

**Choice Set Heterogeneity and the Role of Advertising:  
An Analysis with Micro and Macro Data**

**Michaela Draganska**  
Stanford University  
Graduate School of Business  
Stanford, CA 94305-5015  
draganska\_michaela@gsb.stanford.edu

**Daniel Klapper**  
Goethe University  
Lehrstuhl für Marketing  
D-60054 Frankfurt am Main  
dklapper@wiwi.uni-frankfurt.de

March 10, 2010  
track'v9

# **Choice Set Heterogeneity and the Role of Advertising: An Analysis with Micro and Macro Data**

## **Abstract**

We show how to use micro-level survey data from a tracking study on brand awareness in conjunction with data on sales and advertising expenditures to improve the specification, estimation, and interpretation of aggregate demand systems. We depart from the commonly made full information assumption and incorporate limited information in the form of choice sets to capture the fact that consumers may not be aware of all available brands at purchase time. The detailed modeling of the processes by which advertising influences brand choice and ultimately sales improves our understanding of how advertising works and thus enhances our ability to make policy recommendations. In addition, we are able to show that advertising has a direct effect on brand awareness (inclusion in choice set) in addition to its effect on consumer preferences (increase in utility).

Key Words: advertising, choice sets, demand estimation, econometric models.

# 1 Introduction

Over the past decades we have witnessed an unprecedented proliferation of products, ranging from customized credit cards to dozens of flavors of vanilla ice cream. In such crowded markets with many brands and products, it matters as much *which* brands consumers think about when making their purchase decisions as it matters *what* they think about them. Recognizing this, companies advertise to boost sales not only by persuading consumers of the superiority of their products, but also by ensuring that their products are salient in the minds of consumers in the first place. In addition to a substitution effect (if the brand is in the choice set of the consumer, will it be chosen?), advertising therefore has an availability effect (is the brand in the choice set?) that can be just as important.

In this paper we take a closer look at these two roles of advertising by separately examining the availability effect of advertising on brand awareness and the substitution effect on consumer preferences. Understanding whether advertising affects sales through choice sets or utility is important as it has implications for the intensity of competition in the market and its profitability. If advertising increases awareness and thus the number of brands consumers choose from, then increased advertising can decrease profits (for an overview of the literature see Bagwell 2005). From an empirical perspective, if advertising is mistakenly specified as affecting utility rather than choice sets (as it often is), then estimated price elasticities can be severely biased, causing suboptimal managerial decision-making.

Distinguishing econometrically between the ways in which advertising can affect sales is crucial, but it is also very challenging if only aggregate market share data or revealed preference data on brand choices are available (Akerberg 2001, Narayanan & Manchanda 2008). Intuitively, we cannot tell if a consumer did not buy a product because she does not like it or because she is not aware of it. To disentangle the effects of advertising on choice sets and utility, the existing literature has therefore relied heavily on functional form assumptions (see, e.g., Andrews & Srinivasan 1995,

Siddarth, Bucklin & Morrison 1995, Bronnenberg & Vanhonacker 1996, van Nierop, Paap, Bronnenberg, Franses & Wedel forthcoming). To more directly assess the ways in which advertising can affect sales we make use of a new data source. In particular, we bring in micro-level survey data from a tracking study, which provides us with measures of the level of information possessed by individual consumers such as the subset of brands that they recall with or without aid. Although companies frequently collect this type of data to help them understand how brand perceptions evolve over time, its use in academic research has to date been limited.<sup>1</sup> We show how to incorporate information on brand awareness from a tracking study into a standard model of aggregate demand in order to improve its specification, estimation, and interpretation.

Direct measures of brand awareness can improve demand estimation because consumers are typically not aware of – let alone consider – all available brands when making their purchase decisions and because choice sets vary across consumers and over time. Decades of research in marketing have established that consumers employ different decision heuristics to simplify the brand choice process. Empirical evidence suggests that buyers select from only a subset of the available brands and that screening rules may be invoked to simplify choices between as few as four alternatives (Gensch 1987). There is also evidence that even in a store, when faced with the actual choice options, people do not evaluate all alternatives (Hoyer 1984, Mitra & Lynch 1995) and in fact only notice attributes such as prices or promotions after they have decided to consider a brand for purchase. Several studies have established that estimated brand choice parameters and thus inferred price elasticities based on a full choice set can be very different from the ones based on a limited choice set (see e.g., Swait & Ben-Akiva 1987, Bronnenberg & Vanhonacker 1996). In an aggregate demand estimation setting, Sovinsky Goeree (2008) also finds that failing to account

---

<sup>1</sup>Recently Srinivasan, Vanhuele & Pauwels (forthcoming) have fitted VAR time-series models to data from a tracking study in order to analyze the correlation patterns in these “mindset metrics” and the sales response to them.

for choice sets leads to severely biased parameter estimates. The broader point here is that erroneously assuming that a consumer chooses amongst all products in the market when making a purchase decision, leads to a misspecified demand system (see e.g., Swait & Ben-Akiva 1987, Horowitz & Louviere 1995, Bajari & Benkard 2005).

Furthermore, choice sets vary across consumers and over time for a number of reasons. Choice sets may simply reflect exogenous constraints on choice that arise from consumers' socio-economic characteristics. Moreover, as consumers differ in their propensity to search for products and learn about brands, the size and composition of choice sets will be different for different consumers (Shocker, Ben-Akiva, Boccara & Nedungadi 1991). Of particular interest to marketers is that choice sets are not fixed but can be influenced by the marketing mix (Allenby & Ginter 1995, Mitra 1995, Siddarth et al. 1995). This further increases the heterogeneity in choice sets because different consumers respond differently to marketing activities and because these activities vary over time.

Previous attempts to account for the heterogeneity in choice sets in aggregate demand estimation have been limited. Albuquerque & Bronnenberg (2009) combine macro data on sales and marketing mix with micro data on past purchases. Lacking direct measures of brand awareness, they define a consumer's choice set to consist of the brands that the consumer bought in the past. This approach raises an initial conditions problem and requires strong stationarity assumptions which may not always hold (e.g., new product introductions). Closest in spirit to our paper is Sovinsky Goriee (2008) who models the probability of choice set membership as a function of demographics and exposure to advertising by assuming that choice set membership is independent across brands. Moreover, because she does not have data on choice sets, she cannot identify the differential impact of advertising on choice sets and utility but simply assumes that advertising operates entirely through choice sets.

We contribute to the extant literature on demand estimation by showing how to use micro-level measures of brand awareness to account for the fact that consumers

choose among a subset of the available brands when making their purchase decisions and for the heterogeneity of choice sets. Our starting point is a model of aggregate demand that accounts for consumer heterogeneity by modeling their tastes in a random coefficients specification (Berry 1994, Berry, Levinsohn & Pakes 1995). We depart from the full information assumption in these models and incorporate limited information in the form of choice sets. We acknowledge that choice sets vary across consumers and time, thus adding a second, potentially equally important source of heterogeneity.

We develop three ways of using the micro-level survey data to account for the heterogeneity in choice sets. We first treat the choice set of a consumer as a random coefficient and integrate over its distribution in the population of consumers. This approach takes the stated choice sets as data. To the extent that choice sets are latent, a probabilistic model of choice set formation may be more appropriate. Our second way of using the tracking study is therefore to obtain nonparametric estimates for the aggregate probabilities of the various possible choice sets. Finally, we formulate a model which links the choice set of a consumer to her demographics and to the advertising expenditures of the brands in the product category.

We illustrate these three different empirical strategies using data from the ground coffee category in Germany. For frequently purchased consumer packaged goods such as ground coffee, where there is a vast number of brands on supermarket shelves, the inclusion of a brand in the choice set of a consumer is a function of its top-of-mind awareness. Brand awareness itself is driven by advertising and also by the consumer's purchase history (arguably also a function of brand awareness at past points in time). We do not observe what consumers have bought in the past but this is probably not as relevant because ground coffee is a mature product category where we can safely assume that consumers have at some point of time considered most brands. What shifts the salience levels now is mostly advertising. To capture this effect, we follow the literature on consumer behavior and view the size and composition

of choice sets as being driven by noncompensatory processes (Shocker et al. 1991) as opposed to the systematic evaluation of the utility of all brands in the market (Hauser & Wernerfelt 1990, Roberts 1989, Roberts & Lattin 1991). We then relate the probability that the consumer has a certain choice set back to her demographics and the advertising expenditures of the various competing brands.

Our estimates show that ignoring choice set heterogeneity biases the parameter estimates of aggregate demand models in several ways. First, estimated brand constants are typically too low. After accounting for limited information in the form of choice sets, consumers are seen to have stronger brand preferences. This nicely illustrates the fundamental point about identification: Without data on choice sets, we cannot tell if a product is not bought because the consumer does not know about it or because she does not like it.

Second, failing to account for choice set heterogeneity causes price elasticities to be underestimated. This has obvious managerial implications for pricing policy, in particular, prices may be set too high. The reason for the understated price elasticities is quite intuitive. The price response in an aggregate demand model is calibrated to rationalize observed quantity changes by the price changes recorded in the data. With limited information some consumers do not react to a price change at all. Hence, those consumers who are aware of a product must have a stronger reaction to a change in its price to render the movements in quantities and prices in the data consistent with one another.

Third, by combining data on sales and advertising expenditures with tracking data on brand awareness, we are able to distinguish between two basic functions of advertising. Our estimates indicate that, besides contributing to the utility that a consumer derives from a brand, advertising makes it more likely that the brand is in a consumer's choice set at the time of purchase.

Fourth, we investigate if it is really a consumer's most-highly valued brands that make it into her choice set as assumed by compensatory approaches to choice set

formation (Hauser & Wernerfelt 1990, Roberts 1989, Roberts & Lattin 1991). We find that this is not systematically the case. The fact that the choice set has information beyond the utility function, in turn, helps to explain why it is crucial to account for choice set heterogeneity in aggregate demand estimation.

The remainder of this paper is organized as follows. Section 2 presents the various data sources we use in our empirical application and takes a first look at the data. Sections 3 and 4 develop the theoretical model and our strategy for estimating it. Section 5 presents the results and Section 6 concludes.

## 2 Data

For the empirical application we focus on the ground coffee category in Germany. Our data come from three sources: (1) aggregate sales and marketing mix data from MADAKOM, (2) data on advertising expenditures made available to us by an anonymous manufacturer, and (3) micro-level survey data from a tracking study conducted by a leading German market research company on behalf of the same manufacturer.

