Where has Price Endogeneity in Consumer Prices gone?

A Discrete Choice Analysis of the Market for Detergents with Daily Data on Purchases and Advertising

Christoph Nagel *

December 18, 2008

INCOMPLETE DRAFT - COMMENTS WELCOME

Abstract

This work studies five causes for the well-known price endogeneity bias in brand choice and demand estimation models pioneered by Berry (1994). Previous work tackled the problem by suggesting price endogeneity corrections, but never accounted for all of the possible causes for the endogeneity. I propose to treat all five causes simultaneously. Using a novel German nationwide panel dataset that comprises daily consumer purchase and TV viewing history, I estimate discrete choice models augmented by additional information provided from the dataset. The results suggest that the traditional endogeneity corrections only work well in restrictive models, that cannot be detailed further due to crude data and cannot address all five causes simultaneously. However, if we have sufficiently precise data and can control for the five causes, the endogeneity corrections are rendered unnecessary. This result implies that endogeneity corrections are only necessary in the case of incomplete data, that does not allow to estimate models that address all five causes. Furthermore, TV advertising does not affect the consumer purchase decision in the short run, but seems to be fully dominated by long run brand fixed effects.

KEYWORDS: advertising, brand choice, discrete choice, price elasticity, endogeneity correction

*Universität Mannheim, Economics Department, L7, 3-5, 68131 Mannheim, Germany; contact: cnagel@gmx.de

1 Introduction

Within differentiated product markets the existence of diverse product characteristics is necessary to motivate the presence of abundant varieties. For a market setting the empirical IO literature starting from Berry (1994) has emphasized the importance of an endogeneity correction for prices if unobservable characteristics (to the econometrician) are present, when only market outcome data is available to estimate demand, see Berry, Levinsohn, and Pakes (1995). This method has been introduced and applied to the marketing literature by Chintagunta, Dubé, and Goh (2005) who focus on individual consumers and naturally use individual level data. Petrin and Train (2006) use an alternative approach that works for any aggregation level of data, i.e. either market or individual level. Note that all papers treat product price as the only endogenous variable.

Theoretically all of them are based on McFadden's (1974) probabilistic discrete choice model, but result methodologically in different estimation techniques and importantly in different economic interpretations of endogeneity for prices. In the literature several explanations are common to motivate existence and relevance of the endogeneity problem of prices. The "Berry" techniques lead to statistically and economically significant different results. This usually means that the absolute value of estimates for price coefficients and elasticities are larger with the correction than without it.

There are several causes that can lead to endogenous prices in the case of individual level data, where I will focus on the major five ones: (i) "variety characteristics", i.e. time constant unobserved product characteristics for brand varieties within brands (ii) "marketing activity", i.e. time varying unobserved product characteristics at the retail outlet (iii) individual household level inventories (iv) TV advertising per brand at the household level (v) "data aggregation", usually aggregation to weekly observations.

Papers like Chintagunta, Dubé, and Goh (2005) and Petrin and Train (2006) take care of points (i)-(ii), where they argue that a brand fixed effect will solve problem (i) and the proposed endogeneity correction will solve problem (ii). Problem (iii)-(v) is not being addressed explicitly.

Interestingly, doing the data aggregation and the endogeneity correction captures the rest of problem (i) that is not resolved by using brand fixed effects, a point I will return to in section 2. I will argue that problems (i)-(v) are not correctly addressed in previous work.

Firstly, I will argue intuitively what happens in problem (i). The papers mentioned above ignore brand varieties (apart from price and size), and concentrate on brand choice. Assume that a model estimates a brand intercept in which all varieties of a brand are lumped together, so only price and size information remain to distinguish varieties of a brand. Furthermore, assume that there is no one-to-one mapping from price-size-combinations to variety characteristics, i.e. one cannot distinguish varieties of a brand when seeing solely price-size-combinations. Now assume that consumers do not always buy the same variety within each brand. Then, if the consumer buys two different varieties of a brand, the model, and thus the econometrician, will miss unobserved variety characteristics. This possibly systematic variation may be correlated with prices, and thereby lead to biased estimates of price coefficients. In addition, it is common practice to concentrate only on major brands, thus ignore parts of the market and products, that may share some of the same variety characteristics. With this non-random omission, another source of bias may be present, as the choice set the consumer faces is incomplete and biased towards market share leaders.

