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-60323 Frankfurt am Main
dklapper@wiwi.uni-frankfurt.de

April 26, 2010
track˙v11

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 discrete choice models of demand. In a departure from the commonly made full information assumption, we incorporate limited information in the form of choice sets to reflect the fact that consumers may not be aware of all available brands at purchase time. We derive theoretically and show empirically that both the estimated brand constants and the price coefficient are biased downward when consumer heterogeneity in choice sets is ignored. These biased estimates lead to costly mistakes in firms' price setting.

In addition, the tracking data allow us to identify separately two processes by which advertising influences market shares. We find that advertising has a direct effect on brand awareness (inclusion in choice set) in addition to its effect on consumer preferences (increase in utility). This improved understanding of how advertising works enhances our ability to make policy recommendations.

Key Words: choice set heterogeneity, advertising impact, demand estimation, econometric models.

1 Introduction

A maintained assumption when estimating aggregate discrete choice models of demand in the tradition of Berry, Levinsohn & Pakes (1995) is that all consumers select from the same set of all available brands in a given market. Decades of research in marketing, however, has established that this is not the case and that consumers frequently select from a restricted choice set. There is evidence that even in a store, when faced with the actual choices, people do not evaluate all alternatives (e.g. 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. Choices are thus often limited to a subset of the brands offered and this subset varies across consumers. Ignoring this source of heterogeneity leads to biased estimates of the demand primitives (see, e.g., Bajari & Benkard 2005, Sovinsky Goeree 2008).

Additionally, because there is no recognition of the fact that consumers may not be aware of all brands, the role of advertising in these aggregate discrete choice models is reduced to shifting consumer preferences, without affecting the set of options they select from. Choice sets, however, are not fixed but vary over time and can be influenced by the marketing mix (Allenby & Ginter 1995, Mitra 1995, Siddarth, Bucklin & Morrison 1995). Recognizing that it matters as much *which* brands consumers think about as *what* they think about them, firms advertise to boost sales by persuading consumers of the superiority of their brands and by ensuring that consumers are aware of their brands in the first place. Because awareness is a necessary condition for purchase, choice sets give rise to a second channel through which advertising affects sales in addition to the traditional role of advertising as a factor in shaping consumer preferences.

The difficulty in incorporating choice sets and considering the effects of advertising on the number and identity of brands consumer select from lies mainly in identification. With aggregate data on sales and marketing mix alone, it is impossible to determine without strong assumptions if a brand was not purchased because the consumer did not like it or simply because she was not aware of it. As we discuss in Section 2, the stationarity and independence assumptions typically invoked in the literature, are routinely violated in many empirical settings.

In this paper we bring in a new micro-level data source on brand awareness that facilitates aggregate demand estimation and allows us to dispense with these strong assumptions. Companies frequently collect this type of data in so-called tracking studies to determine the subset of brands consumers recall with or without aid and how consumers perceive the brands that they are aware of.[1] These new individual-level data enable us to contribute to the extant literature in two ways. First, we propose a method to improve the specification and estimation of aggregate discrete choice models by explicitly considering consumer heterogeneity in brand awareness. Second, our data allows us to separately identify the availability effect of advertising on brand awareness (is the brand in the choice set) from its substitution effect on consumer preferences (if a brand is in the choice set, will it be selected) and thus shed light on the dual role of advertising.

Our starting point is a model of aggregate demand that accounts for consumer heterogeneity by modeling their tastes in a random coefficients specification (Berry et al. 1995, Nevo 2000). We depart from the full-information assumption in these models and incorporate limited information in the form of choice sets. Our assumption is that a consumer's choice set is restricted to the set of brands she is aware of. We acknowledge that brand awareness varies across consumers and time and fully account for this potentially equally important source of heterogeneity in the estimation. We illustrate our proposed approach using data from the ground coffee category in Germany, where we combine aggregate sales and price data with individual-level survey data on brand awareness and perceptions.

We first treat brand awareness as a random coefficient and integrate over its distribution in the population of consumers. This approach takes the stated choice sets as data and we do not have to take a stance on the origin of choice sets in order to obtain unbiased estimates of the demand primitives. However, as we are interested in the determinants of choice sets and in particular in the role of advertising, we also 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.

[1] Although companies frequently conduct tracking studies, their use in academic research has to date been limited. Recently Srinivasan, Vanhuele & Pauwels (forthcoming) have fitted VAR time-series models to data from a tracking study in order to analyze the correlation patters in these "mindset metrics" and the sales response to them.

We show analytically and empirically that failing to account for choice set heterogeneity causes the price elasticities to be underestimated. This result 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. Our theoretical derivations and empirical results yield a conclusion opposite to the one advanced in a closely related paper, Sovinsky Goeree (2008), who finds that the bias is towards overestimating price sensitivities. We attribute this different result to the fact that we have data on the sets of brands consumers are aware of and do not rely on functional form to identify them.

By combining data on sales and advertising expenditures with tracking data on brand awareness, we are able to shed light on the dual role of advertising. Our empirical results 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. Hence, if a firm advertises, there are two different effects on demand: (1) advertising increases the willingness to pay of individual consumers, and (2) advertising draws new consumers into the market for that product (Becker & Murphy 1993). This latter channel is quite important, especially in cases where the newly drawn-in consumers differ systematically from a firm's existing customer base (Erdem, Keane & Sun 2008). Directly modeling the effect of advertising on choice sets, may thus provide an explanation for the effect of advertising on consumers' price sensitivity.

2 Background

Choice Set Heterogeneity and Demand Estimation

While recent aggregate demand models in the spirit of Berry et al. (1995) have placed a lot of emphasis on incorporating consumer heterogeneity in tastes through random coefficients, heterogeneity in choice sets is routinely ignored. This is troubling because there is considerable heterogeneity in the size and composition of these choice sets across consumers (see, e.g., Mehta, Rajiv & Srinivasan 2003, Chiang, Chib & Narasimhan 1999). Compounding this problem is the fact that different consumers respond differently to advertising and promotions and firms' spending on these activities varies over time. Thus, the heterogeneity in choice sets increases further in response to the marketing mix.

Erroneously assuming that a consumer is fully informed and chooses among all products in the market when making a purchase decision leads to a misspecified demand system and biased estimates of the demand primitives (see, e.g., Bajari & Benkard 2005, Sovinsky Goeree 2008). The literature has developed two approaches to dealing with this problem. The first approach tries to supplement the aggregate demand data with auxiliary information. For example, Bruno & Vilcassim (2008) consider estimating aggregate discrete choice models when consumers make choices from different choice sets due to differential product availability across stores. Because they do not have data on the actual set of products available at the time of purchase, they simulate the potential assortments that consumers might have faced based on the marginal probabilities, assuming that product availability is independent across brands. In a closely related study, Tenn (2006) proxies for differences in promotions across stores by assuming that only one product is promoted at a given time. Albuquerque & Bronnenberg (2009) combine macro data on sales 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. As the authors point out, their approach raises an initial conditions problem and requires strong stationarity assumptions which may not always hold (e.g., new product introductions). Pancras (2010) similarly tries to create proxies for brand awareness from consumers' purchase histories. Ultimately, this approach to accounting for choice set heterogeneity is limited by the fact that although we can infer that a brand must have been in the consumer's choice set if she bought it, the fact that the consumer did not buy the brand does not imply that the brand was not in her choice set.

