13.

Roberts, John H., and James M. Lattin. 1997. "Consideration: Review of Research and Prospects for Future Insights." *Journal of Marketing Research*, 34(3): 406–410.

Seiler, Stephan. 2013. "The Impact of Search Costs on Consumer Behavior: a Dynamic Approach." *Quantitative Marketing and Economics*, 11(2): 155–203.

Siddarth, S., Randolph E. Bucklin, and Donald G. Morrison. 1995. "Making the Cut: Modeling and Analyzing Choice Set Restriction in Scanner Panel Data." *Journal of Marketing Research*, 32(3): 255–266.

Smith, Howard. 2004. "Supermarket Choice and Supermarket Competition in Market Equilibrium." *Review of Economic Studies*, 71(1): 235–263.

Sovinsky, Michelle. 2008. "Limited Information and Advertising in the US Personal Computer Industry." *Econometrica*, 76(5): 1017–1074.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation.* Cambridge University Press.

Van Nierop, Erjen, Bart Bronnenberg, Richard Paap, Michel Wedel, and Philip H. Franses. 2010. "Retrieving Unobserved Consideration Sets from Household Panel Data." *Journal of Marketing Research*, 47(1): 63–74.

Villas-Boas, Sofia B. 2007. "Vertical Relationships Between Manufacturers and Retailers: Inference with Limited Data." *Review of Economic Studies*, 74(2): 625–652.

Vuong, Quang H. 1989. "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses." *Econometrica*, 57(2): 307–333.

Wakefield, Kirk L., and J. Jeffrey Inman. 2003. "Situational Price Sensitivity: the Role of Consumption Occasion, Social Context and Income." *Journal of Retailing*, 79(4): 199–212.

8 Appendix

8.1 Household Characteristics

Table A.1: Summary Statistics: Net Monthly Household Income

Income Group	No. Households	%
COFFEE		
1 (income <1750 euros)	1,781	36.08
2 (1750 euros $\leq$ income $\leq$ 2750 euros)	1,871	37.91
3 (2750 euros < income)	1,284	26.01
MILK		
1 (income <1750 euros)	9,186	29.27
2 (1750 euros $\leq$ income $\leq$ 2750 euros)	11,703	37.29
3 (2750 euros < income)	10,498	33.45

8.2 Control Function Approach

In this section, we describe our application of the control function approach proposed by Petrin and Train (2010). The key idea is that if we can derive a proxy variable that captures the part of the price that depends on the error term, then the remaining variation in the price will be independent of the error and thus allow standard estimation. In the first step, we use an ordinary least squares estimator to regress the potentially endogenous price on a number of instruments and exogenous variables:

$$p_{jt} = \delta J_{jt} + \gamma W_{jt} + \eta_{jt}. \tag{18}$$

J_{jt} and W_{jt} are vectors of product characteristics and cost-shifters, respectively. J_{jt} includes the fat content, a private label dummy, a fresh milk dummy, an organic dummy and retailer dummies. W_{jt} includes the price indices for raw milk, diesel, and electricity. η_{jt} is an i.i.d. mean-zero error

term. Table A.2 displays the regression results which are all consistent with economic intuition.

Table A.2: Estimation Results: Control Function

Variable	Mean	Standard Error
German Raw Milk Price Index	0.217***	(0.0101)
Diesel Price Index	0.0177*	(0.00856)
Electricity Price Index	0.528***	(0.00762)
Private Label	-16.16***	(0.110)
Fresh Milk	-0.348***	(0.0452)
Fat Content (in %)	2.933***	(0.0211)
Organic	34.11***	(0.129)
Retailer 2	0.298***	(0.0893)
Retailer 3	0.0934	(0.0963)
Retailer 4	0.536***	(0.0629)
Retailer 5	-0.232**	(0.0773)
Retailer 6	0.00378	(0.0927)
Retailer 7	1.385***	(0.0895)
N	37799	
adj. R^2	0.995	

Standard errors are in parentheses. The symbols *, ** and *** denote significance at the 1%, 5%, and 10% significance level, respectively.