Secondly, problem (v) arises, as purchase data is aggregated to yield weekly observations. This prohibits intraweek variation and does not allow more than a single unit purchase per week for the simple discrete choice models to work. In the data I observe both. [1]

Thirdly, even if problem (ii) is solved using the usual endogeneity corrections cited above, it is important to understand that economically the source of the endogeneity is thought to be unobserved retail or marketing activity. Examples are TV advertising as in problem (iv) or, alternatively, if the whole sample is restricted to a narrow geographical area as in previous papers, this is problematic because unobserved retail activity is then shared by virtually all agents in the sample. None of the two previous papers use other marketing information I have available, such as TV ads. Obviously, TV ad activity can be correlated with prices, as the manufacturer typically sets advertising levels and wholesale prices for his products.

Finally, problem (iii), not accounting for individual level household inventories, can lead to serious errors for estimating price effects as has been shown by Hendel and Nevo (2006) for a simple storable consumer good: detergents.

In this work I want to study how additional information may solve problems (i)-(v) for individual level data. Due to availability of a novel, very detailed dataset on consumer purchases of detergents I am able to take an alternative approach. The dataset from A.C. Nielsen comprises daily detergent purchases and TV viewing history of nationwide sampled German Households across different retail chains. I have precise variety information for the products sold. A wide array of demographic variables is available for

[1] Of course, this in in principle doable in a discrete choice model, but (i) it may lead to a huge choice set, if pairs and triples of brand need to be considered and (ii) the interpretation of the choice objects from (i) is not clear.

all households in the sample. The advertising data is available for a fraction of the national sample and is detailed down to the individual viewing contact of the household with a specific advertising spot. The motive of the spot is known, so that besides brand other spot characteristics are available, such as length or theme of the spot.

I can address each issue in the following way: (i) variety information is available and is included in the estimation to distinguish the choice of different products and brands. Perhaps a consumers cares more about the "color" effect than about the brand. (ii) retail activity is controlled by feature and display variables and by adding appropriate dummies for each chain. Different from previous work I have a nationwide sample, local effects are not relevant for the whole sample of consumers as in previous work. (iii) duration since last purchase is used to control for individual household inventories. (iv) the effect of TV advertising is controlled for by using the TV viewing history for adverts per brand, as this category is frequently advertised on TV and usually campaigns for new product placements are run. (v) data from actual purchases is not aggregated, but kept on a daily basis in the data.[2]

Conducting all of these steps with my dataset and comparing to the price endogeneity correction of Petrin and Train (2006), I find that the endogeneity correction has no effects for the price coefficients. It remains to precisely trace out which if the leading cause, or if the interplay of them is necessary. Relating to the overdispersion results of Chintagunta, Dubé, and Goh (2005) I find, that the price endogeneity correction does not reduce the standard deviations of the random coefficients on price. This finding is in line with the arguments laid out in Horsky, Misra, and Nelson (2006) that find a reduction once stated-preference data is available.

The paper proceeds as follows. After a following literature review, in section 2 the empirical model is specified and the relevant causes for price endogeneity are laid out. Section 3 describes the data and conducts simple analysis. Then in section 4, the results are presented. The final section 5 concludes.

1.1 Related Literature

I am of course not the first to look at the role of TV advertising for the purchase decision in a consumer goods category. But so far the focus has not been on its relation to the price endogeneity problem for individual level data. The marketing literature has been active on the empirical side to assess the impact of advertising on prices and price sensitivity. See the paper of

[2]There are rarely multiple purchases per day. I exclude these days from the estimation.