The second approach to accounting for choice set heterogeneity is to augment the aggregate demand model with a model of choice set formation. This approach offers the ability to explore the role of advertising in driving brand awareness. The best-known paper in this vein is Sovinsky Goeree (2008), where the probability that a consumer would be aware of a given brand is expressed as a function of her demographics and exposure to advertising. Independence is invoked to compute the joint distribution of awareness across brands from the marginal distributions. The main limitation of this study is related however to a fundamental identification problem: Suppose a consumer buys a product. Then she must have known about it and she must have liked it more than any other product in her choice set. But because we do not know what other brands are in her choice set, it could either be that the utility of the chosen product was very high and she had many products in the choice set, or it could be that the utility of the chosen product was very low and she had just a few products in her choice set. Any attempt to separately estimate consumer preferences and brand awareness without individual-level data on the latter must therefore rely on functional form assumptions to disentangle these two cases.

Because we incorporate micro-level data on brand awareness into aggregate demand estimation, we do not need to impute choice sets from purchase data as in Albuquerque & Bronnenberg (2009) or Bruno & Vilcassim (2008). Neither do we have to make assumptions on the choice set formation process as in Sovinsky Goeree (2008) to account for the heterogeneity of consumers with respect to brand awareness. Instead, our micro-level data on consumers' choice sets enables us to directly obtain unbiased estimates of the demand parameters.[2]

Our theoretical and empirical findings point to the usefulness of supplementing sales and price data with information on the particular choice environments that individual consumers face. Recently Conlon & Mortimer (2010) have used secondary data sources to bound the occurrence of stockouts in vending machines. They argue that variation in choice sets can be used to identify consumers' tastes even if there is no variation in prices. The choice environment can also be easily manipulated and observed in an experimental setting such as conjoint studies (see, e.g., Zeithammer & Lenk 2009). Musalem, Olivares, Bradlow, Terwiesch & Corsten (2010) develop a structural demand model that captures the effect of

[2]In addition to our model-free approach, we advance a model linking awareness to advertising but the model is not required for the bias correction.

out-of-stocks on customer choice by using secondary data on partial product availability. Our paper differs in two respects. First, we are concerned with consumers' brand awareness, not with the physical availability of the product in vending machines or on store shelves. Second, because we have direct information on the brand awareness of individual consumers, we do not have to simulate the choice environment that a consumer might have faced.

Dual Role of Advertising

Understanding whether advertising affects sales through increasing brand awareness or consumer preferences is important as it has implications for the intensity of competition in the market and its profitability. For example, if advertising increases brand 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).

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. The difficulty again lies in identifying the amount of information that is available to consumers. Ackerberg (2001) and Narayanan & Manchanda (2008) proxy for available information with usage experience, but because usage experience is often not directly observable, this approach is largely limited to newly introduced brands. In the more mature market for PCs, Sovinsky Goeree (2008) cannot identify the differential impact of advertising on brand awareness and consumer preferences. Because she does not have data on choice sets she instead assumes that advertising operates entirely through raising awareness. While this may be a justifiable assumption for the PC market where there are a lot of products and purchases are reasoned, it may be less so for other markets. Other studies have relied heavily on functional form assumptions to disentangle the effects of advertising on choice sets and utility (see, e.g., Andrews & Srinivasan 1995, Siddarth et al. 1995, Bronnenberg & Vanhonacker 1996, van Nierop, Bronnenberg, Paap, Wedel & Franses 2010).

Our micro-level data on brand awareness allows us to distinguish between two basic functions of advertising. In particular, we can directly estimate the effect of advertising on awareness by relating a brand's advertising to whether or not this brand is in the set of brands a consumer is aware of. We are thus able to complement existing experimental (Mitra & Lynch 1995) and cross-sectional studies (Clark, Doraszelski & Draganska 2009) that have studied the question whether advertising affects market share through information or persuasion. Using scanner panel data, Erdem et al. (2008) also find for 17 of their 18 brands that advertising leads to a flattening of the demand curve by drawing new consumers into the market for a brand.

The remainder of this paper is organized as follows. Section 3 describes the data sources we use in our empirical application and presents a first look at the data. Sections 4 and 5 develop the theoretical model and our strategy for estimating it. We discuss the results in Section 6 presents the results and conclude with directions for future research in Section 7.

3 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. We focus on the five major national brands, Jacobs, Melitta, Dallmayr, Tchibo, and Eduscho, which together comprise about 70% of the market. There is weekly information on the sales, prices, and promotional activities (in-store communication and features which we aggregate to a single promotion variable for estimation purposes) 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 market shares (inside shares; outside good reported separately), marketing mix, and advertising expenditures. Averages with standard deviations below in parentheses.

	market share	price	promotion	advertising
Jacobs	33.63	6.84	0.24	681.64
	(12.47)	(0.51)	(0.19)	(416.06)
Melitta	21.53	6.30	0.22	806.94
	(12.59)	(0.45)	(0.20)	(430.62)
Dallmayr	13.45	7.52	0.16	725.42
	(8.76)	(0.44)	(0.18)	(237.30)
Tchibo	16.69	7.99	0.17	874.30
	(4.67)	(0.43)	(0.12)	(510.59)
Eduscho	14.70	6.79	0.22	587.49
	(5.59)	(0.40)	(0.14)	(444.96)
outside good	88.34			
	(1.76)			

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 across the weeks of a month using exponential smoothing. 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. 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 λ measures the persistence of goodwill over time. The goodwill of a brand affects the utility that a consumer derives from the brand and it may also affect consumer's brand awareness.

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. As can be expected in a mature product category, aided brand recall is generally high and does not to change much. Similar to previous research (Horowitz & Louviere 1995), we thus focus on top-of-mind awareness as captured by unaided recall in our empirical investigation. The idea is that consumers only include in their choice set brands which are highly salient. All other measures that are provided in the survey data are conceptually closer to consideration rather than awareness, as they include some reference to consumer preferences and are therefore not consistent with our modeling assumption that choice sets are independent thereof.