In the second step, we obtain the residual from (18) and plug it into the utility function:

$$U_{ijt} = \alpha_i p_{jt} + x_{jt}\beta_i + \tau\hat{\eta}_{jt} + \bar{\varepsilon}_{ijt}, \tag{19}$$

where ε_{ijt} equals $\bar{\varepsilon}_{ijt} + \tau\hat{\eta}_{jt}$ and is extreme-value I distributed. Equation 19 can now be estimated with standard methods. In our robustness checks we find that the coefficient of the control variable is statistically significant but economically irrelevant and does not affect our model selection results.

8.3 Grid-Search: Filling in Prices

In order to fill in price p_{jtig} of good j on day t paid by household i living in postcode g, we search for households in the same postcode that purchased the product on the same day. If we find such households, we replace p_{jtig} with the average price paid by households in the postcode area. If we do not find any such households, we increase the searched time period to a week and repeat the procedure. We gradually and alternatingly increase both the time period and the geographical area until we find a matching household that purchased the same product.

8.4 Simulated Maximum Likelihood

One complication of the mixed logit model is that there is no analytic solution to the integral in Equations 6 and 11. We approximate both equations via simulation. The simulated probability is:
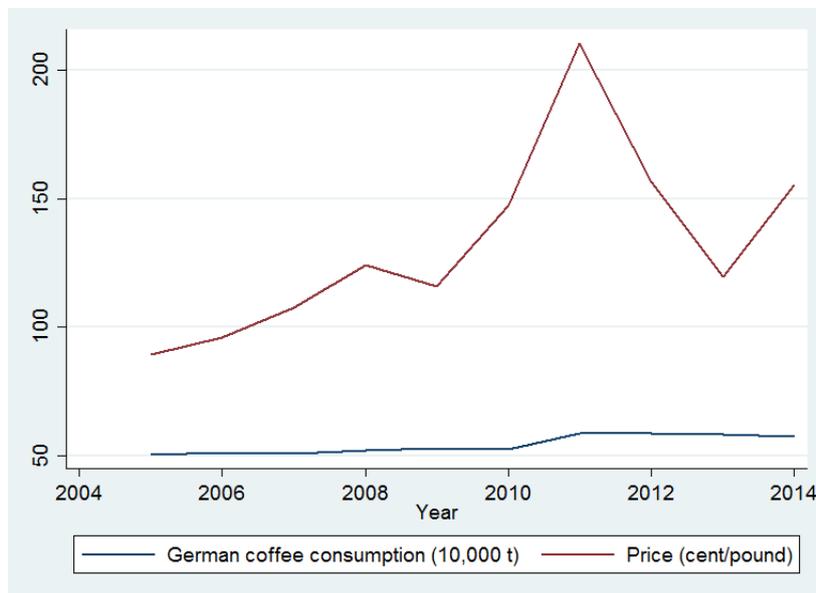
$$SP_i = \sum_{r=1}^{R} L_i(\nu^r), \tag{20}$$

where R is the number of simulations and ν^r is the rth draw from the standard-normal distribution. We use Halton draws for faster convergence. The simulated log-likelihood function is

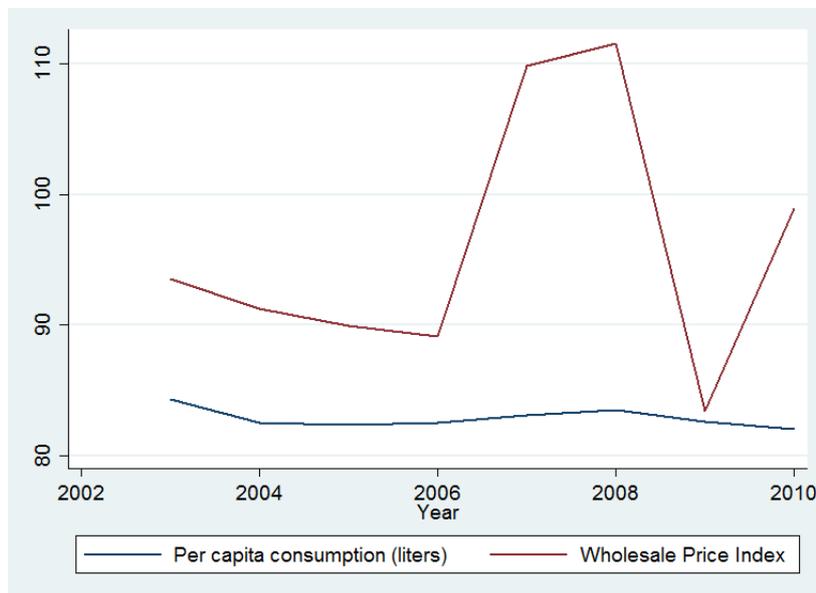$$SLL = \sum_{i=1}^{N} \ln(SP_i). \tag{21}$$

8.5 Coffee and Milk Consumption

Figure A.1: Coffee Prices and Coffee Consumption

The figure shows variation in the coffee world market price and in German coffee consumption. Data Source: International Coffee Organization, German Coffee Association.