Kaul and Wittink (1995) for a survey of the results in this literature.

A very similar dataset has been used by Ackerberg (2001, 2003) and has a very good data quality. His work is not concerned with price endogeneity, but with the prestigious and informational aspects of advertising. Erdem and Keane (1996) estimates models similar to my work, but has to use a dataset, where the advertising information is known to be of a lower quality than in this paper. Also her focus lies not assessing the impact of price endogeneity corrections. The two papers dealing with these issues by Chintagunta, Dubé, and Goh (2005) and Petrin and Train (2006) do not have national samples and advertising information as we have. Keane (1997) does a very detailed analysis on the interplay of consumer taste heterogeneity and state dependence using also data without advertising. I will later discuss the existing specialties of each. Depending on the focus of the work these differences are more or less important.

This work zooms in on detergents, the product category that has also been studied in a recent paper by Hendel and Nevo (2006). They quantify the influence of individual inventory holdings on demand analysis, as detergents constitute a storable product category. They find large corrections of their procedure for price elasticities, i.e. state that usual models overestimate these elasticities, the opposite direction of usual endogeneity corrections. Different from their multi-stage approach, I will use a pure and simple discrete choice specification and try to control for inventory holdings.

In this paper the inclusion of advertising will lead to a effect that is in line with Horsky, Misra, and Nelson (2006). The inclusion will reduce the need for the model to capture the consumer heterogeneity randomly. Advertising levels will vary per consumer, per product and over time, thus introducing some more observed heterogeneity that needs not be accommodated by the parametric distribution of the random coefficients. This helps the model to fit the data without leading to overdispersion in the random coefficients, an effect that is also achieved by having any additional information on the consumer level, such as stated-preference information like in Horsky, Misra, and Nelson (2006). **[to be completed...]**

2 The Model

Before I come to the estimated models, I want to illustrate the level of detail in my data and then discuss the relevant sources of endogeneity for prices. A consumer i, precisely speaking a representative of a household, goes shopping for detergents in a known store s in a known location l on a given day t to buy a specific variety v of a brand j. It is known which TV adverts the household had contact to anytime it is in the sample. The price consumer i faces at the purchase time t is thus:

$$p_{j_v lst} \tag{1}$$

where $j = 1, \dots J$ denotes brand, j_v denotes variety v of brand j and $v \in V_j$, which is the set of all varieties of brand j. l denotes location, which is given as geographic code, i.e. postal code or county (formal: "Landkreis"), and s denotes the store/chain. In the appendix section A some of the classical important models in unified notation are summarized.

2.1 Sources of Price Endogeneity

In this section I want to lay out the structure in which consumers face their choices. Given this structure, I then discuss consumer and industry specific effects that may lead to endogeneity of prices in the individual consumer choice problem. The literature on price endogeneity in IO has emphasized that usually the positive correlation between demand shocks/unobserved product attributes (typically "quality" is the example) and prices will lead to an attenuation bias of the price coefficient, meaning that the effect of price is understated without endogeneity correction. This is mainly motivated by thinking in a market setting, where the researcher only has access to market level data, whereas I focus on the individual level due to access to the necessary data.

2.1.1 Industry Structure

I assume there is a dichotomy between manufacturers and retailers. Thus, the price from equation 1 can be decomposed to:

$$p_{j_v lst} = p^W_{j_v lst} + p^R_{j_v lst} \tag{2}$$

Manufacturers, i.e. producers of detergent, are assumed to set wholesale prices p^W and plan/run TV advertising campaigns. Retailers set markup p^R and have discretionary power to do in-store advertising, typically captured by the feature and display variables in scanner data.[3] Retail chains are assumed to operate independently of each other. Note that this decomposition