We calculated the ϕ coefficients between unaided brand recall and possible measures of brand preference such as relevant set, price-independent preference, first choice, second

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

	Jacobs	Melitta	Dallmayr	Tchibo	Eduscho
unaided brand recall	0.8390	0.6489	0.4218	0.6697	0.5897
	(0.3675)	(0.4773)	(0.4939)	(0.4704)	(0.4919)
aided brand recall	0.9719	0.9135	0.9144	0.9042	0.8230
	(0.1653)	(0.2811)	(0.2797)	(0.2944)	(0.3817)
relevant set	0.5353	0.3535	0.3546	0.3545	0.3039
	(0.4988)	(0.4781)	(0.4784)	(0.4784)	(0.4599)
reject set	0.4304	0.3252	0.2486	0.2437	0.2397
	(0.4952)	(0.4685)	(0.4322)	(0.4294)	(0.4269)
price-independent preference	0.4048	0.2478	0.3152	0.2570	0.2398
	(0.4909)	(0.4317)	(0.4646)	(0.4370)	(0.4270)
recent purchase	0.3516	0.1648	0.1762	0.2473	0.1233
	(0.4775)	(0.3710)	(0.3810)	(0.4315)	(0.3288)
purchase in the last 12 months	0.5306	0.3164	0.3238	0.4071	0.2572
	(0.4991)	(0.4651)	(0.4679)	(0.4913)	(0.4371)
first choice	0.2638	0.1077	0.1214	0.1981	0.0875
	(0.4407)	(0.3100)	(0.3266)	(0.3986)	(0.2826)
second choice	0.1584	0.1208	0.0923	0.1610	0.0818
	(0.3651)	(0.3259)	(0.2895)	(0.3675)	(0.2741)

choice, and recent purchase.[3] The ϕ coefficients reported in Table 3 clearly show that unaided recall is only very moderately related to the preference measures.

In addition, the tracking study includes a rich set of demographics and measures of usage behavior for the surveyed consumers. Table 4 provides an overview. Approximately 20% of respondents are from East Germany which has 17 million out of a total of 82 million citizens. Three quarters of the respondents state that they watch TV every day, a number that seems plausible given that the average German citizen watches more than 100 minutes of TV a day. However, more than 40% of respondents report not being very interested or even not watching commercials at all. A slight majority of respondents classify themselves as heavy users of ground coffee, which is consistent with the fact that more than 94% of Germans drink coffee. In our empirical model, we use this information on demographics and usage behavior to allow consumers to systematically vary in their awareness level of the different

[3]The ϕ-coefficient measures the association between two qualitative measures. They may be regarded as the ordinary Pearson correlation between attributes A and B when the categories are associated with scores of zero and one.

Table 3: Correlation between unaided recall and preference measures.

	Jacobs	Melitta	Eduscho	Tchibo	Eduscho
relevant set	0.21	0.29	0.30	0.25	0.26
reject set	0.02	-0.03	-0.11	0.05	0.02
price-independent preference	0.17	0.22	0.23	0.19	0.20
recent purchase	0.20	0.23	0.28	0.24	0.19
purchase in the last 12 months	0.23	0.29	0.31	0.29	0.24
first choice	0.18	0.20	0.25	0.23	0.16
second choice	0.10	0.18	0.12	0.17	0.15

brands.

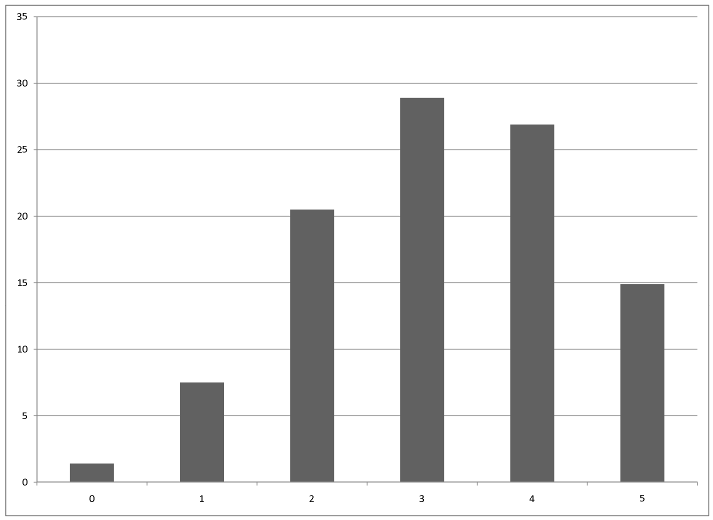
Table 4: Summary statistics for tracking study: Usage behavior and demographics.

	Respondents	Percentage
Region		
West Germany	12046	78.97
East Germany	3208	21.03
Behavior when commercials are aired		
"I watch commercials"	8720	57.17
"I switch TV stations when commercials are on air"	6347	41.61
TV-usage behavior		
never or rarely	3860	25.31
almost every day or every day	11207	73.47
Usage of ground coffee		
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 awareness as captured by unaided brand recall varies across consumers. Figure 1 shows the distribution of choice set sizes (measured by number of brands recalled) 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.

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. 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. Similarly, Bruno & Vilcassim (2008) make an independence assumption because they only have access to a proxy for the marginal distribution of availability.

Our data, however, 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 5. 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 χ^2 distribution with 26 degrees 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 awareness. Overall, this shows that even if data on the marginal distribution of choice set membership for the various brands is available (and it seldom is), then this may not be enough to reliably account for choice set heterogeneity; instead, the entire joint distribution is required.

Table 5: 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	

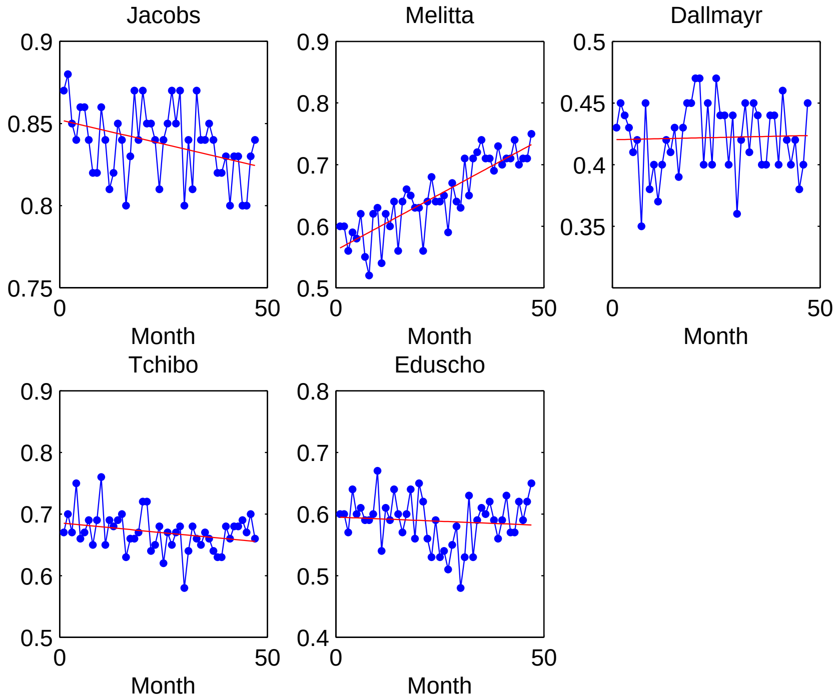
Brand awareness not only varies over consumers but also over time, as evident from Figure 2 where we depict the time path of brand awareness (averaged across consumers) for our five major national brands. The brand awareness of Melitta exhibits a clear upward trend, those of Jacobs and Tchibo trend down. All brands show considerable fluctuations in awareness. The fact that brand awareness varies over time calls into question the stationarity assumption in papers such as Albuquerque & Bronnenberg (2009) and Chiang et al. (1999).