**Sales and marketing mix.** The data was collected by MADAKOM, Germany, from a national sample of stores belonging to six major retail chains, Edeka, Markant, Metro, Rewe, Spar, and Tengelmann. The data covers more than 50 brands. We focus on five major national brands, Jacobs, Melitta, Dallmayr, Tchibo, and Eduscho, which together comprise about 70% of the market. The next largest brands are Idee, Onko, and Röstfein. We exclude these brands as well as a number of brands with negligible market shares below 0.5% from the analysis.

We have weekly information on the sales, prices, and promotional activities (in-store communication and features) for all brands in all retail chains in the ground coffee category from the first week in 2000 to the last week in 2001. Week 25 in 2000 is missing due to data collection problems so that we have a total of 103 weeks. Because



buying the same brand at a different retail chain may provide for a different purchase experience, the unit of observation in the empirical analysis is a brand in a given retail chain and week. Table 1 gives an overview of the data. As can be seen, Jacobs is the largest brand followed by Melitta. The remaining brands, Dallmayr, Tchibo, and Eduscho, have about half the market share of Jacobs. There is a considerable amount of variation within and across brands in the marketing mix variables.

Table 1: Summary statistics for sales, marketing mix, and advertising expenditures. Averages with standard deviations underneath in parentheses.

	market share	price	communication	feature	advertising
Jacobs	33.63 (12.47)	6.84 (0.51)	0.17 (0.09)	0.36 (0.20)	681.64 (416.06)
Melitta	21.53 (12.59)	6.30 (0.45)	0.18 (0.12)	0.35 (0.19)	806.94 (430.62)
Dallmayr	13.45 (8.76)	7.52 (0.44)	0.14 (0.09)	0.29 (0.20)	725.42 (237.30)
Tchibo	16.69 (4.67)	7.99 (0.43)	0.20 (0.06)	0.02 (0.04)	874.30 (510.59)
Eduscho	14.70 (5.59)	6.79 (0.40)	0.26 (0.07)	0.09 (0.08)	587.49 (444.96)

For the empirical analysis we include an outside good. To calculate its share, we use the total sales within each week in each retail chain. From the *Lebensmittel Zeitung* (2006) we collected data about the average amount spent per shopping trip in each of the six major retail chains in 2000 and 2001. We use this information to estimate retail chain traffic and calculate the size of the potential market.

**Advertising expenditures.** We received monthly brand-level advertising expenditures for all brands in the ground coffee category in Germany from an anonymous manufacturer. Advertising expenditures are available for different media, TV, radio, newspapers, magazines, and billboards. Because in the ground coffee category TV is by far the most important media, we focus on it. We spread the monthly advertising

expenditures uniformly across the weeks of a month. In Table 1 we present summary statistics on weekly advertising expenditures for TV commercials in 1000 DEM.

The full impact of advertising may be realized over time. Hence, we use advertising expenditures to construct a measure of goodwill. Goodwill accumulates over time as a function of firms’ investment in advertising and depreciates in the absence thereof. The goodwill of a brand then affects the utility that a consumer derives from the brand. It may also affect the consumer’s top-of-mind awareness when we model the process of choice set formation. As in Doganoglu & Klapper (2006), we specify a Cobb-Douglas production function for goodwill

$$g_{jt} = g_{jt-1}^\lambda a_{jt}^{(1-\lambda)},$$

where  $g_{jt}$  and  $g_{jt-1}$  are the current and past goodwill of brand  $j$  and  $a_{jt}$  are its advertising expenditures. The parameter  $\lambda$  measures the persistence of goodwill over time.

**Tracking study.** The tracking study extends over a period of 47 months from January 1999 to November 2002. It was conducted by a leading German market research company on behalf of an anonymous manufacturer. Each month approximately 320 consumers are interviewed regarding their awareness of and attitude towards various brands in the market for ground coffee in Germany. The data set consists of repeated cross-sections of consumers with a total of 15254 consumers. The market research company provides consumer-specific weights to ensure that the respondents are a nationally representative sample. Throughout this paper we use these weights when sampling consumers from the tracking study.

The tracking study includes measures of aided and unaided brand recall, brands considered when making a purchase decision (“relevant set”), brands that a consumer would never buy (“reject set”), and brands that a consumer has purchased in the last 12 months. A description of these variables can be found in Table 2.

Given that there are more than 50 brands in the German ground coffee category,

it is implausible to assume that a consumer is aware of – let alone considers – all of them when making her purchase decision. It seems more likely that she chooses between a handful of brands that are on top of her mind. The tracking study reveals that aided brand recall is generally high and does not to change much (see again Table 2), as can perhaps be expected in a mature product category. Similar to previous research (Horowitz & Louviere 1995), we thus focus on top-of-mind awareness or brand salience in our empirical investigation.

Table 2: Summary statistics for tracking study: Awareness and attitude. Averages with standard deviations underneath in parentheses.

	Melitta	Tchibo	Dallmayr	Jacobs	Eduscho
unaided brand recall	0.6489 (0.4773)	0.6697 (0.4704)	0.4218 (0.4939)	0.8390 (0.3675)	0.5897 (0.4919)
aided brand recall	0.9135 (0.2811)	0.9042 (0.2944)	0.9144 (0.2797)	0.9719 (0.1653)	0.8230 (0.3817)
relevant set	0.3535 (0.4781)	0.3545 (0.4784)	0.3546 (0.4784)	0.5353 (0.4988)	0.3039 (0.4599)
reject set	0.3252 (0.4685)	0.2437 (0.4294)	0.2486 (0.4322)	0.4304 (0.4952)	0.2397 (0.4269)
purchase in last 12 months	0.3164 (0.4651)	0.4071 (0.4913)	0.3238 (0.4679)	0.5306 (0.4991)	0.2572 (0.4371)

The tracking study additionally includes a rich set of demographics and measures of usage behavior for the surveyed consumers. Table 3 provides an overview. It shows that more than 40% of respondents are not very interested or may not even like watching commercials on TV. Three quarters of respondents state that they watch every day TV, a number that seems plausible given that the average German citizen watches more than 100 minutes of TV a day. Approximately 20% of respondents are from East Germany which has 17 million out of a total of 82 million citizens. More than half of respondents have no kids in their household. Our respondents have an average age of 40.38. A slight majority of respondents classify themselves as heavy users of ground coffee, which seems reasonable given the fact that more than 94% of

Germans drink coffee. When modeling choice set formation we use this information on demographics and usage behavior to allow consumers to be systematically more or less aware of the different brands of ground coffee.

Table 3: Summary statistics for tracking study: Demographics and usage behavior.

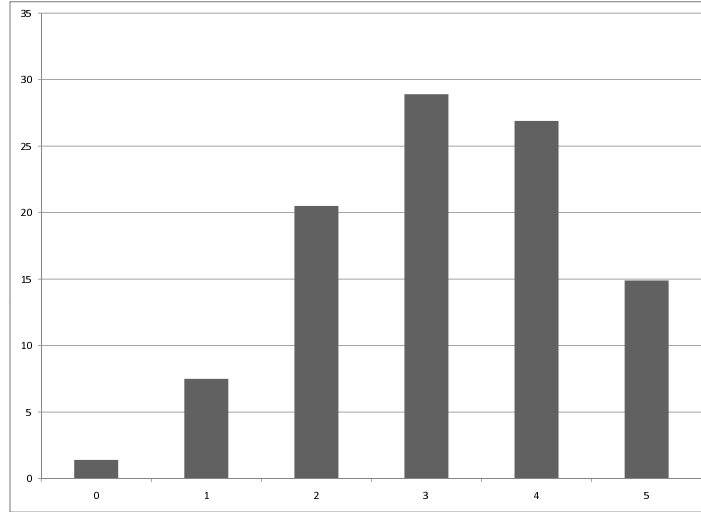
	Respondents	Percentage
<u>Behavior when commercials are aired</u>		
missing	187	1.23
“I watch commercials”	8720	57.17
“I switch TV stations when commercials are on air”	6347	41.61
<u>TV-usage behavior</u>		
missing	187	1.23
never or rarely	3860	25.31
almost every day or every day	11207	73.47
<u>Region</u>		
West Germany	12046	78.97
East Germany	3208	21.03
<u>Children in the household</u>		
children	6975	46.39
no children	8178	53.61
<u>Usage of ground coffee</u>		
missing	151	0.99
light user	6418	42.07
heavy user	8685	56.94
	Mean	Std. dev.
Age	40.38	10.88

**Heterogeneity in choice sets.** We close this section with a preliminary inspection of the data in order to establish the key facts that our model has to account for. We focus on choice sets as a source of additional heterogeneity in aggregate demand models.

Brand salience varies across consumers. Figure 1 shows the distribution of choice set sizes (measured by number of brands) across consumers. Consumers are clearly far from fully informed about brands. The bulk of consumers has between 2 and 4 brands in their choice set. The most likely possibility is that a consumer is aware

of 3 out of 5 major national brands. This finding is in line with evidence presented by Hauser & Wernerfelt (1990) that choice sets tend to be small with most people considering far fewer than the total number of available brands.

Figure 1: Distribution of choice set sizes across consumers.



Since different consumers have different choice sets, this additional source of heterogeneity has to be taken into account when modeling and estimating demand. This is not an easy task because the number of possible choice sets increases exponentially with the number of brands in the market. The previous literature has therefore made simplifying assumptions on the process of choice set formation in order to render the problem tractable, especially since it often lacked direct measures of brand salience and choice sets.

Sovinsky Goeree (2008) assumes that choice set membership is independent across brands, so that it suffices for her to model the marginal probability that a given brand is in the choice set of a given consumer. However, our data reveals that the probability that a consumer is aware of one brand is not independent of the probability that she is aware of another. This interdependence can already be seen from the Pearson

correlation coefficients in Table 4. We have further conducted an analysis of variance using a likelihood ratio test to compare the restricted model assuming independence with the unrestricted (saturated) model. The test statistic has a  $\chi^2$  distribution with 26 degree of freedom. Its value of 3914.36 with a  $p$ -value of less than 0.0001 confirms that indeed there are important interdependencies in brand salience.