[3]This assumption may not hold for several reasons. Firstly, in the case of product introductions, as then all marketing instruments are likely to be used in a concerted action. Secondly, for larger retail chains, manufacturers tightly interfere in the retailers' domain, in the sense that manufacturers buy more shelf store from the retailer.

is merely for expositional purposes, thus writing it in a linear form is no restriction. The price components p^W and p^R are unobserved.[4]

2.1.2 Advertising

Tv advertising is taken to be exogenous. Advertising should not influence whether agent goes to the store, but should work on a different level. Assume there are two decisions the consumer takes: first is whether to buy in the category and second what brand to choose. Advertising can influence the consumer in the first stage in the store as do all other covariates, i.e. prices and promotions, too. Then it can work at the second stage, where the consumer chooses a brand. I observe all on air-adverts from detergent products. Perhaps the aggregate level of advertising influences him at the first stage, and only own brand advertising is relevant for the brand choice stage. Referring to Bagwell (2005) or Lauga (2008) advertising enters the consumption decision in many ways. A crude concept distinguishes informational versus prestigious advertising, the point of interest in the work of Ackerberg (2001, 2003). Nevo (2001, p. 326 footnote 24) finds that once the usual endogeneity corrections are done, effects of advertising for price sensitivities are negligible. I want to stress that this could be due to the nature of aggregation in the model, as the author does not have individual level consumer data.

As TV advertising is set by the manufacturer, it is correlated with prices, at least through wholesale prices p^W. It may be uncorrelated to other demand side factors, that still contaminate prices, as the remaining marketing variables operate on retailer level and are not directly influenced by the manufacturer.[5]

Under these assumptions, the mere availability of advertising alleviates part of the price endogeneity problem.

2.1.3 Direction of Advertising Bias

Tv advertising is assumed to be set by the manufacturer, and is likely to be set simultaneously and according to some unobserved decision calculus. Consider the following two thought experiments to understand the interplay

[4] Note that in Germany prices between retailer and manufacturer are directly negotiated and unobservable to the researcher. In the US there are firms such as *Promodata Leemis Services* that collect this information for the US. Chintagunta, Dubé, and Goh (2005) use this information as instrument for retail prices p.

[5] These are usually measured as feature and display conditions at time of purchase in scanner datasets.

of advertising and prices.[6]

Firstly, suppose manufacturers set prices and TV advertising such that we would observe a positive correlation in the number of advertising contacts and price level variables over time. Then, omitting advertising, this introduces attenuation bias. This is the kind of bias that the conventional price endogeneity correction à la Berry should fix. Does this correction alleviate the problem of omitting advertising fully, or will it still understate the effect of prices? If so, adding the observed advertising information should increase the price effect.

This case of positive correlation may arise as firms try to increase their prestige or brand recognition to create local market power as in monopolistic competition, so that they can price discriminate and charge higher prices.

Secondly, suppose firms set prices and TV advertising such that we would observe a negative correlation in the number of advertising contacts and price level. This would cause an amplification bias. Then, omitting advertising, and doing the usual price endogeneity correction we may in total overstate the price effect. Instead, if we account for the observed advertising data and do the correction, the price coefficient will be reduced again, perhaps may even overcompensate the effect of the conventional price endogeneity correction. In this case of negative correlation, firms may inform the consumer by advertising about pricing events, such as sales or certain rebates or jubilees.

The first case appeals more to what one would intuitively expect the role of TV advertising to be, whereas the second case seems to match well the in-store advertising captured by the usual measured feature or display variables. The latter two variables have been already used in most studies. This informal argument would lead to a positive correlation of TV advertising and prices, whereas one would expect a negative correlation between in-store advertising and prices. I will pursue this thought in the basic analysis part of section 3.[still do this!!!]

Note that for our discussion TV advertising is treated as exogenous, although it is obviously set by firms along with prices. With advertising added as additional endogenous variable, finding instruments for both endogenous variables will be will be more challenging.