Advertising is a key driver of brand awareness. A reduced-form regression of brand awareness on its share in advertising expenditures (share of voice) yields a significant positive effect for 3 out of 5 brands.[4] This first look at the data confirms previous findings that choice sets are not fixed but can 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. In what follows, we show how to use tracking data to account

[4]Table not included for brevity and is available from the authors upon request.

Figure 2: Evolution of brand awareness over time.



for choice set heterogeneity in aggregate demand estimation and to distinguish between two basic functions of advertising.

4 Modeling Framework

We allow for the possibility that a consumer may not be aware of all available brands and assume that her choice is restricted to the set of brands that she is aware of. We thus model the probability that consumer n chooses brand j in period t as

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

where $\Pr_{nt}(j|\iota_{nt})$ is the probability that the consumer chooses the brand given choice set ι_{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.

Although our model allows us to take the stated choice sets as data, it does not require us to be able to impute choice sets with certainty. By specifying $\Pr_{nt}(\iota_{nt})$ to be the probability that consumer n has choice set ι_{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, Ben-Akiva, Boccara & Nedungadi 1991). Thus, following Manski (1977), we model the process of choice set formation in a probabilistic fashion.

Consumer utility. Let consumers (households) be indexed by $n = 1, \ldots, N$, brands by $j = 0, \ldots, J$,[5] and time periods (purchase occasions) by $t = 1, \ldots, T$. 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 a brand-specific constant that captures any time-invariant characteristics of the brand and promotional activities. p_{jt} denotes the price of brand j in period t and g_{jt} its accumulated goodwill. We consider 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 $\nu_n = (\nu_{n1}, \ldots, \nu_{nK}) \sim N(0, \Sigma)$ her tastes. Together with the idiosyncratic shocks $(\varepsilon_{n0t}, \ldots, \varepsilon_{nJt})$, 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 $(\xi_{1t}, \ldots, \xi_{Jt})$ 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, ξ_{jt} reflects time-varying factors like coupon availability and shelf space allocation

[5] By brand we mean here brand-chain combination

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 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 $\iota_{nt} = (\iota_{n1t}, \ldots, \iota_{nJt}) \in \{0,1\}^J$ 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 ι_{nt}, the probability that consumer n chooses brand j at purchase occasion t is

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

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.

Direction of bias for 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 the 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 leads to flawed pricing decisions. Our example demonstrates that the resulting profit losses can be substantial.

Although 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 that ensure that each individual consumer demands less than she actually does. 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 utility of those consumers who know about the brand is $1 - p + \epsilon$ and assuming extreme-value, iid errors, their demand is given by the logit form $\frac{\exp(1-p)}{1+\exp(1-p)}$. There are ϕ_1 of those consumers. Naturally, the $\phi_0 = 1 - \phi_1$ consumers who do not know about the brand, do not 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 ϕ_1. As can be seen, the full-information assumption leads to underestimation of both the brand value a and the price

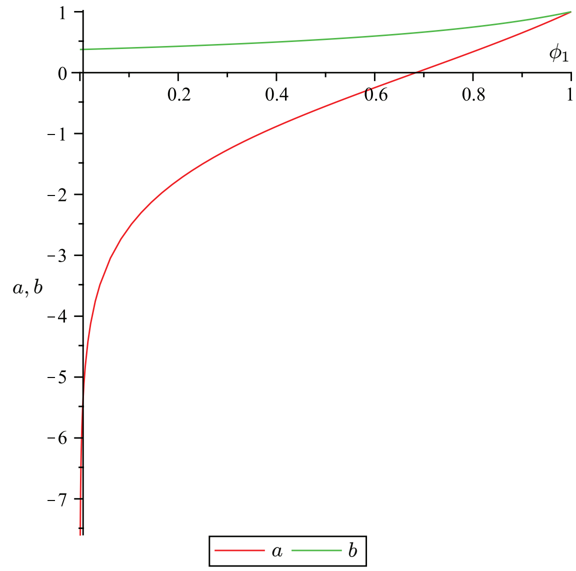


Figure 3: Estimated brand value a and price sensitivity b. True value is 1 for both.

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.

Note that our derivations are completely independent of the utility of those consumers who do not know about the brand. In fact, we never specify the utility of those consumers who do not know about the brand – it could be $1 - p + \epsilon$ or $0 - p + \epsilon$ or anything else. In case of $0 - p + \epsilon$, we have a situation where consumers who like the brand less are also less likely to have it in their choice sets. Yet the bias remains exactly the same.

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. 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 θ is a conduct parameter whose value depends on the specific model of oligopolistic competition. Continuing with our example, a routine calculation shows that at a price of $p = 1$, the full-information demand model implies an elasticity as low as -0.38 (at $\phi_1 = 0$), 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 ϕ_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 of up to 2.62. This happens because in the full-information demand model consumers appear to be less price sensitive than they actually are. Making flawed pricing decisions is detrimental for profitability as Figure 5 illustrates. Therefore, because the full-information assumption fails to recover the correct demand primitives, basing pricing decisions on the full-information model when in truth demand is generated from a limited-information model can be very costly.

5 Empirical Strategy

We first treat the choice set of a consumer as a random coefficient in a model-free approach to incorporating the heterogeneity of brand awareness across consumers. Then, 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), these two ways of incorporating micro-level survey data from the tracking study amount to difference specifications of $\Pr_{nt}(\iota_{nt})$, the probability of consumer n having choice set ι_{nt} in period t.

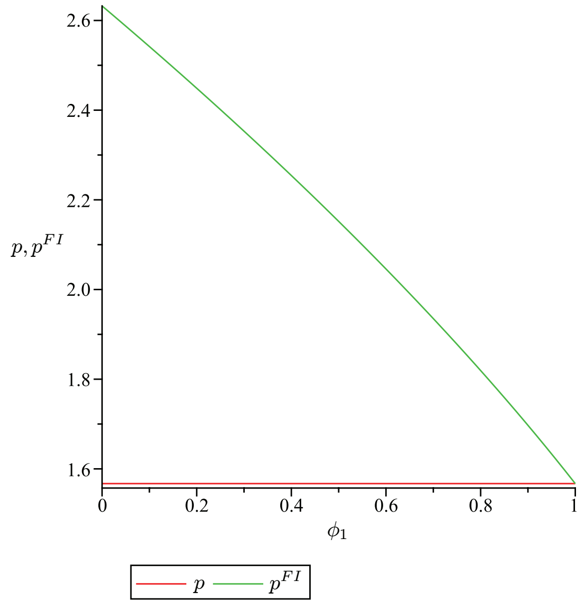


Figure 4: Optimal price for true limited-information demand model (p) and misspecified full-information model (p^{FI}).