Figure A.2: Milk Prices and Milk Consumption

The figure shows variation in the German wholesale price index for milk (base year = 2010) and in the annual German per capita milk consumption. Data Source: German Federal Statistical Office.

8.6 Model Selection Test

Table A.3: Regression Results of Marginal Costs on Cost-Shifters for Milk

Variables	Single-Stage Model	Two-Stage Model
German Raw Milk Price Index	-0.897	-575.8
	(10.12)	(2,175)
Diesel Price Index	-2.716	2,898***
	(3.153)	(677.9)
Electricity Price Index	4.052	-5,691***
	(8.446)	(1,816)
Labor Cost Index	0.937	-775.3
	(2.477)	(532.6)
Paper Price Index	2.146	-8,626***
	(14.78)	(3,177)
Fresh	3.651	6,026
	(31.19)	(6,707)
Fat Content	-1.375	1,251
	(10.20)	(2,193)
Private Label	-7.037	-1,369
	(32.95)	(7,085)
Organic	12.68	-8,802
	(53.43)	(11,489)
Retailer 2	-20.32	46,135***
	(50.54)	(10,868)
Retailer 3	-11.98	46,670***
	(46.47)	(9,992)
Retailer 4	-10.60	42,404***
	(47.48)	(10,208)
Retailer 5	-3.569	49,770***
	(49.29)	(10,598)
Retailer 6	-8.975	51,460***
	(65.54)	(14,092)
Retailer 7	-15.23	44,046***
	(56.98)	(12,252)
Constant	377.5	-69,068
	(996.7)	(214,307)
Observations	600	600
R-squared	0.153	0.167

Product dummies are not displayed due to their large number. Standard errors are in parentheses.
The symbols *, ** and *** denote significance at the 1%, 5%, and 10% level, respectively.

Table A.4: Regression Results of Marginal Costs on Cost-Shifters for Coffee

Variables	Single-Stage Model	Two-Stage Model
Arabica Beans	-13.67	-15.01
	(16.43)	(16.40)
Robusta Beans	57.77	62.44*
	(36.36)	(36.30)
Diesel Price Index	-0.405	-0.239
	(0.957)	(0.956)
Electricity Price Index	-20.19*	-21.40**
	(10.50)	(10.48)
Labor Cost Index	0.474	0.593
	(0.541)	(0.540)
Paper Price Index	5.632	5.887
	(3.718)	(3.711)
Private Label	-54.66***	-57.51***
	(8.990)	(8.974)
Mild	100.8***	104.0***
	(8.990)	(8.974)
Decaffeinated	86.90***	88.06***
	(12.71)	(12.69)
Retailer 2	15.79	13.05
	(12.71)	(12.69)
Retailer 3	-50.18***	-51.23***
	(12.71)	(12.69)
Retailer 4	46.64***	52.43***
	(8.990)	(8.974)
Retailer 5 (omitted)	-	-
	-	-
Retailer 6	-17.60	-22.64*
	(12.71)	(12.69)
Retailer 7	-43.79***	-48.03***
	(12.71)	(12.69)
Constant	1,674**	1,740**
	(807.8)	(806.3)
Observations	360	360
R-squared	0.833	0.836

Product dummies are not displayed due to their large number. Standard errors are in parentheses. The symbols *, ** and *** denote significance at the 1%, 5%, and 10% level, respectively.

Table A.5: Vuong Test Statistic

Model	Vuong Test Statistic	Result
MILK		
$V(B,A)$	2.807	B $\succ$ A
COFFEE		
$V(B,A)$	0.011	B $\approx$ A

(A) Homogeneous consideration sets. (B) Heterogeneous consideration sets. The test is carried out at a 1% significance level, with the corresponding χ^2-distributed comparison value being 2.326.