2.1.4 Retail Activity

Assuming a strict cut between manufacturers and retailers in term of in-store activities to influence consumers and attribute all of it to the retailer is a questionable assumption. Its validity may depend on retailer type. Big chains

[6]Doing this argument the intuition from a linear model carries through if only one endogenous variable is present.

with large subsidiaries are more likely to cooperate than chains with smaller average store size. We will use this argument later for the construction of price instruments. Due to that I want to control for retail chain effects, as we have also variation in the stores where customers shop. The primary example where retail activity is manufacturer driven are promotions, where this will usually take place in larger subsidiaries. These are not recorded in the data, but usually are quite rare in German stores and also for the detergent category. Then what remains is usual retail activity as measured by the feature and display variables and this may be correlated through the markup p^R with consumers prices p.

2.1.5 Product Characteristics and Variety Information

Detergent is a differentiated product that is in its functioning very homogenous, in the sense that any of them will remove basic stains and from clothes and "refresh" clothing. But products will have different scent compositions (i.e. summer, april,...), come in different consistency (i.e. liquid, powder, tablet, gel), have different optimal purpose of use than basic cleaning (i.e. for color, wool, silk or black clothing). Generally, the price of a product will then not only depend on own characteristics, but also on the characteristics of the nearest neighbor. As Nevo (2001) argues, the markup a firm can set depends on the position of the nearest neighbor which in turn depends on characteristics. Thus, not only taking own product characteristics, but also other product characteristics as exogenous or predetermined and therefore contemporaneously uncorrelated to prices is a strong assumption, but is usually done in the literature. There is a crucial difference compared to the literature that works on market level. There the assumption is necessary to get estimates. In my individual level case it is necessary to get parameter estimates (the βs), but price sensitivities/elasticities can be calculated without the exogeneity assumption on brands. See Nevo (2001) for further explanation. So if we include product characteristics into the estimation, we should not face problems.

Here a crucial point in the usual literature arises that leads to price endogeneity: The focus is only on part of a product category, choices per consumer are aggregated to a week and all product varieties are lumped into one time constant brand coefficient. Suppose price of variety of brand is defined as follows:

$$p_{j_v lst} = p_{jlst} + \nu_{j_v lst} \tag{3}$$

where ν is discount or a markup for variety v in the brand portfolio j. Suppose now a consumer buys on two occasions the same brand, but in a different variety, say "sensitive" and "color" and prices may differ. If all varieties cost the same, $\nu \equiv 0\,\forall v$. If not, then the model attributes any advantage of the chosen variety over the other available choices to the price variable, as the variety information was not included in the standard procedure of the literature. I add the variety information, as almost each brand has a "liquid", "concentrated", "color", et cetera version to avoid this situation. For the argument, it does not matter with what component of equation 2 ν is correlated: it may correlated with p^W because production costs differ or with p^R because the retailer reshifts his range of detergents sold and thereby sells a certain kind with discount.

2.1.6 Individual Inventories

Individual inventory holdings (at the consumer level) are usually only proxied or even ignored in most models, naturally in those that operate on market level data. If proxied, this is performed by adding a variable that captures the duration from last purchase (so called interpurchase time). Whereas in Guadagni and Little (1983) it enters in the utility for each brand, approaches with no-purchase option usually let the only the no-purchase option's utility depend on the duration. Of course, a discrete choice model is not an explicit dynamic model of inventory holding as the modern approaches such as Hendel and Nevo (2006). In the models, I shall vary the way the control enters to see how sensitive the specification is subject to a change. The longer ago a purchase is, the lesser is the price disturbing effect of individual inventories. Then the purchase may be done on the basis of a actual demand (price plays its role in determining brand choice) and not due to a stockpiling motive on the side of the consumer. Interpurchase times may be normalized using the usual median interpurchase time of the household. Following a more formal argument we expect a positive sign on the duration coefficient. Let $P(j)$ be the probability of chosing good j, Inv is the level of individual inventories and Dur be the duration since the last purchase. Then we get:

$$\frac{\partial P(j)}{\partial Inv} < 0, \forall j = 1, \ldots, J \tag{4}$$

$$\frac{\partial Inv}{\partial Dur} < 0, \forall j = 1, \ldots, J \tag{5}$$

$$\Rightarrow \frac{\partial P(j)}{\partial Dur} = \frac{\partial P(j)}{\partial Inv}\frac{\partial Inv}{\partial Dur} > 0, \forall j = 1, \ldots, J \tag{6}$$

$$\tag{7}$$

At the moment the specification is differently implemented. Duration is set to zero forall no-purchase trips, and duration takes in value for all brand purchases. This is the analogue formulation to the explanations and above and so we must expect a negative coefficient on duration.

2.1.7 Consumer Heterogeneity

The heterogeneity of consumers is captured in two different ways. The linear index structure of the DC model will have in most models random coefficients, that capture the unobservable portion of heterogeneity in a parametric way. The parameters will be assumed to follow a standard normal distribution. As demographic variables are available, these are added to the specification and will partly be interacted with economically relevant variables as price and advertising. There is no option for any survey data as in Horsky, Misra, and Nelson (2006) to improve the heterogeneity part of the models by adding stated preference information.

2.2 Instrumental Variables Choice

To find potential instruments in our setting with potentially two endogenous variables we have to discuss the relation of price and advertising as seen in the section beforehand across several dimensions. First, I only consider prices to be an endogenous variable. The literature sees unobserved local retail activity as the leading cause, thus an instrument should at least be valid in this case.

Instruments of the Hausman (1997) type may be a first good choice.[7] Also Nevo (2001) (on page 309 top, 320 bottom) uses essentially the same instruments for his estimation to be valid. For potential weaknesses of Hausman Instruments see Nevo (2001), page 321. The instruments are as follows. For prices, instruments are price for a similar product in another outlet, as all underlying manufacturer specific characteristics are the same, whereas the difference is the retail outlet activity and the agreement between manufacturer and chain. This retail activity is the candidate explanation that Chintagunta, Dubé, and Goh (2005) have for introducing the unobserved brand characteristic in the panel data setting. Such a Hausman instrument would eliminate it. In my case I will restrict myself to the 15 biggest chains, for the analysis and use the mean prices of the remaining 160 chains in the same week, for a "similar" product, in the same geographic region (identified by postal code or county "Landkreis") as instrument.[8] Chintagunta, Dubé,

[7] See Petrin and Train (2006) p.22. Another instrumental variable approach supposed by Petrin are instruments in the manner of Pakes(1994).

[8] "Similar" means that brand, approximate packaging size and consistency match.

and Goh (2005) cannot do this, as they focus on one big market and do not use data from another big market that they have.

Petrin and Train (2006) favors wholesale prices, but informal discussion with market participants show there are only direct unobserved negotiations between retailer and manufacturer in Germany, so that no "wholesale" price exists for a consumer good like detergents as in the US.

Other instruments could be aggregate adverts by a manufacturer, that are correlated with price but not with individual retail activities. Then the adverts per consumer (the aggregate adverts, but filtered, i the adverts a consumer sees) could be used to see the impact of adverts on consumers, while aggregate adverts control for unobserved retail activity.[**to be completed...**]

An alternative avenue to find instruments is the "unobservable instruments" idea from Matzkin (2004) suggested to me by Richard Blundell. There is random variation between total advertising and advertising received by individual consumer. Not seeing a certain spot is random, as TV choice decision is not based on the choice of advertising, but on the TV programme. This may be however correlated to the advertising level, so that it is not an instrument for the ads by a manufacturer. But it is unrelated to the unobserved retail activity which following Chintagunta, Dubé, and Goh (2005) is the primary source of endogeneity in the case of individual panel data.[**to be completed...**]

To find instruments if both prices and advertising are treated endogenous is still an outstanding challenge.