Brand awareness as random coefficient. One approach is to think of the choice set of a consumer as just another random coefficient, which 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

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

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 ι_{nt} stated in the tracking study. If one only had data on the awareness levels of individual brands as opposed to their joint distribution as we do, then the method advanced in Bruno & Vilcassim (2008) could be applied by assuming independence.

To compute predicted market shares, we proceed as follows:

1. Draw an individual from the tracking study (demographics D_n and choice set ι_{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).

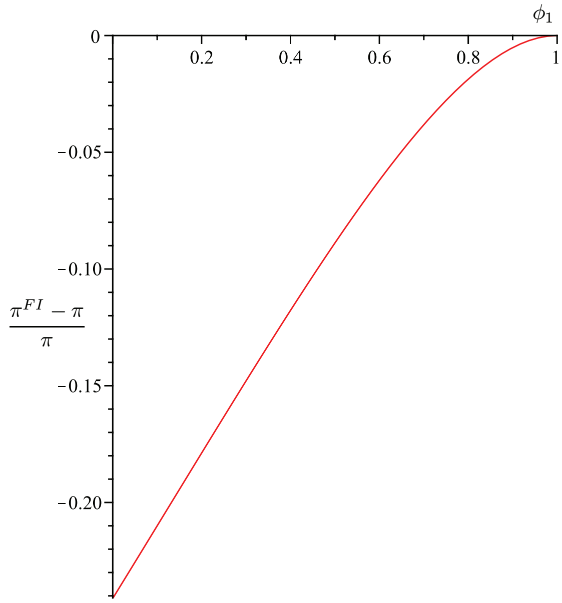


Figure 5: Percentage profit loss from basing pricing decisions on the misspecified full-information model ($\frac{\pi^{FI}-\pi}{\pi}$).

4. Repeat for another individual. Average to obtain the predicted market shares in equation (3).

We use the contraction mapping in Berry (1994) to back out the unobserved demand shocks $(\xi_{1t}, \ldots, \xi_{Jt})$ from the observed market shares $(s_{0t}, s_{1t}, \ldots, s_{jt})$. 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 $(\xi_{1t}, \ldots, \xi_{Jt})$. We use the cost of raw coffee as an instrument for price.[6] The cost of raw coffee is determined in worldwide commodity markets and can thus be taken as exogenous to coffee manufacturers in Germany. We instrument for goodwill with the cost of TV commercials, which again should be independent of the actual advertising behavior of coffee manufacturers.[7] The instruments for price and 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

[6] 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.

[7] 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.

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.

Modeling individual choice set probabilities. To gain insights into the drivers of choice set composition, we next explicitly model the choice set probabilities at the individual level as a function of 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 ι_{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. The market share of brand j in period t is

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

By specifying the choice set probability to be $\Pr_{nt}(\iota_{nt}) = \phi_{\iota_{nt}}(g_t, D_n)$, our model allows for measurement error in the sense that the actual choice sets can differ from the ones stated in the tracking study.

In practice we specify the individual choice set probabilities $\phi_{\iota_{nt}}(g_t, D_n)$ as a multinomial logit model with the 2^J possible choice sets as dependent variable. The index $v_{\iota_{nt}}$ of choice set $\iota_{nt} = (\iota_{n1t}, \dots, \iota_{nJt})$ is given by

$$v_{\iota_{nt}} = \mu_{\iota_{nt}} + (\iota_{nt} \otimes D_n)\,\delta + \left(\sum_{j=1}^{J} \iota_{njt} g_{jt} \right) \gamma + \eta_{\iota_{nt}},$$

where the demographic and usage behavior variables D_n are interacted with brand dummies and the goodwill is summed across all brands in a choice set. We assume that the error terms $\eta_{\iota_{nt}}$ in the choice set probability model are independent (conditional on the observables) of those underlying the market share model (Horowitz & Louviere 1995).

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. (2010).

By allowing the choice set probabilities to be functions of goodwill and demographics, 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. There is some experimental evidence 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 $\Pr_{nt}(\iota_{nt})$, the probability of the consumer having choice set ι_{nt} in period t, and $\Pr_{nt}(j|\iota_{nt})$, the probability that the consumer choose brand j given this choice set, as functions of advertising, we are able to separately examine the availability effect of advertising on brand awareness from the substitution effect on consumer preferences.

We estimate the model in equation (4) in two steps. First, we estimate $\phi_{\iota_{nt}}(g_t, D_n)$ by maximum likelihood using the tracking data. Then we substitute these estimates into equation (4) and 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 (4).

Given the large number of observations, $\phi_{\iota_{nt}}(g_t, D_n)$ is estimated very precisely. Still, we also calculated the standard errors via bootstrap in order to account for possible effects of the prediction error in the first stage on the parameter estimates in the second stage.

6 Estimation Results

In what follows we illustrate the empirical strategies developed in Section 5 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 standard 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 4.

Table 6 displays the coefficient estimates and Tables 8 and 7 display various measures derived from these estimates.[8] In each table, the first column is the standard aggregate demand model assuming full information. The remaining columns correspond to the two ways we suggest for incorporating limited information in the form of choice sets.

Baseline. In the full-information model we obtain reasonable estimates for the parameters of the demand system. Price, goodwill and promotions all have significant effects in the expected direction. 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.[9] The price elasticities implied by our estimates are reasonable and range between -4.12 for Melitta to -4.91 for Tchibo, see the first column of Table 7. These price sensitivities are very comparable to what other researchers have found in the ground coffee category (Guadagni & Little 1983, Krishnamurthi & Raj 1991).

Because buying the same brand at a different retail chain may provide for a different

[8] We conducted a grid search to determine the persistence parameter λ 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.

[9] Adding demographics to the random coefficients specification did not improve the estimates.