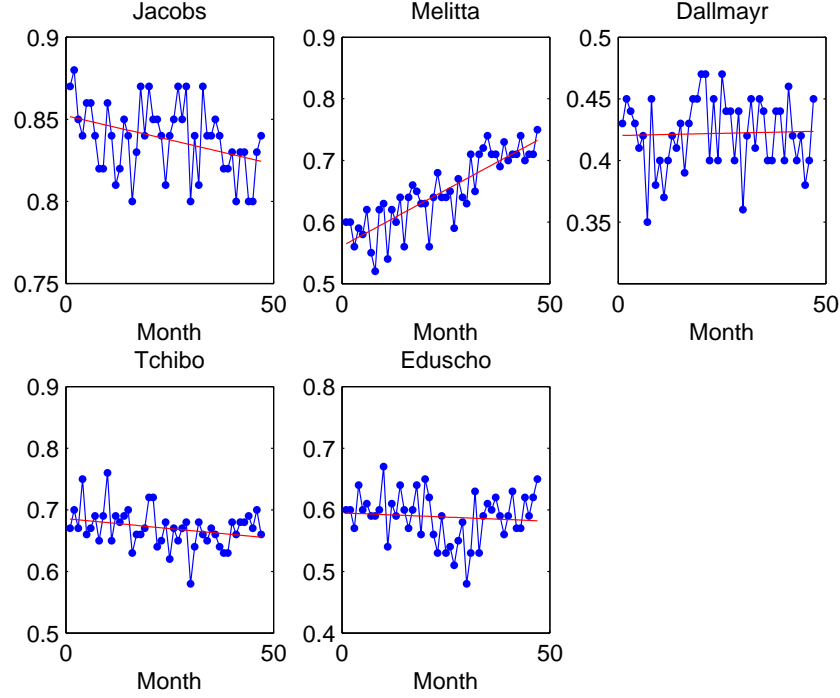
Table 4: Pearson correlation coefficients.  $p$ -values underneath.

	Eduscho	Dallmayr	Jacobs	Tchibo	Melitta
Eduscho	1.0000	0.0968	0.0176	0.4615	0.0287
		<.0001	0.0301	<.0001	0.0004
Dallmayr	0.0968	1.0000	0.0776	0.0924	0.0215
	<.0001		<.0001	<.0001	0.0081
Jacobs	0.0176	0.0776	1.0000	0.0067	0.1176
	0.0301	<.0001		0.4075	<.0001
Tchibo	0.4615	0.0924	0.0067	1.0000	-0.0159
	<.0001	<.0001	0.4075		0.0502
Melitta	0.0287	0.0215	0.1176	-0.0159	1.0000
	0.0004	0.0081	<.0001	0.0502	

Brand salience not only varies over consumers but also over time, as evident from Figure 2 where we depict the time path of brand salience (averaged across consumers) for our five major national brands. The brand salience of Melitta exhibits a clear upward trend, those of Jacobs and Tchibo trend down. All brands show considerable fluctuations in brand salience around its trend. The fact that brand salience can rise as well as fall calls into question the assumption in Albuquerque & Bronnenberg (2009) that choice sets are stable over time (see also Chiang, Chib & Narasimhan 1999).

Advertising is a prime force for the evolution of brand salience. Table 5 displays the estimated coefficients from a regression of the salience of a brand on its share in advertising expenditures (share of voice). As can be seen, we find a significant positive effect for 3 out of 5 brands. The point estimates for Tchibo and Jacobs are not significant, but we also only have 47 data points per brand. Overall, this first look at the data confirms previous findings that choice sets are not fixed but can

Figure 2: Evolution of brand salience over time.



be influenced by the marketing mix (Allenby & Ginter 1995, Mitra 1995, Siddarth et al. 1995).

In sum, the tracking study reveals a considerable amount of heterogeneity in choice sets with most consumers being aware of only a subset of the available brands. Choice sets vary across consumers and time. The independence and stationarity assumptions invoked in the previous literature to simplify modeling the process of choice set formation do not seem warranted. Most important, the standard approach of focusing on brand evaluations in determining choice from a given set of alternatives that the researcher has specified appears to ignore an important element of the consumer decision making, namely the existence of choice sets.

Table 5: Estimation results for regression of brand salience on share in advertising expenditures. Standard errors in parentheses.

	constant	advertising
Jacobs	0.8392* (0.0067)	-0.0003 (0.0277)
Melitta	0.5982* (0.0177)	0.2327* (0.0729)
Dallmayr	0.3951* (0.0137)	0.1296* (0.0635)
Tchibo	0.6674* (0.0088)	0.0108 (0.0366)
Edusho	0.5733* (0.0099)	0.1013* (0.0511)

### 3 Empirical Framework

We model a consumer’s choice of a brand. Importantly, the consumer may not consider all available brands; instead we allow for the possibility that her choice is restricted to a smaller set of brands. Much research views the formation of these choice sets as part of a multistage process that ultimately leads to choice (Shocker et al. 1991). The stages of this process are widely believed to be of a different nature. Whereas choice decisions in the final stage of brand selection are thought to be elaborate and compensatory in nature, the size and composition of consideration sets are viewed as arising from more simplistic, noncompensatory processes.

We draw on the existing literature and posit a two-stage process for brand choice:

1. Choice sets are determined.
2. Purchase decisions are made.

We thus model the probability that consumer  $n$  chooses brand  $j$  in period  $t$  as

$$\Pr_{nt}(j) = \sum_{\iota_{nt}} \Pr_{nt}(j|\iota_{nt})\Pr_{nt}(\iota_{nt}), \quad (1)$$



where  $\Pr_{nt}(j|\iota_{nt})$  is the probability that the consumer chooses the brand given choice set  $\iota_{nt}$  and  $\Pr_{nt}(\iota_{nt})$  is the probability of the consumer having this particular choice set. The brand choice probability  $\Pr_{nt}(j)$  then follows from the law of total probability.

Our model does not require us to be able to impute choice sets with certainty from the data. Instead, by specifying  $\Pr_{nt}(\iota_{nt})$  to be the probability that consumer  $n$  has choice set  $\iota_{nt}$  in period  $t$ , we allow for the possibility that choice sets themselves are latent. Some researchers have argued that a probabilistic model of choice set formation is more realistic because consumers may be unable to disclose the content of these sets or possibly even to understand the concept of a choice set (see, e.g., Shocker et al. 1991). Thus, following Manski (1977), we model the process of choice set formation in a probabilistic fashion.

Choice sets may simply reflect exogenous constraints on choice that arise from consumers' socio-economic characteristics. An individual's choice set may also depend on her subjective attitudes and perceptions (Shocker et al. 1991). There is a wide array of justifications for the existence of choice sets. The differentiation between long- and short-term memory in psychology is consistent with a reduction process where only items that are relevant to an immediate purpose are retrieved from storage and made accessible. The concept of bounded rationality (March & Simon 1958) and the economics of information (Stigler 1961) provide further rationales for a consumer not considering all available brands at each purchase occasion.

**Consumer utility.** Let consumers (households) be indexed by  $n = 1, \dots, N$ , brands by  $j = 0, \dots, J$ , and time periods (purchase occasions) by  $t = 1, \dots, T$ . The utility of consumer  $n$  from purchasing brand  $j$  at purchase occasion  $t$  is given by

$$u_{njt} = x_{jt}\beta_n + \gamma_n g_{jt} - \alpha_n p_{jt} + \xi_{jt} + \varepsilon_{njt}$$

and their utility from not purchasing (opting for the outside good) is given by

$$u_{n0t} = \varepsilon_{n0t}.$$

The observed characteristics  $x_{jt}$  of brand  $j$  in period  $t$  include our measures of promotional activities and a brand-specific constant that soaks up any time-invariant characteristics of the brand.  $p_{jt}$  denotes the price of brand  $j$  in period  $t$  and  $g_{jt}$  its accumulated goodwill. We capture heterogeneity in consumer preferences through a random coefficients specification

$$\begin{pmatrix} \alpha_n \\ \beta_n \\ \gamma_n \end{pmatrix} = \theta D_n + \nu_n,$$

where  $D_n$  are the demographics of consumer  $n$  and  $\nu_n = (\nu_{n1}, \dots, \nu_{nK}) \sim N(0, \Sigma)$  her tastes. Together with the idiosyncratic shocks  $(\varepsilon_{n0t}, \dots, \varepsilon_{nJt})$ , which we assume to be iid extreme value distributed, the random coefficients specification allows consumers to differ in their brand perceptions and in their sensitivities to marketing mix variables.

The demand shocks  $(\xi_{1t}, \dots, \xi_{Jt})$  are common across consumers and represent the characteristics of the various brands that are unobserved by the researcher. Because unobserved characteristics such as product quality and brand image are captured in the brand-specific constants,  $\xi_{jt}$  reflects time-varying factors like coupon availability and shelf space allocation that are unobserved to us but known to market participants. These demand shocks give rise to an endogeneity problem to the extent that the market participants condition their decisions on them. In our estimation we allow for price and advertising/goodwill to be endogenous.

**Choice sets.** We next incorporate choice sets into the above model. We acknowledge that choice sets vary across consumers and time and let  $\iota_{nt} = (\iota_{n1t}, \dots, \iota_{nJt}) \in \{0, 1\}^J$  indicate whether consumer  $n$  is aware of brand  $j$  at time  $t$ . Conditional on her tastes  $(\alpha_n, \beta_n, \gamma_n)$  and her choice set  $\iota_{nt}$ , the probability that consumer  $n$  chooses brand  $j$  at purchase occasion  $t$  is

$$\Pr_{nt}(j|\iota_{nt}) = \frac{\exp(x_{jt}\beta_n + \gamma_n g_{jt} - \alpha_n p_{jt} + \xi_{jt})\iota_{njt}}{1 + \sum_{k=1}^J \exp(x_{kt}\beta_n + \gamma_n g_{kt} - \alpha_n p_{kt} + \xi_{kt})\iota_{nkt}}. \quad (2)$$

In particular, if  $\iota_{njt} = 0$ , then the probability that brand  $j$  is chosen is zero. If  $\iota_{njt} = 1$ , then the probability that brand  $j$  is chosen depends on the utilities of only those brands that are also in the choice set of consumer  $n$  in period  $t$ . For example, if  $J = 3$  but consumer  $n$  is aware of only brands 1 and 2, then the probability that she chooses brand 1 is given by

$$\Pr_{nt}(1|(1, 1, 0)) = \frac{\exp(x_{1t}\beta_n + \gamma_n g_{1t} - \alpha_n p_{1t} + \xi_{1t})}{1 + \exp(x_{1t}\beta_n + \gamma_n g_{1t} - \alpha_n p_{1t} + \xi_{1t}) + \exp(x_{2t}\beta_n + \gamma_n g_{2t} - \alpha_n p_{2t} + \xi_{2t})}.$$

In the aggregate, the demand for brand 1 is effectively composed of different segments of consumers, namely those who only consider brand 1, those who consider both brands 1 and 2, those who consider both brands 1 and 3, and those who consider all three brands. Because we incorporate choice sets, our brand choice model in equation (1) is a mixture model.