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

	Full information		Limited information			
			CS-as-RC		Ind-CS-Prob	
Price	-0.605*	(0.038)	-0.635*	(0.048)	-0.632*	(0.036)
Promotion	0.377*	(0.126)	0.421*	(0.110)	0.527*	(0.118)
Goodwill	0.220*	(0.074)	0.190*	(0.060)	0.177*	(0.049)
Trend	-0.059*	(0.009)	-0.047*	(0.008)	-0.071*	(0.010)
RC Price	0.007	(0.103)	0.077*	(0.027)	0.010	(0.086)
RC Promotion	0.892*	(0.145)	0.994*	(0.133)	0.892*	(0.169)
RC Goodwill	0.041	(0.078)	0.054*	(0.024)	0.001	(0.048)
RC Trend	0.063*	(0.018)	0.059*	(0.027)	0.080*	(0.018)
Jacobs/Edeka	-4.901*	(1.306)	-4.124*	(0.789)	-3.547*	(0.640)
Melitta/Edeka	-5.256*	(0.927)	-4.401*	(0.761)	-3.902*	(0.639)
Dallmayr/Edeka	-4.538*	(0.988)	-3.300*	(0.712)	-2.883*	(0.869)
Tchibo/Edeka	-4.542*	(0.711)	-3.343*	(0.766)	-4.063*	(1.128)
Eduscho/Edeka	-4.731*	(1.111)	-3.974*	(0.847)	-3.313*	(0.633)
Jacobs/Markant	-4.943*	(1.302)	-4.169*	(0.788)	-3.599*	(0.635)
Melitta/Markant	-5.039*	(0.921)	-4.200*	(0.762)	-3.689*	(0.636)
Dallmayr/Markant	-4.695*	(0.985)	-3.453*	(0.711)	-3.034*	(0.866)
Tchibo/Markant	-4.571*	(0.710)	-3.374*	(0.765)	-4.094*	(1.127)
Eduscho/Markant	-4.529*	(1.109)	-3.776*	(0.846)	-3.110*	(0.631)
Jacobs/Metro	-5.001*	(1.305)	-4.250*	(0.793)	-3.637*	(0.640)
Melitta/Metro	-5.029*	(0.941)	-4.203*	(0.769)	-3.644*	(0.653)
Dallmayr/Metro	-4.941*	(0.999)	-3.719*	(0.719)	-3.245*	(0.856)
Tchibo/Metro	-4.909*	(0.710)	-3.715*	(0.765)	-4.434*	(1.128)
Eduscho/Metro	-4.431*	(1.109)	-3.679*	(0.846)	-3.010*	(0.631)
Jacobs/Rewe	-4.892*	(1.296)	-4.128*	(0.791)	-3.541*	(0.634)
Melitta/Rewe	-4.862*	(0.935)	-4.017*	(0.763)	-3.495*	(0.643)
Dallmayr/Rewe	-4.902*	(0.994)	-3.679*	(0.719)	-3.212*	(0.855)
Tchibo/Rewe	-4.154*	(0.713)	-2.983*	(0.771)	-3.680*	(1.127)
Eduscho/Rewe	-4.416*	(1.109)	-3.659*	(0.847)	-3.001*	(0.633)
Jacobs/Spar	-4.965*	(1.305)	-4.205*	(0.791)	-3.597*	(0.640)
Melitta/Spar	-5.046*	(0.939)	-4.200*	(0.765)	-3.666*	(0.648)
Dallmayr/Spar	-4.293*	(1.009)	-3.077*	(0.725)	-2.585*	(0.862)
Tchibo/Spar	-4.392*	(0.710)	-3.196*	(0.766)	-3.912*	(1.126)
Eduscho/Spar	-4.707*	(1.109)	-3.952*	(0.847)	-3.282*	(0.633)
Jacobs/Tengelmann	-4.617*	(1.293)	-3.859*	(0.791)	-3.264*	(0.631)
Melitta/Tengelmann	-4.181*	(0.925)	-3.330*	(0.759)	-2.813*	(0.642)
Dallmayr/Tengelmann	-4.148*	(0.987)	-2.919*	(0.714)	-2.466*	(0.856)
Tchibo/Tengelmann	-4.706*	(0.713)	-3.520*	(0.769)	-4.225*	(1.127)
Eduscho/Tengelmann	-4.886*	(1.114)	-4.134*	(0.846)	-3.462*	(0.633)
RC Jacobs	1.528*	(0.782)	1.234*	(0.368)	1.174*	(0.448)
RC Melitta	0.428	(1.168)	0.211	(0.542)	0.225	(1.762)
RC Dallmayr	0.340	(0.975)	0.124	(0.323)	0.532	(1.055)
RC Tchibo	1.693*	(0.596)	1.274*	(0.257)	2.567*	(0.589)
RC Eduscho	0.334	(1.378)	0.708	(0.475)	0.321	(0.884)

Table 7: Price elasticities.

	Full information	Limited information	
		CS-as-RC	Ind-CS-Prob
Jacobs	-4.199	-4.372	-4.428
Melitta	-4.117	-4.310	-4.311
Dallmayr	-4.910	-5.200	-5.158
Tchibo	-4.757	-5.093	-4.990
Eduscho	-4.220	-4.446	-4.404

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 8, the average value of the brand across all retail chains is -4.886 for Jacobs, -4.902 for Melitta, -4.586 for Dallmayr, -4.546 for Tchibo, and -4.617 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. However, the estimated random coefficients for the brand values show that there is considerable heterogeneity in brand valuation across consumers.

Table 8: Mean brand values.

	Full information	Limited information	
		CS-as-RC	Ind-CS-Prob
Jacobs	-4.886	-4.122	-3.531
Melitta	-4.902	-4.058	-3.535
Dallmayr	-4.586	-3.358	-2.904
Tchibo	-4.546	-3.355	-4.068
Eduscho	-4.617	-3.862	-3.196

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 8 to the remaining columns, brand values increase as choice set heterogeneity is accounted for. The reason is that the full-information model must explain the observation that a consumer did not buy a brand by assigning that brand a particularly low utility. In contrast, the limited-information models can rationalize this observation without lowering the value of the brand if the consumer is not aware of the brand. The difference in the estimated brand values shows clearly that some consumers do not buy a product not because they do not like it but because they do not know it.

Biased price sensitivities. Next, as can already be seen from Table 6, the price coefficient changes as choice set heterogeneity is accounted for. As shown in in Section 4, wrongly assuming full information biases the estimated price coefficient downward (i.e., it is less negative). Table 6 empirically confirms our derivations for the direction of the bias in price sensitivities.

Comparing the first column of Table 7 to the remaining columns shows that ignoring choice set heterogeneity leads to incorrect price elasticities. There are two reasons for this. First, the price coefficient changes as choice set heterogeneity is accounted for. Second, the functional form for aggregate demand changes. To see this, suppose there were no random coefficients. Then the full-information model reduces to a logit model and, consequently, suffers from the independence of irrelevant alternatives property. The limited-information models, in contrast, are mixture models and therefore do not suffer from this property. Hence, price elasticities necessarily change.

In line with our theoretical derivations in Section 4, the estimated price elasticities in Table 7 are systematically understated in the full-information model. This finding is in contrast with Sovinsky Goeree (2008) who argues 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 4, 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 quantity response to a price change in the data, the full-information model has to elicit a smaller response from each individual consumer than the limited-information model. Hence, the full-information model can be expected to systematically understate the price sensitivity of consumers. In line with our results, Mehta et al. (2003) find using a search model of consideration set formation that price elasticities are much larger when consideration sets are accounted for (e.g., an elasticity of Wisk of -2.057 versus -1.360 in a full-information model).

Table 9: Per period profit loss from ignoring choice set restrictions. Limited Information (LI) versus Full Information (FI).