In this paper, we take a consumer's choice set to be a primitive. We define the choice set as the set of brands in a product category that the consumer chooses amongst when making a purchase decision. In other words, these are the brands that have – ex ante – a purchase probability of more than zero. We then relate the probability that a consumer has a certain choice set back to her demographics, usage behavior, and the advertising expenditures of the various brands.

**Biased brand values and price sensitivity.** We close this section with a discussion of the importance of accounting for choice set heterogeneity in empirical studies of demand. We proceed in two steps. First, we show that consumer-level estimates of both brand values and price sensitivities are biased if we wrongly assume that consumers are fully informed. We analytically derive the direction of the bias in the demand primitives in a simple example. Second, we show that the assumption of full information causes firms to make flawed pricing decisions. Our example demonstrates that the resulting losses can be substantial.

While there are many methods to estimate demand, intuitively all of them calibrate the brand values to explain the level of demand in the data and the price

sensitivity to rationalize the change in quantity that results from a change in price. Throughout we assume that the true demand model is limited information with certain brand values and price sensitivity. By construction these primitives fully explain the level of demand at a given price and the change in demand resulting from a change in price.

To see the direction of the bias, suppose we wrongly assume that all consumers are aware of all brands. Because now all consumers demand all brands, in order to explain the level of demand in the data, we have to pick parameters ensuring that each individual consumer demands less than she actually is. Making the assumption of full information thus forces us to lower the brand values in order to match the data. Estimates of brand values will therefore be *downward biased* if we wrongly assume full information. Next consider a change in price. Since the demand of all consumers drops as the price rises, the full information model predicts a larger demand response than what we see in the data unless we make each individual consumer less price sensitive than she actually is. Estimates of price sensitivity will therefore also be *downward biased* if we wrongly assume full information.

To formalize the above arguments, we consider a simple example. Suppose there is just one brand (in addition to an outside good) and two groups of consumers. One group knows about the brand whereas the other does not. The demand of those consumers who know about the brand is  $q = \frac{\exp(1-p)}{1+\exp(1-p)}$  and there are  $\phi_1$  of those consumers. Naturally, the  $\phi_0 = 1 - \phi_1$  consumers who do not know about the brand, don't demand it. Total demand for the brand is thus  $Q = \phi_1 \cdot \frac{\exp(1-p)}{1+\exp(1-p)} + \phi_0 \cdot 0 = \phi_1 \frac{\exp(1-p)}{1+\exp(1-p)}$ . As price rises from 0 to 1, the demand for the brand drops from  $0.73\phi_1$  to  $0.5\phi_1$ .

To see how the full-information assumption can bias the estimated demand parameters, suppose that we fit a demand model of the form  $Q^{FI} = \frac{\exp(a-b \cdot p)}{1+\exp(a-b \cdot p)}$  to this data, where  $a$  is the brand value and  $b$  the price sensitivity to be estimated. To explain the data and accurately reflect demand before and after the price change, the

unknown coefficients must satisfy  $0.73\phi_1 = \frac{\exp(a-b \cdot 0)}{1+\exp(a-b \cdot 0)}$  (before the price change) and  $0.5\phi_1 = \frac{\exp(a-b \cdot 1)}{1+\exp(a-b \cdot 1)}$  (after). Figure 3 presents the estimates for  $a$  and  $b$  for various values of  $\phi_1$ . As can be seen, the full-information assumption causes us to underesti-

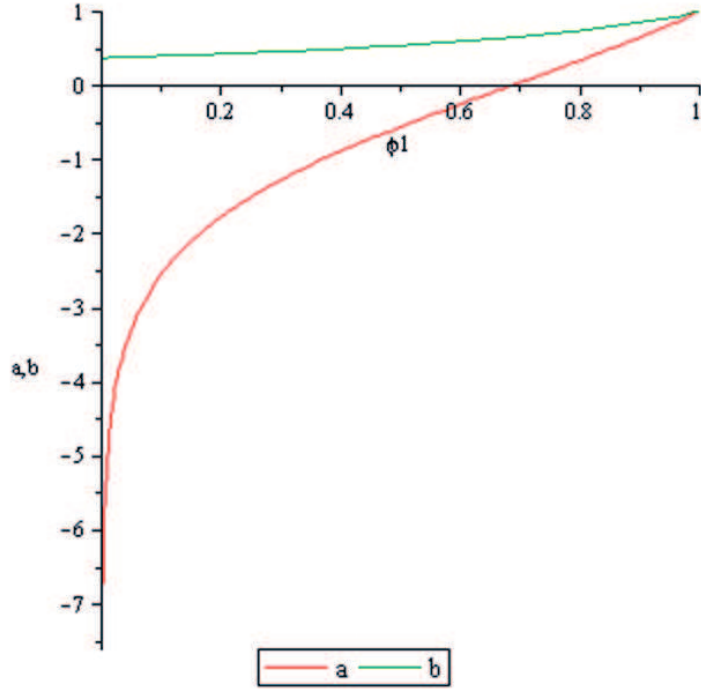


Figure 3: Estimated brand value  $a$  and price sensitivity  $b$ . True value is 1 for both.

mate both the brand value  $a$  and the price sensitivity  $b$ . Only if  $\phi_1 = 1$  so that the true demand model in fact satisfies the full-information assumption do we recover the correct consumer-level estimates of both brand value and price sensitivity.

To further develop our intuition for how wrongly assuming full-information biases price sensitivity, suppose for simplicity that consumers are identical and that products are identical. The limited information model can rationalize that identical consumers purchase different products with identical characteristics through differences in awareness. The full information model, in contrast, must rationalize this purchase pattern through the idiosyncratic error terms. But since the variance of the idiosyncratic error terms is normalized (or, equivalently, since we estimate the price

coefficient up to the scale of these error terms), this means that the estimated price coefficient becomes smaller in absolute value. Again, we see that in the full-information model consumers are estimated to be *less* price sensitive than they actually are.

**Managerial implications.** Next we show that the assumption of full information causes firms to make flawed pricing decisions. Recall that in equilibrium the inverse elasticity rule  $\frac{p-c}{p} = \frac{\theta}{|\varepsilon|}$  holds, where  $|\varepsilon|$  is the absolute value of the price elasticity of the demand and  $\theta$  is a conduct parameter whose value depends on the specific model of oligopolistic competition. Given that we have already shown that the full-information assumption yields biased estimates of the demand primitives, we can hardly expect to obtain the correct elasticity of demand from the misspecified demand model. Continuing with our example, a routine calculation shows that at a price of  $p = 1$ , the full-information demand model implies an elasticity between -0.38 (at  $\phi_1 = 0$ ) and -0.5 (at  $\phi_1 = 1$ ), compared to an elasticity of -0.5 in the true limited-information demand model. Consequently, the firm makes a flawed pricing decision. Figure 4 presents, for various values of  $\phi_1$ , the optimal price as derived from either the true limited-information demand model or the misspecified full-information demand model.

As can be seen, with the true limited-information demand model at hand, the firm sets a price of 1.57. In contrast, if the firm wrongly assumes full information (and calibrates that model to fit the data), then it generally sets a much higher price ranging from 2.62 to 1.57 depending on the value of  $\Phi_1$ . This is because in the full-information demand model consumers appear to be less price sensitive than they actually are. Unsurprisingly, making flawed pricing decisions is detrimental for profitability. Indeed, profit losses can be quite substantial as Figure 5 illustrates.

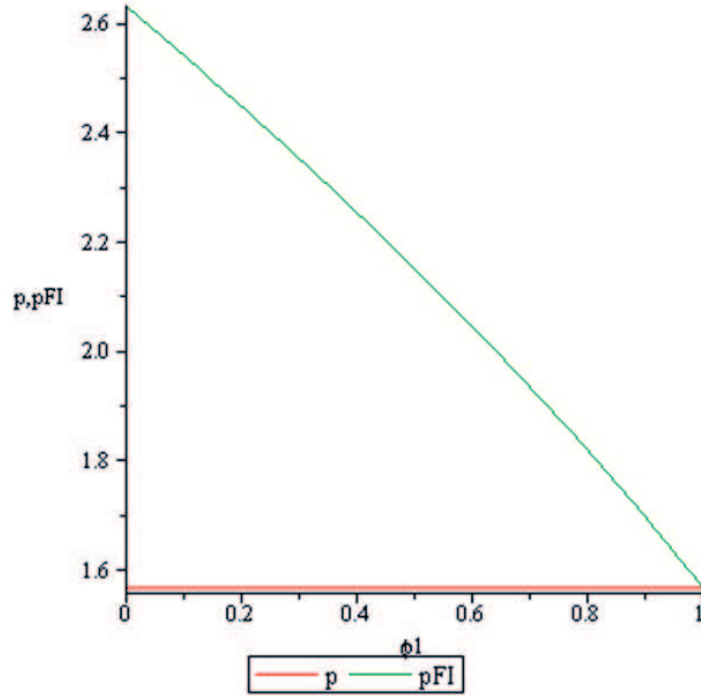


Figure 4: Optimal price for true limited-information demand model ( $p$ ) and misspecified full-information model ( $p^{FI}$ ).