		Jacobs	Melitta	Dallmayr	Tchibo	Eduscho
Markup	LI	2.305	1.871	1.837	2.874	1.831
	FI	2.531	1.994	1.934	3.207	1.937
Price	LI	7.384	6.821	8.145	8.363	7.207
	FI	7.705	7.024	8.304	8.773	7.386
Share	LI	3.363	1.973	1.147	2.281	1.605
	FI	2.506	1.207	0.463	1.298	0.906
Profit	LI	12628	5630	3433	8417	4646
	FI	10350	3700	1469	5373	2781

Table 9 illustrates the consequences of wrongly assuming full information for profitability. To do this, we first use the limited information demand system (individual choice probabilities model, equation (4)) and an assumption on firm conduct (Manufacturer Stackelberg) to recover firms' marginal cost. Then we recompute the price equilibrium under the full-information demand system. Because the full-information model understates price sensitivities, firms choose higher prices and markups than appropriate. This pricing policy entails substantial profit losses, again in line with our simple example in Section 4. For example, if Jacobs were setting equilibrium prices under the incorrect full information assumption the lower price coefficient would lead to an almost 10% higher markup than in the limited information case (2.531 instead of 2.305). This would translate into 5% higher prices at the retailer (7.705 instead of 7.384) and 25% lower shares (3.363 to 2.506). Per period manufacturer profits would decline by 22% from 12628 to 10350.

Table 10: Estimation results for individual choice set probabilities model. Standard errors in parentheses. Choice set constants omitted for brevity.

goodwill	0.449* (0.054)				
	Jacobs	Melitta	Dallmayr	Tchibo	Eduscho
region	0.936*	0.152*	0.289*	-0.595*	-0.717*
(1=east, 0=west)	(0.065)	(0.041)	(0.039)	(0.044)	(0.044)
age	-0.003	-0.003*	0.012*	0.007*	0.002
	(0.002)	(0.001))	(0.001)	(0.002)	(0.002)
ad viewing	0.119*	0.166*	0.093*	-0.034	0.095*
(1=watch commercials, 0=no)	(0.042)	(0.032)	(0.031)	(0.037)	(0.036)
TV frequency	-0.086*	-0.136*	0.026	0.096*	-0.056
(1=rarely, 0=daily)	(0.047)	(0.037)	(0.036)	(0.044)	(0.042)
coffee consumption	0.077*	0.145*	0.055*	-0.079*	0.093*
(1=heavy user, 0=light user)	(0.043)	(0.032)	(0.032)	(0.037)	(0.036)

Choice set drivers: 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 10 indicate what makes it more or less likely that a brand is included in the choice set.

The estimates 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 likely related to the period prior to the reunification of Germany, when 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 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 and less likely to be aware of the more expensive brand Tchibo.

Choice set drivers: Advertising. Importantly, as Table 10 illustrates, goodwill has a significant positive effect on the choice set composition. 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.

To further investigate the separate effect of advertising on awareness and consumer brand preferences, we calculate market shares based on the estimated demand parameters and observed prices, advertising and marketing-mix and then vary the amount of advertising and the channel through which it operates. As Table 12 shows, the impact of advertising on brand awareness can be just as important as its impact on utility, lending support to the informative view of advertising (see also Erdem et al. 2008). For Dallmayr, Tchibo, and Eduscho the effect on awareness is even larger than the effect on utility. Moreover, these two effects compound each other in the overall impact of advertising on market shares. Our finding that the availability effect is as important than the substitution effect is in line with previous empirical work on advertising. For example, Mitra & Lynch (1995) and Clark et al. (2009) show that advertising can have a much stronger effect on brand awareness than on the relative strength of consumer preferences.

Table 11: Advertising elasticities.

	Full information	Limited information	
		CS-as-RC	Ind-CS-Prob
Jacobs	0.0111	0.0095	0.0091
Melitta	0.0117	0.0100	0.0094
Dallmayr	0.0118	0.0102	0.0095
Tchibo	0.0108	0.0095	0.0088
Eduscho	0.0113	0.0097	0.0091

At first glance, the effect of advertising on market share may seem small, despite being well in line with the reported elasticities in the literature (compare also the elasticities in Table 11). For example, Sethuraman & Tellis (1991) report an average of 0.1 over 262 studies but also that "several advertising elasticities ... are close to zero (on the order of 0.001)". Even if small in magnitude, the effect is economically significant. A simultaneous increase in advertising budgets of all brands by 50% increases the category market share by 1.5 percentage points. Considering that the total market for ground coffee at the time of our study was approximately DEM 8.2 billion, a 1.5 percentage point increase in share corresponds to approximately DEM 120 million increase in revenue for all firms.

Table 12 also reveals that the effect of advertising via brand awareness and consumer preferences varies across brands. Dallmayr and Eduscho – two brands with a relatively low level of awareness – exhibit the strongest effect. On the other hand, Jacobs, which has an average unaided brand recall of more than 80%, has the lowest relative impact of an increase in advertising on share. A 50 percent increase in advertising would increase the share of Dallmayr by 19.3 percent whereas the same increase in advertising budget by Jacobs would shift its market share by only 10.7 percent. In this respect unaided recall seems to be an appropriate predictor for the success of intensified advertising. If brand awareness is already high, then the effect of advertising comes predominately through utility. However brands that have a lower awareness level can benefit from both channels and get a higher advertising response for their investment.

A further noteworthy phenomenon is that substantial increases in advertising expenditures of all firms simultaneously do not change the ranking of market shares relative to the outside good but shifts the market shares within the group of the five brands. Therefore, some brands can gain at the expense of competitors. This becomes obvious if we focus on the within groups shares before and after an increase in advertising. A brand like Jacobs with high brand awareness cannot benefit (would actually lose 3%) as opposed to a lesser-known brand such as Dallmayr, which could improve its within group market share by 5%. Interestingly, Dallmayr is one of the biggest advertisers. Obviously their commercials do not make it into consumers mind and in this respect their advertisements seems to be inefficient; brand awareness relative to advertising expenditures is lowest among our brands. On the other hand, Eduscho, a brand with moderate awareness and relatively low expenditures could benefit a lot from increased advertising because it would not only increase the market

Table 12: Channels of advertising impact. Change in overall market shares (relative to outside good) and within market shares due to simultaneous increase in advertising expenditures.

		base	utility	choice set	total
Overall Share	Jacobs	3.631	3.839	3.725	4.020
	Melitta	2.225	2.362	2.356	2.506
	Dallmayr	1.287	1.365	1.444	1.535
	Tchibo	2.305	2.418	2.491	2.612
	Eduscho	1.651	1.754	1.816	1.935
Total		11.099	11.737	11.831	12.609
Within Share	Jacobs	32.712	32.704	31.483	31.885
	Melitta	20.049	20.124	19.917	19.877
	Dallmayr	11.598	11.625	12.203	12.177
	Tchibo	20.770	20.601	21.051	20.716
	Eduscho	14.872	14.946	15.347	15.345
Total		100.00	100.00	100.00	100.00

share but also the within-group market share. Big advertisers like Melitta and Tchibo with average awareness levels would benefit from increased market share if they were to advertise more but would not be able to improve their position relative to their competitors.