## 4 Models

An important strand of the literature on choice set formation assumes that consumer preferences exist without reference to choice sets. Choice sets reflect the costs and benefits of acquiring and processing information about the available brands (Hauser & Wernerfelt 1990, Roberts 1989, Roberts & Lattin 1991). Then a two-stage process is considered where, in the first stage, brands are selected into the choice set according to their utility; in particular, the most highly valued brands are included in the choice set and the other brands are excluded. In the second stage, the consumer chooses which of the brands in her choice set to purchase. Because choice set formation is based on utility, the choice set provides no information that is not available from the utility function. As a consequence, empirical implementations of this type of models have been shown to offer no substantial advantages over a straight choice model in

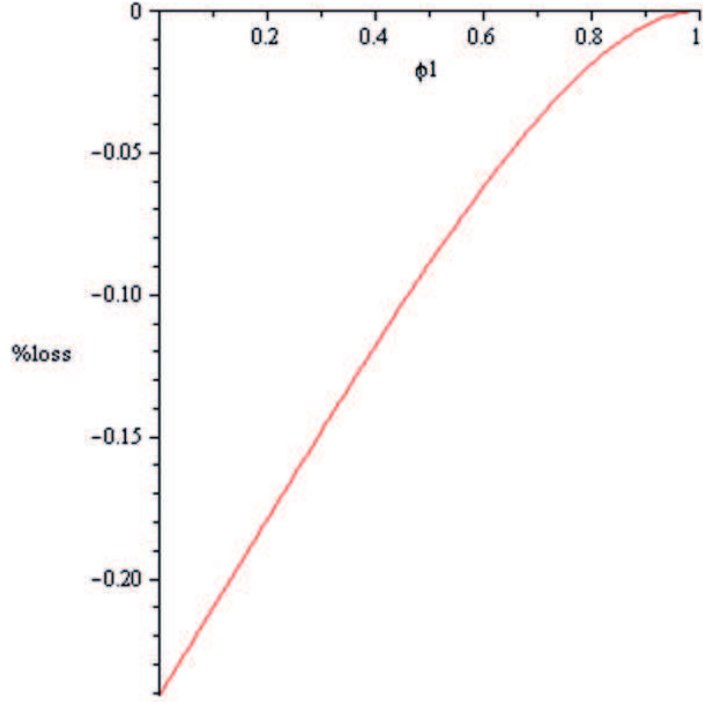


Figure 5: Percentage profit loss from basing pricing decisions on the misspecified full-information model  $(\frac{\pi^{FI}-\pi}{\pi})$ .

terms of forecasting demand or explaining consumer behavior (see, e.g., Horowitz & Louviere 1995).

We develop three ways of using the tracking survey to account for choice set heterogeneity in the population of consumers. First we treat the choice set of a consumer as a random coefficient. Second, we obtain nonparametric estimates for the aggregate probabilities of the various possible choice sets from the tracking study. Third, we formulate a model that links the choice set of a consumer to her demographics, usage behavior, and the advertising expenditures of the various brands. Viewed through the lens of our brand choice model in equation (1), the three different ways of incorporating micro-level survey data from the tracking study amount to different specifications of  $\Pr_{nt}(\iota_{nt})$ , the probability of consumer  $n$  having choice set  $\iota_{nt}$  in period  $t$ .



**Model 1: Brand awareness as random coefficient.** One approach is to think of the choice set of a consumer as just another random coefficient. This random coefficient captures the heterogeneity of the population with respect to brand awareness, just as the traditional random coefficients specification captures differences in tastes across consumers. Integrating out over both types of heterogeneity, the market share of product  $j$  in period  $t$  is

$$s_{jt} = \int \frac{\exp(x_{jt}\beta_n + \gamma_n g_{jt} - \alpha_n p_{jt} + \xi_{jt})\iota_{njt}}{\underbrace{1 + \sum_{k=1}^J \exp(x_{kt}\beta_n + \gamma_n g_{kt} - \alpha_n p_{kt} + \xi_{kt})\iota_{nkt}}_{\Pr_{nt}(\iota_{nt})}} dF(\nu_n, D_n, \iota_{nt}). \quad (3)$$

This model is a special case of the brand choice model in equation (1) where  $\Pr_{nt}(\iota_{nt})$  puts a point mass on the choice set  $\iota_{nt}$  stated in the tracking study.

To compute predicted market shares, we proceed as follows:

1. Draw an individual from the tracking study (demographics  $D_n$  and choice set  $\iota_{nt}$ ).
2. Draw the random taste component  $\nu_n \sim N(0, \Sigma)$ .
3. Compute choice probabilities  $\Pr_{nt}(j|\iota_{nt})$  from equation (2).
4. Repeat for another individual. Average to obtain the predicted market shares in equation (3).

To reduce the variance of the simulation we employ an importance sampling algorithm as suggested by Berry et al. (1995). The key idea is to oversample consumers that have a high likelihood of making a purchase (instead of opting for the outside good). As we simulate more consumers that make a purchase, the prediction of the inside shares becomes more accurate.

We use the contraction mapping in Berry (1994) to invert out the unobserved demand shocks  $(\xi_{1t}, \dots, \xi_{Jt})$  from the observed market shares  $(s_{0t}, s_{1t}, \dots, s_{Jt})$ . As in Berry et al. (1995), estimation is then based on the moments

$$E(\xi_t|Z_t) = 0,$$

where  $Z_t$  is a vector of instruments that are orthogonal to the demand shocks  $(\xi_{1t}, \dots, \xi_{Jt})$ .

We use the cost of raw coffee as an instrument for price.<sup>2</sup> The cost of raw coffee is determined in world-wide commodity markets and can thus be taken as exogenous to coffee manufacturers in Germany. We instrument for advertising/goodwill with the cost of TV commercials, which again should be independent of the actual advertising behavior of coffee manufacturers.<sup>3</sup> The instruments for price and advertising/goodwill along with the other exogenous demand shifters are interacted with brand and retail-chain dummies. The  $R^2$  of the first-stage regression for price is 0.8097, and the  $F$ -test of the significance of the instruments in explaining price is 4.8826 with a  $p$ -value of 0.0000. The  $R^2$  of the first-stage regression for advertising is 0.2059, and the corresponding  $F$ -test of the significance is 1.7635 with a  $p$ -value of 0.0064.

**Model 2: Aggregate choice set probabilities.** The above model takes the choice sets stated in the tracking study at face value. This may not be entirely plausible since actual purchase decision are made in a store environment. Further evidence that the stated choice sets should not be taken too literally are various inconsistencies in the variables in the tracking study (see again Table 2). For example, 8165 respondents consider Jacobs but 718 of them do not recall the brand, and similarly for the other brands (880 out of 5392 for Melitta, 2058 out of 5409 for Dallmayr, 912 out of 5407 for Tchibo, and 1011 of 4635 for Eduscho). This suggests treating choice sets as latent in a probabilistic model of choice set formation.

As a first step, we assume that our direct measures of brand salience and choice sets in the tracking study are correct on average. Let  $\phi_{it}$  denote the probability of

---

<sup>2</sup>We obtained commodity prices for coffee from the New York Stock Exchange. There are different types of contracts, and we selected the contract with the highest correlation with shelf prices (coffee price mean high second near by). We then adjusted for the exchange rate.

<sup>3</sup>We obtained the average price to reach 1000 viewers across all TV stations with a market share greater than 0.1%, weighted by the market shares of those TV stations.

choice set  $\iota_t$  in period  $t$ . The market share of brand  $j$  in period  $t$  is

$$\begin{aligned}
s_{jt} &= \int \left( \sum_{\iota_t} \underbrace{\phi_{\iota_t}}_{\Pr_{nt}(\iota_{nt})} \frac{\exp(x_{jt}\beta_n + \gamma_n g_{jt} - \alpha_n p_{jt} + \xi_{jt})\iota_{jt}}{\underbrace{1 + \sum_{k=1}^J \exp(x_{kt}\beta_n + \gamma_n g_{kt} - \alpha_n p_{kt} + \xi_{kt})\iota_{kt}}_{\Pr_{nt}(j|\iota_{nt})}} \right) dF(\nu_n, D_n) \\
&= \sum_{\iota_t} \left( \phi_{\iota_t} \int \frac{\exp(x_{jt}\beta_n + \gamma_n g_{jt} - \alpha_n p_{jt} + \xi_{jt})\iota_{jt}}{1 + \sum_{k=1}^J \exp(x_{kt}\beta_n + \gamma_n g_{kt} - \alpha_n p_{kt} + \xi_{kt})\iota_{kt}} dF(\nu_n, D_n) \right). \quad (4)
\end{aligned}$$

Note that the second line follows from the first by the fact that the choice set probabilities are the same across consumers. The model in equation (4) is a special case of the brand choice model in equation (1) where  $\Pr_{nt}(\iota_{nt}) = \Pr_{nt}(\iota_t) = \phi_{\iota_t}$  is the same across consumers. Equation (4) nicely illustrates that aggregate demand is effectively composed of different segments of consumers that differ in their choice sets. The brand choice model is thus a mixture model.

Practically the problem is that the number of possible choice sets is  $2^J$  and thus increases exponentially with the number of brands  $J$ . A lot of data is required to obtain precise period-specific estimates of the choice set probabilities. To facilitate the estimation, we assume that the choice set probabilities are stationary throughout the sample period. Setting  $\phi_{\iota_t} = \phi_{\iota}$ , we estimate  $\phi_{\iota}$  by the relative frequency with which choice set  $\iota$  occurs in the tracking study.

Finally, we compute predicted market shares as follows:

1. Draw an individual from the tracking study (demographics  $D_n$ ).
2. Draw the random taste component  $\nu_n \sim N(0, \Sigma)$ .
3. Compute choice probabilities  $\Pr_{nt}(j|\iota_{nt}) = \Pr_{nt}(j|\iota)$  from equation (2). Multiply together with choice set probability  $\phi_{\iota}$  and sum over all possible choice sets.
4. Repeat for another individual. Average to obtain the predicted market share in equation (4).

**Model 3: Individual choice set probabilities.** To gain insights into the process of choice set formation, we next explicitly model the choice set probabilities at the individual level as a function demographics, usage behavior, and the advertising expenditures of the various brands. Let  $\Pr_{nt}(\iota_{nt}) = \phi_{\iota_{nt}}(g_t, D_n)$  denote the probability that consumer  $n$  in period  $t$  has choice set  $\iota_{nt}$ .  $g_t = (g_{1t}, \dots, g_{Jt})$  is the goodwill of the various brands in period  $t$  and, as above,  $D_n$  are the demographics of consumer  $n$  in period  $t$ , including her usage behavior. Like in the previous approach, our model allows for measurement error in the sense that the actual choice sets can differ from the ones stated in the tracking study. Unlike the previous approach, we account for heterogeneity in choice sets.

The market share of brand  $j$  in period  $t$  is

$$s_{jt} = \int \left( \sum_{\iota_{nt}} \underbrace{\phi_{\iota_{nt}}(g_t, D_n)}_{\Pr_{nt}(\iota_{nt})} \underbrace{\frac{\exp(x_{jt}\beta_n + \gamma_n g_{jt} - \alpha_n p_{jt} + \xi_{jt})\iota_{njt}}{1 + \sum_{k=1}^J \exp(x_{kt}\beta_n + \gamma_n g_{kt} - \alpha_n p_{kt} + \xi_{kt})\iota_{nkt}}}_{\Pr_{nt}(j|\iota_{nt})} \right) dF(\nu_n, D_n). \quad (5)$$

Because the choice set probabilities do not rely on the assumption that consumer preferences exist without reference to choice sets, the empirical model we specify for  $\phi_{\iota_{nt}}(g_t, D_n)$  is largely descriptive.

In practice we specify  $\phi_{\iota_{nt}}(g_t, D_n)$  as a multinomial logit model with the  $2^J$  possible choice sets as dependent variable. Consistent with our two-stage process for brand choice, we assume that the error terms in the second stage of the brand choice process are independent (conditional on the observables) of those underlying the multinomial logit model in the first stage (Horowitz & Louviere 1995).

This approach is in line with the literature on consumer behavior which strongly favors noncompensatory models for the determination of choice sets (for an excellent summary, see, Shocker et al. 1991). It is also consistent with the economics literature which assumes that a product is relevant for the consumer's ultimate purchase decision for largely exogenous reasons such as exposure to advertising (Butters 1977, Stegeman

1991, Robert & Stahl 1993). An advantage of our approach is that it provides us with the opportunity to investigate if a consumer’s choice set contains her most-highly valued brands.

Moreover, we are able to evaluate empirically if there are influence factors that make it more or less likely that a brand makes it into the choice set of a consumer without affecting her preferences for the brand. The literature on consumer behavior suggests that this is the case. For example, Nedungadi (1990) demonstrates an effect on choice probabilities by changing consideration probabilities, without altering brand evaluations, by differential prompting of brands in product categories with known structures. Our focus is on advertising and the ways in which it can affect sales. By parameterizing both  $\Pr_{nt}(j|\iota_{nt})$ , the probability that consumer  $n$  choose brand  $j$  in period  $t$  given choice set  $\iota_{nt}$ , and  $\Pr_{nt}(\iota_{nt})$ , the probability of the consumer having this particular choice set, as functions of advertising, we are able to separately examine the availability effect of advertising on brand salience from the substitution effect on consumer preferences.

The drawbacks of using a multinomial logit model for the choice set probabilities are well-known and include the independence from irrelevant alternatives property and, in this setting, the curse of dimensionality as the number of possible choice sets increases exponentially with the number of brands. Further, demographics and measures of usage behavior cannot be easily incorporated and must be interacted with brand dummies because they otherwise cancel out. An alternative is to specify a threshold-crossing model for choice set membership along the lines of van Nierop et al. (forthcoming). The drawback of this approach is that the resulting parameter estimates are difficult to interpret because the interdependency of the choice sets is specified only through the correlations in the error terms.

We estimate the model in equation (5) in two steps (models 1 and 2 are estimated in one step). First, we estimate  $\phi_{\iota_{nt}}(g_t, D_n)$  by maximum likelihood using the tracking study. Then we substitute these estimates into equation (5). Next, we compute

predicted market shares as follows:

1. Draw an individual from the tracking study (demographics  $D_n$ ).
2. Draw the random taste component  $\nu_n \sim N(0, \Sigma)$ .
3. Compute choice probabilities  $\Pr_{nt}(j|\iota_{nt})$  from equation (2). Compute choice set probabilities  $\phi_{\iota_{nt}}(g_t, D_n)$ . Multiply together and sum over all possible choice sets.
4. Repeat for another individual. Average to obtain the predicted market share in equation (5).

## 5 Estimation Results

In what follows we illustrate the empirical strategies developed in Section 4 using data from the ground coffee category in Germany. To establish a baseline we begin with a model of aggregate demand that accounts for consumer heterogeneity by modeling their tastes using a random coefficients specification (Berry 1994, Berry et al. 1995). We then depart from the full information assumption and incorporate limited information into the model in order to demonstrate the importance of accounting for choice set heterogeneity in aggregate demand estimation. We show that wrongly assuming full information biases the estimates of the demand primitives in the direction derived in Section 3.

Table 6 displays the coefficient estimates and Tables 7 and 8 display various measures derived from these estimates.<sup>4</sup> In each table, the first column labeled “No CS” is the standard aggregate demand model assuming full information. The remaining

---

<sup>4</sup>We conducted a grid search to determine the persistence parameter  $\lambda$  for goodwill. It turned out that a value of 0.8 is appropriate to explain market share variations across time when goodwill is part of the utility function. This value is also in line with the estimates presented in Dube, Hitsch & Manchanda (2005) and Doganoglu & Klapper (2006). We also investigated whether a brand-specific persistence parameter helps to explain market shares, but this was not the case.

columns correspond to the three ways we suggest for incorporating limited information in the form of choice sets.

**Baseline.** In the model without CS we obtain reasonable estimates for the parameters of the demand system. Price and goodwill have significant effects, as do in-store communication and feature. There is a significant and negative time trend in line with industry evidence that shows that yearly per capita consumption of coffee in Germany has fallen by 10% from 1990 to 2002.

The random coefficients are jointly significant, indicating that there are considerable differences in tastes across consumers.<sup>5</sup> The estimated covariance of the random coefficients for price and goodwill is negative and captures the idea that consumers who are price sensitive are also less prone to pay attention to advertising.

Because buying the same brand at a different retail chain may provide for a different purchase experience, we include a constant for each possible combination of brand and retail chain in the model. As can be seen from the first column of Table 7, the average value of the brand across all retail chains for Jacobs is -1.74 for Jacobs, -2.60 for Melitta, -2.43 for Dallmayr, -1.31 for Tchibo, and -2.37 for Eduscho. These estimates match with brand managers' expectations about how consumers value the various brands. Tchibo is viewed as a strong brand that during our sample period did not have to rely too much on promotional support to push sales. Jacobs, the leading brand in terms of sales, on the other hand, generated much of its sales through promotions (see Table 1). Melitta, the cheapest brand, which also engaged heavily in promotions, has the lowest average brand value.

The price elasticities implied by our estimates are reasonable and range between -3.71 for Melitta to -4.71 for Tchibo, see the first column of Table 8. These price sensitivities are very comparable to what other researchers have found in the ground coffee category (Guadagni & Little 1983, Krishnamurthi & Raj 1991).

---

<sup>5</sup>Adding demographics to the random coefficients specification did not improve the estimates.

Table 6: Estimation results for aggregate demand model. Standard errors in parentheses.

	No CS	CS-as-RC	Agg-CS-Prob	Ind-CS-Prob
Price	-0.5778* (0.0453)	-0.6616* (0.0591)	-0.6344* (0.0571)	-0.6286* (0.0504)
Communication	0.1923* (0.0372)	0.2112* (0.0519)	0.2074* (0.0400)	0.1923* (0.0376)
Feature	1.1387* (0.0395)	1.2705* (0.0541)	1.1806* (0.0421)	1.1543* (0.0404)
Goodwill	1.1794* (0.2193)	1.3271* (0.3151)	1.2052* (0.2469)	1.1838* (0.2282)
Trend	-0.0387* (0.0022)	-0.0301* (0.0034)	-0.0407* (0.0023)	-0.0411* (0.0022)
RC Price	0.0360 (0.0494)	0.0553 (0.0512)	0.0709 (0.0535)	0.0455 (0.0506)
RC Goodwill	2.0661* (0.3902)	1.6295* (0.5453)	2.0171* (0.4427)	2.0138* (0.4171)
Cov RC Price RC Goodwill	-0.1086* (0.0221)	-0.1857* (0.0225)	-0.1244* (0.0243)	-0.1203* (0.0234)
Jacobs/Edeka	-1.7816* (0.2084)	-1.3326* (0.3165)	-1.1725* (0.2449)	-1.2854* (0.2224)
Melitta/Edeka	-2.9556* (0.1911)	-2.2966* (0.2873)	-2.1245* (0.2264)	-2.2381* (0.2045)
Dallmayr/Edeka	-2.3978* (0.2201)	-1.2380* (0.3351)	-1.0878* (0.2561)	-1.1895* (0.2337)
Tchibo/Edeka	-1.3208* (0.2283)	-0.5430 (0.3451)	-0.4487 (0.2642)	-0.5567* (0.2423)
Eduscho/Edeka	-2.4709* (0.2014)	-1.4736* (0.3080)	-1.5087* (0.2368)	-1.6213* (0.2147)
Jacobs/Markant	-1.8040* (0.2011)	-1.4093* (0.3034)	-1.2257* (0.2363)	-1.3405* (0.2143)
Melitta/Markant	-2.7267* (0.1902)	-2.1147* (0.2865)	-1.9071* (0.2248)	-2.0185* (0.2033)
Dallmayr/Markant	-2.5199* (0.2182)	-1.3706* (0.3296)	-1.2130* (0.2549)	-1.3139* (0.2321)
Tchibo/Markant	-1.3674* (0.2271)	-0.5881 (0.3430)	-0.5011 (0.2634)	-0.6102* (0.2412)
Eduscho/Markant	-2.2956* (0.1992)	-1.3033* (0.3059)	-1.3334* (0.2346)	-1.4455* (0.2126)
Jacobs/Metro	-1.9145* (0.2167)	-1.4579* (0.3257)	-1.3014* (0.2528)	-1.4143* (0.2303)
Melitta/Metro	-2.8624* (0.2066)	-2.1878* (0.3098)	-1.9783* (0.2421)	-2.0918* (0.2201)
Dallmayr/Metro	-2.8412* (0.2333)	-1.6075* (0.3500)	-1.4978* (0.2700)	-1.6013* (0.2473)
Tchibo/Metro	-1.6653* (0.2265)	-0.8845* (0.3428)	-0.7988* (0.2624)	-0.9102* (0.2405)
Eduscho/Metro	-2.1553* (0.2008)	-1.1727* (0.3078)	-1.1920* (0.2360)	-1.3038* (0.2142)
Jacobs/Rewe	-1.7125* (0.2070)	-1.2692* (0.3127)	-1.1095* (0.2424)	-1.2272* (0.2202)
Melitta/Rewe	-2.5646* (0.1999)	-1.9078* (0.3016)	-1.7045* (0.2355)	-1.8195* (0.2134)
Dallmayr/Rewe	-2.7630* (0.2324)	-1.5724* (0.3512)	-1.4201* (0.2689)	-1.5176* (0.2462)
Tchibo/Rewe	-0.8562* (0.2305)	-0.0763 (0.3480)	0.0079 (0.2662)	-0.0985 (0.2442)
Eduscho/Rewe	-2.0922* (0.2011)	-1.1171* (0.3078)	-1.1343* (0.2363)	-1.2457* (0.2143)
Jacobs/Spar	-1.7179* (0.2188)	-1.2946* (0.3299)	-1.1190* (0.2545)	-1.2330* (0.2322)
Melitta/Spar	-2.7150* (0.2059)	-2.0433* (0.3094)	-1.8527* (0.2420)	-1.9644* (0.2195)
Dallmayr/Spar	-2.1865* (0.2415)	-0.9283* (0.3640)	-0.8071* (0.2783)	-0.9094* (0.2557)
Tchibo/Spar	-1.1705* (0.2286)	-0.3830 (0.3450)	-0.3018 (0.2641)	-0.4110 (0.2424)
Eduscho/Spar	-2.4529* (0.2035)	-1.4753* (0.3112)	-1.4926* (0.2389)	-1.6030* (0.2171)
Jacobs/Tengelmann	-1.4897* (0.2085)	-0.9924* (0.3149)	-0.8391* (0.2445)	-0.9621* (0.2220)
Melitta/Tengelmann	-1.7781* (0.1886)	-1.1136* (0.2856)	-0.9288* (0.2241)	-1.0493* (0.2021)
Dallmayr/Tengelmann	-1.9018* (0.2233)	-0.6970* (0.3382)	-0.5703* (0.2598)	-0.6766* (0.2372)
Tchibo/Tengelmann	-1.4633* (0.2319)	-0.6721 (0.3499)	-0.5930* (0.2679)	-0.7024* (0.2459)
Eduscho/Tengelmann	-2.7725* (0.2011)	-1.7279* (0.3059)	-1.7405* (0.2366)	-1.8537* (0.2149)
Objective	439.5768	417.7987	453.0381	468.1468