In sum, our advertising results show that brands with relatively low brand awareness may benefit the most from an increase in advertising. They can gain market share and also improve their relative market share within the group. However this recommendation needs to be reevaluated depending on the firms' absolute spending levels. If advertising expenditures are already high, a firm may first consider changing advertising quality (e.g., change or modify the campaign) in order to increase advertising efficiency.

Choice set and utility function. 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. We investigate if it is

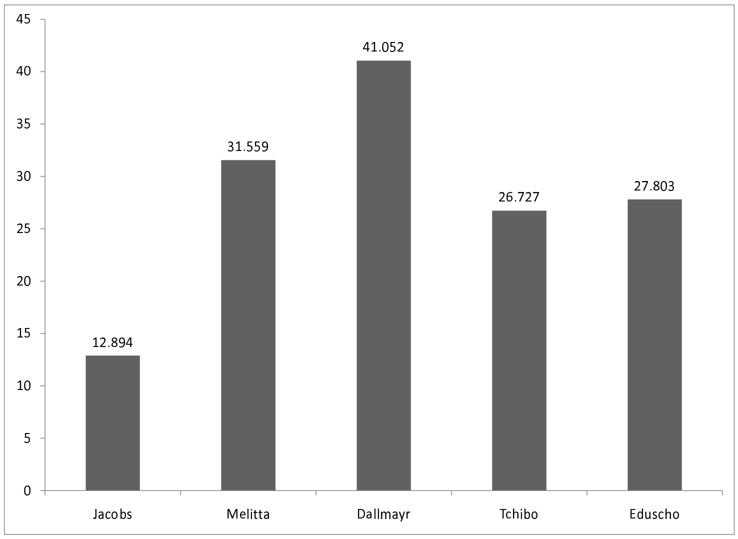


Figure 6: Percentage of best brand not included in choice set.

really a consumer's most-highly valued brands that make it into her choice set as assumed by these compensatory approaches to choice set formation. To this end, we compute the systematic part of an individual's utility for all brands using our coefficient estimates from the individual choice probabilities model (i.e., we do not include the logit errors, reflecting the idiosyncratic tastes). 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 as Figure 6 shows. According to our estimates, therefore, consumers systematically evaluate only the utility of the brands that make it into her choice set. The fact that the choice set contains information that is not already captured by the utility function, in turn, helps to explain why it is crucial to account for choice set heterogeneity in aggregate demand estimation.

7 Conclusions

This paper contributes to the extant literature on aggregate demand estimation by recognizing that consumers often choose among a limited set of brands when making their purchase decisions. Choice sets that vary across consumers and over time add another source of heterogeneity to the demand model, in addition to the taste differences that are typically captured by a random coefficients specification. Limited information in the form of choice sets calls into question the commonly made full-information assumption that all consumers select from the same set of all available brands in a given market.

To capture the heterogeneity in choice sets, we combine macro data on market shares and prices with micro-level survey data from a tracking study. This latter data provides us with direct information on the brand awareness of individual consumers. We find that ignoring choice set heterogeneity biases the parameter estimates of aggregate demand models in several ways. First, estimated brand values are typically too low. After accounting for choice sets, consumers are seen to have stronger brand preferences. This finding 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, ignoring choice set heterogeneity causes consumer price sensitivity to be underestimated. These biased estimates can lead to costly mistakes in firms' price setting

The additional data from the tracking study also allows us to model and estimate two processes by which advertising influences market shares without relying on overly stringent functional form assumptions or hard-to-generalize identification strategies. 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. Indeed, the impact of advertising on brand awareness can be just as important as its impact on utility. Besides lending further support to the informative view of advertising, separately quantifying the impact of advertising on brand awareness and brand preferences improves our understanding of the way advertising works and enhances our ability to make policy recommendations.

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. Because these low awareness/high utility consumers may be relatively easy to convert into sales if the firm could only reach them, our finding is potentially important for designing advertising strategies. Our model of choice set formation links demographics with brand awareness and hence can potentially be used by the firm to identify this group of low awareness/high utility consumers. If so, then the firm may be able to pursue a much more targeted and effective communication strategy.

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. The awareness set is further divided into brands the consumer may consider for purchase, sometimes called the evoked set or consideration set, and those that are not considered. Thus, developing and estimating a richer model of the brand choice process that progresses from awareness over consideration to choice could yield additional interesting insights.

In sum, in this paper we show how to combine micro-level survey data on brand awareness 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 brand awareness.

References

Ackerberg, 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.

Becker, G. & Murphy, K. (1993). A simple theory of advertising as a good or bad, *Quarterly Journal of Economics* **108**(4): 941–964.

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.

Bronnenberg, B. & Vanhonacker, W. (1996). Limited choice sets, local price response, adn implied measures of price competition, *Journal of Marketing Research* **33**: 163–173.

Bruno, H. & Vilcassim, N. (2008). Structural demand estimation with varying product avalailability, *Marketing Science* **27**(6): 1126–1131.

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.

Conlon, C. & Mortimer, J. (2010). Demand estimation under incomplete product availabilits, *Technical report*, Harvard University.

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.

Erdem, T., Keane, M. & Sun, B. (2008). The impact of advertising on consumer price sensitivity in experience goods markets, *Quantitative Marketing and Economics* **6**: 139–176.

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 Genussmittlebranche*, 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.

Mehta, N., Rajiv, S. & Srinivasan, K. (2003). Price uncertainity and consumer search: A structural model of consideration set formation, *Marketing Science* **22**(1): 58–84.

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.

Musalem, A., Olivares, M., Bradlow, E., Terwiesch, C. & Corsten, D. (2010). Structural estimation of the effect of out-of-stocks, *Management Science* .

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 consiedarion sets: Influencing choice without altering brand evaluations, *Journal of Consumer Research* **17**(3): 263–276.

Nevo, A. (2000). Mergers with differentiated products: The case of the ready-to-eat cereal industry, *RAND Journal of Economics* **31**(3): 395–421.

Pancras, J. (2010). A framework to determine the value of consumer consideration set information for firm pricing strategies, *Computational Economics* .

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.

Sethuraman, R. & Tellis, G. (1991). An analysis of the tradeoff between advertising and price discounting, *Journal of Marketing Research* **28**(2): 160–174.

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* .

Tenn, S. (2006). Avoiding aggregation bias in demand estimation: A multivariate promotional disaggregation approach, *Quantitative Marketing and Economics* **4**: 383–405.

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.

Zeithammer, R. & Lenk, P. (2009). Statistical benefits of choices from subsets, *Journal of Marketing Research* **46**(4): 816–831.