Table 7: Brand values.

	No CS	CS-as-RC	Agg-CS-Prob	Ind-CS-Prob
Jacobs	-1.74	-1.23	-1.13	-1.24
Melitta	-2.60	-1.94	-1.75	-1.86
Dallmayr	-2.43	-1.24	-1.10	-1.20
Tchibo	-1.31	-0.52	-0.44	-0.55
Eduscho	-2.37	-1.38	-1.40	-1.51

Table 8: Price elasticities.

	No CS	CS-as-RC	Agg-CS-Prob	Ind-CS-Prob
Jacobs	-3.9519	-4.6042	-5.0313	-4.3326
Melitta	-3.7172	-4.3366	-4.7213	-4.0817
Dallmayr	-4.4168	-5.2862	-5.6208	-4.8525
Tchibo	-4.7070	-5.1610	-5.9362	-5.1134
Eduscho	-4.0701	-4.6021	-5.1598	-4.4431

**Biased brand values.** As evident from comparing the first column of Table 6 to the remaining columns, ignoring choice set heterogeneity results in biased estimates. To begin with, there is a difference in the estimated brand values (measured by the average value of the brand constant across retail chains). As can be seen from comparing the first column of Table 7 to the remaining columns, brand values tend to increase as choice set heterogeneity is accounted for. Put differently, wrongly assuming full information biases the estimated brand values downward, just as we derived in Section 3.

**Biased price sensitivities.** Comparing the first column of Table 8 to the remaining columns shows that ignoring choice set heterogeneity leads to incorrect price elasticities. There are two reasons for this. First, as can be seen from Table 6, the price coefficient changes as choice set heterogeneity is accounted for. As shown in in Section 3, wrongly assuming full information biases the estimated price coefficient downward

(i.e., it is less negative). Second, the functional form for aggregate demand changes. To see this, suppose there were no random coefficients. Then the model without CS reduces to a logit model and, consequently, suffers from the independence of irrelevant alternatives property. The models with CS, in contrast, are mixture models and therefore do not suffer from this property. Hence, price elasticities necessarily change.

Our empirical model, of course, incorporates heterogeneity in consumer preferences through a random coefficients specification. The estimated price elasticities in Table 8 indicate that the model without CS systematically understates the price sensitivity of consumers. This may seem surprising at first glance, especially since Sovinsky Goeree (2008) argues to the contrary that estimated price elasticities may be too high. Her reasoning is that because traditional demand models assume that consumers are aware of – and hence choose among – all brands in the market, they overstate the degree of competition in the market when in actuality most consumers are aware of only a small subset of brands. Sovinsky Goeree’s (2008) argument mixes supply and demand side aspects and does not stand up to a closer examination. In particular, in light of our discussion in Section 3, it becomes evident that her argument misses the key point that demand is estimated to match the levels and changes in quantities in the data. To explain the price-quantity movements in the data, the full-information model has to elicit a smaller response from each individual consumer than the limited-information model. Hence, the model without CS can be expected to systematically understate the price sensitivity of consumers.

**Choice set formation: Demographics and usage behavior.** Having established the importance of accounting for choice set heterogeneity in aggregate demand estimation, we turn our attention to the model of individual choice set probabilities that links the choice set of a consumer to her demographics, usage behavior, and the advertising expenditures of the various brands. The estimated coefficients for our multinomial logit model in Table 9 indicate what makes it more or less likely that a

brand is included in the choice set.

The estimates again have face validity. Older consumers are more likely to be aware of Dallmayr, Tchibo, and Eduscho. East Germans are more likely to be aware of Jacobs, Melitta, and Dallmayr but less likely to be aware of Tchibo and Eduscho. This effect is related to the period prior to the reunification of Germany. During this period Tchibo and Eduscho had exclusive coffee store outlets in West Germany that helped both brands to establish a high degree of salience in West Germany. The other brands competed with Tchibo and Eduscho through TV commercials that could also be seen in East Germany. Consumers who have kids are more likely to be aware of Melitta, Dallmayr, and Eduscho. In the case of Melitta a possible explanation may lie in the TV commercials themselves that told funny stories about a father and his son that may have appealed to families with kids. Consumers who pay attention to TV commercials are more likely to be aware of Jacobs and Melitta and, to some extent, also Dallmayr and Eduscho. Consumer who watch more TV have a lower probability of having Tchibo in their choices set, but a higher probability of having Jacobs and Melitta. Finally, consumers who report themselves as drinking a lot of coffee are more likely to be aware of the cheaper brands Melitta and Eduscho.

**Choice set formation: Advertising.** Importantly, as Table 9 makes clear, goodwill has a significant positive effect on the probability of choice set membership. Besides contributing to the utility that a consumer derives from the brand by building goodwill, advertising makes it more likely that the consumer is aware of the brand in the first place. Hence, according to our estimates, advertising plays a dual role of impacting both brand salience and consumer preferences.

Table 10 decomposes the change in market shares due to a simultaneous increase in the advertising expenditures of all brands. It shows both the substitution effect on consumer preferences and, ultimately, market shares, and the availability effect of advertising on choice sets as well as their combined effect. Interestingly, the two

effects can go in opposite directions. For example, the substitution effect causes the market share of Melitta to increase whereas the availability effect causes it to decrease. This finding suggests that Melitta is a relatively weak competitor in the sense that it depends on sales to consumers that have no (or few) other brands in their choice sets. In contrast, Dallmayr, Tchibo, and perhaps also Eduscho stand to gain from more fully informed consumers.

Table 10 further suggests the impact of advertising on choice sets is weaker than that on utility. This does not necessarily mean that the first channel is less important than the second because it may be the case that much of the impact of advertising is already captured in the choice set constants and the other explanatory variables (such as age) of our model of individual choice set probabilities. Without a measure of the advertising exposure of the different demographic groups, it is hard to disentangle the direct effect of advertising from its indirect effect through demographics.

The weaker impact through choice sets may also be driven by supply-side considerations. In theory, the marginal impact of a dollar should be equalized across the two channels. But while increasing brand preference always boosts price and profits, increasing the share of consumers who are aware of one's brand may not. In fact, as Fershtman & Muller (1993) and Boyer & Moreaux (1999) show, doing so may provoke a competitive response that puts pressure on prices, and profits may fall as a result. Hence, the return to awareness-increasing advertising may be quite low (or even negative), causing a brand to design advertising campaigns that are aimed at elevating consumer preference rather than brand salience. Our estimates are consistent with a supply-side story like this, but we acknowledge that this finding may be highly-specific to the product category we study. Since Mitra & Lynch (1995) and Clark, Doraszelski & Draganska (2009) show that advertising can have a much stronger effect on the size of the choice set than on the relative strength of consumer preferences, we believe that applying our empirical techniques to other categories besides ground coffee may be a fruitful venue for future research.

Table 9: Estimation results for individual choice set probabilities model. Standard errors in parentheses. Choice set constants omitted for brevity.

	Jacobs	Melitta	Dallmayr	Tchibo	Eduscho
age	-0.026 (0.018)	-0.021 (0.012)	0.129* (0.014)	0.069* (0.012)	0.025* (0.011)
region (1=east, 0=west)	0.9202* (0.061)	0.1511* (0.029)	0.2735* (0.029)	-0.5947* (0.035)	-0.7146* (0.037)
children (1=no, 0=yes)	-0.043 (0.030)	-0.115* (0.033)	-0.113* (0.032)	-0.018 (0.025)	-0.071* (0.025)
ad viewing (1=watch commercials, 0=no)	0.118* (0.036)	0.166* (0.032)	0.093* (0.032)	-0.034 (0.023)	0.095* (0.024)
TV frequency (1=rarely, 0=daily)	-0.083* (0.030)	-0.126* (0.036)	0.035 (0.036)	0.096* (0.027)	-0.049 (0.029)
coffee consumption (1=heavy user, 0=light user)	0.071* (0.036)	0.131* (0.032)	0.044 (0.032)	-0.081* (0.026)	0.082* (0.026)
goodwill	0.054* (0.017)	0.054* (0.017)	0.054* (0.017)	0.054* (0.017)	0.054* (0.017)

Table 10: Channels of advertising impact. Change in market shares due to simultaneous increase in advertising expenditures.

	base	utility	choice set	total
Jacobs	33.0156	32.2465	32.8591	32.0847
Melitta	19.4704	19.6662	19.4047	19.5972
Dallmayr	11.2121	11.8980	11.4208	12.1167
Tchibo	21.2764	22.1976	21.2928	22.2153
Eduscho	15.0255	13.9917	15.0226	13.9861

**Choice set and utility function.** We investigate if it is really a consumer’s most-highly valued brands that make it into her choice set as assumed by compensatory approaches to choice set formation (Hauser & Wernerfelt 1990, Roberts 1989, Roberts & Lattin 1991). To this end, we compute the systematic part of an individual’s utility for all brands using our coefficient estimates. Then we check whether the brand with the highest utility is a member of the consumer’s choice set. In some cases a substantial percentage of individuals fails this check. Table 11 shows the percentage of individuals that have their best brand in their choice set. According our estimates, therefore, it does not appear that a consumer systematically evaluates the utility of all brands even if those brands do not make it into her choice set. The fact that the choice set contains information that is not already captured by the the utility function, in turn, helps to explain why it is crucial to account for choice set heterogeneity in aggregate demand estimation.

Table 11: Percentage of best brand being in consideration set.

	CS-as-RC	Agg-CS-Prob	Ind-CS-Prob
Jacobs	97.49	85.03	85.90
Melitta	93.08	69.92	69.26
Dallmayr	85.05	55.00	55.90
Tchibo	93.42	68.83	72.38
Eduscho	84.44	64.29	66.44

## 6 Conclusions

This paper contributes to the extant literature on aggregate demand estimation by recognizing that consumers choose among a limited set of brands when making their purchase decisions. The fact that choice sets vary across consumers and over time adds another source of heterogeneity to the taste differences that are commonly captured by a random coefficients specification.

To account for the heterogeneity in choice sets, we combine macro data on market shares and prices with micro-level survey data from a tracking study. We find that ignoring choice set heterogeneity biases the parameter estimates of aggregate demand models in several ways. First, estimated brand constants are typically too low. Second, failing to account for choice set heterogeneity causes price elasticities to be underestimated.

We further model and estimate the various processes by which advertising influences sales in an attempt to improve our understanding of the way advertising works and thus enhance our ability to make policy recommendations. Our estimates indicate that, besides contributing to the utility that a consumer derives from a brand, advertising makes it more likely that the consumers considers the brand at the time of purchase.

Finally, we investigate if a consumer's choice set contains her most-highly valued brands, as compensatory models of consideration set formation suggest (Hauser & Wernerfelt 1990, Roberts 1989, Roberts & Lattin 1991). We find that sometimes a substantial percentage of individuals fails to have their best brand in their choice set.

Like any model, our brand choice model is a simplified account of consumer decision making. In particular, we largely abstract from the distinction between awareness and consideration that is often made in the literature on consumer behavior. This literature starts by dividing the available brands into those the consumer is aware of and those she is not aware of. This awareness set is further divided into brands the consumer may consider for purchasing, sometimes called the evoked set or consideration set, and those that are not considered. This suggest that developing and estimating a richer three-stage model of the brand choice process that progresses from awareness over consideration to choice could yield additional interesting insights.

While we have used the tracking study in an aggregate demand estimation setting, its potential uses extend much further. Besides providing us with direct measures of brand salience and choice sets, the tracking study includes attitudinal measures that

capture consumers' perceptions of and attitudes towards the advertising campaigns of the various brands. Enriching the usual input-oriented measures of advertising (expenditures, GRPs) with output-oriented ones may help us to understand to what extent the quality of an advertising campaign matters for sales and profits beyond the sheer quantity of advertising.

In sum, in this paper we show how to combine micro-level survey data on brand salience and choice sets with macro data on aggregate demand. We depart from the commonly made full information assumption and incorporate limited information in the form of choice sets in order to improve the specification, estimation, and interpretation of aggregate demand systems. Moreover, because we combine data on sales and advertising expenditures with tracking data, we are able to separate the effect of advertising on consumer preferences from its effect on choice sets.



## References

- Akerberg, D. (2001). Empirically distinguishing between informative and prestige effects of advertising, *Rand Journal of Economics* **32**(2): 316–333.
- Albuquerque, P. & Bronnenberg, B. (2009). Estimating demand heterogeneity using aggregated data: An application to the frozen pizza category, *Marketing Science* **28**(2): 356–372.
- Allenby, G. & Ginter, J. (1995). The effects of in-store displays and feature advertising on consideration sets, *International Journal of Research in Marketing* **12**: 67–80.
- Andrews, R. & Srinivasan, T. (1995). Studying consideration effects in empirical choice models using scanner panel data, *Journal of Marketing Research* **32**(1): 30–41.
- Bagwell, K. (2005). The economic analysis of advertising, *Working paper*, Columbia University, New York.
- Bajari, P. & Benkard, L. (2005). Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach, *Journal of Political Economy* **113**(6): 1239–1276.
- Berry, S. (1994). Estimating discrete-choice models of product differentiation, *RAND Journal of Economics* **25**: 242–262.
- Berry, S., Levinsohn, J. & Pakes, A. (1995). Automobile prices in market equilibrium, *Econometrica* **63**: 841–890.
- Boyer, M. & Moreaux, M. (1999). Strategic underinvestment in informative advertising: The cases of substitutes and complements, *Canadian Journal of Economics* **32**(3): 654–672.

- Bronnenberg, B. & Vanhonacker, W. (1996). Limited choice sets, local price response, and implied measures of price competition, *Journal of Marketing Research* **33**: 163–173.
- Butters, G. (1977). Equilibrium distributions of sales and advertising prices, *Review of Economic Studies* **44**(3): 465–491.
- Chiang, J., Chib, S. & Narasimhan, C. (1999). Markov chain monte carlo and models of consideration set and parameter heterogeneity, *Journal of Econometrics* **89**: 223–248.
- Clark, R., Doraszelski, U. & Draganska, M. (2009). The effect of advertising on brand awareness and perceived quality: An empirical investigation using panel data, *Quantitative Marketing and Economics* **7**(2): 207–236.
- Doganoglu, T. & Klapper, D. (2006). Goodwill and dynamic advertising strategies, *Quantitative Marketing and Economics* **4**(1): 5–29.
- Dube, J., Hitsch, G. & Manchanda, P. (2005). An empirical model of advertising dynamics, *Quantitative Marketing and Economics* **3**: 107–144.
- Fershtman, C. & Muller, E. (1993). The benefits of being small: Duopolistic competition with market segmentation, *Review of Industrial Organization* **8**: 101–111.
- Gensch, D. (1987). A two-stage disaggregate attribute choice model, *Marketing Science* **6**(3): 223–239.
- Guadagni, P. & Little, J. D. C. (1983). A logit model of brand choice calibrated on scanner data, *Marketing Science* **2**(3): 203–238.
- Hauser, J. & Wernerfelt, B. (1990). An evaluation cost model of consideration sets, *Journal of Consumer Research* **16**: 393–408.

- Horowitz, J. & Louviere, J. (1995). What is the role of consideration sets in choice modeling?, *International Journal of Research in Marketing* **12**: 39–54.
- Hoyer, W. (1984). An examination of consumer decision making for a common repeat purchase product, *Journal of Consumer Research* **11**: 822–829.
- Krishnamurthi, L. & Raj, S. (1991). An empirical analysis of the relationship between brand loyalty and consumer price elasticity, *Marketing Science* **10**(2): 172–183.
- Lebensmittel Zeitung (2006). *Markt- und Strukturzahlen der Nahrungs- und Genussmittelbranche*, LZ Report 2005/2006, Deutscher Fachverlag, Frankfurt am Main, Germany.
- Manski, C. (1977). The structure of random utility models, *Theory and Decision* **8**: 229–254.
- March, J. & Simon, H. (1958). *Organizations*, John Wiley & Sons.
- Mitra, A. (1995). Advertising and the stability of consideration sets over multiple purchase occasions, *International Journal of Research in Marketing* **12**(1): 81 – 94.
- Mitra, A. & Lynch, J. G. (1995). Toward a reconciliation of market power and information theories of advertising effects on price elasticity, *Journal of Consumer Research* **21**: 644–659.
- Narayanan, S. & Manchanda, P. (2008). Heterogeneous learning and the targeting of marketing communication for new products, *Marketing Science* **forthcoming**.
- Nedungadi, P. (1990). Recall and consumer consideration sets: Influencing choice without altering brand evaluations, *Journal of Consumer Research* **17**(3): 263–276.

- Robert, J. & Stahl, D. (1993). Informative price advertising in a sequential search model, *Econometrica* **61**(3): 657–686.
- Roberts, J. (1989). A grounded model of consideration set size and composition, *Advances in Consumer Research* **16**: 749–757.
- Roberts, J. & Lattin, J. (1991). Development and testing of a model of consideration set composition, *Journal of Marketing Research* **28**(4): 429–440.
- Shocker, A., Ben-Akiva, M., Boccara, P. & Nedungadi, P. (1991). Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions, *Marketing Letters* **2**(3): 181–197.
- Siddarth, S., Bucklin, R. & Morrison, D. (1995). Making the cut: Modeling and analyzing choice set restriction in scanner panel data., *Journal of Marketing Research* **32**(3): 255–266.
- Sovinsky Goeree, M. (2008). Limited information and advertising in the U.S. personal computer industry, *Econometrica* **76**(5): 1017–1074.
- Srinivasan, S., Vanhuele, M. & Pauwels, K. (forthcoming). Mindset metrics in market response models: An integrative approach, *Journal of Marketing Research* .
- Stegeman, M. (1991). Advertising in competitive markets, *American Economic Review* **81**(1): 210–223.
- Stigler, G. (1961). The economics of information, *Quarterly Journal of Economics* **69**(3): 213–225.
- Swait, J. & Ben-Akiva, M. (1987). Incorporating random constraints in discrete models of choice set generation, *Transportation Research B* **21B**(2): 91–102.
- van Nierop, E., Bronnenberg, B., Paap, R., Wedel, M. & Franses, P. H. (2010). Retrieving unobserved consideration sets from household panel data, *Journal of Marketing Research* **47**(1): 63–74.