2.3 Models under Consideration

The models used are standard logit and random coefficient logit models. I discuss the different specifications that I will pursue.

2.3.1 The Empirical Model

The first model is motivated by Guadagni and Little (1983), who introduced a logit model without random coefficients used extensively in marketing applications. It is nested within the specification of the standard random coefficients model that is described in the following. Removing the i subscript on the parameters (i.e. α, β, γ) gives the first model.

Each consumer i at time t derives utility from a choice among one brand $j = 1, \ldots, J$ or a no-purchase $J + 1$ according to the following utility form.

$$U_{ijt} = \beta^1_{ij} + \beta^2_i X^A_{ijt} + \beta^3 X^S_{ijt} + \beta^4 X^V_{ijt} + \alpha_i p_{ijt} + \gamma_i a_{ijt} + \epsilon_{ijt} \quad (8)$$

$$U_{i(J+1)t} = \epsilon_{i(J+1)t} \quad (9)$$

where X^A contains marketing variables, X^S contains state dependence variables, X^V contains product characteristics, the "variety information", p is price, a is TV advertising and ϵ is an iid extreme value error, $\beta^i, i = 1, 2, 3, 4$ and especially α_i, γ_i contain the parameters of interest. The subscript for location l and s disappear, as the individual i can only be at one place at a time. So the corresponding values of prices at l, s are selected and thereby I save some exuberant notation. Note also that X^S is constructed using observations from other time periods (at least form the one before), so that we have a dynamic model. Dynamic is to be understood in the sense of a reduced form approach, as there is no unique behavioral explanation given the way dynamics are introduced. I employ as specification for state dependence the exponentially smoothed weighted average of past purchases suggested by Guadagni and Little (1983), also used as base case in Keane (1997). The k-th component of the vector X^S_{ijt} is defined by:

$$\begin{aligned} X^{Sk}_{ijt} &= \phi_k X^{Sk}_{ijt-1} + (1-\phi_k) d^k_{ijt-1} && (10) \\ d^1_{ijt} &= \mathbf{1}\{\text{purchase of } j \text{ by } i \text{ at time } t\} && (11) \\ d^2_{ijt} &= \mathbf{1}\{\text{ad of } j \text{ by } i \text{ at time } t\} && (12) \end{aligned}$$

This adds parameters $\phi_k, k = 1, 2$. The model features only state dependence where past purchases and ads of a brand matter, but dependence on product sizes of past purchases is excluded. Both variables enter in a form of exponentially weighted average of past values. Setting the ϕs to zero gives only first-order dependence. According to Guadagni and Little (1983) there are two options for setting ϕ. Calibrating ϕ or taking it as another parameter in ML estimation.[9] Keane (1997) includes ϕ as parameter in his joint estimation procedure. He initializes $X^{Sk}_{ijt} = 0\,\forall i, j, k$ and $t = 0$. To stay more comparable to Chintagunta, Dubé, and Goh (2005) we set $\phi = 0$ and drop the $k = 2$ part of equation 10 **[to be completed...]**.

The random coefficients of the model parameters in the utility function are defined by:

$$\begin{pmatrix} \alpha_i \\ \beta^1_{i1} \\ \vdots \\ \beta^1_{iJ} \\ \beta^2_i \\ \beta^3_i \\ \beta^4_i \\ \gamma_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta^1_j \\ \vdots \\ \beta^1_J \\ \beta^2 \\ \beta^3 \\ \beta^4 \\ \gamma \end{pmatrix} + \begin{pmatrix} \kappa_\alpha X^D_{i,\alpha} \\ \kappa_{\beta^1_j} X^D_{i,\beta^1_j} \\ \vdots \\ \kappa_{\beta^1_J} X^D_{i,\beta^1_J} \\ \kappa_{\beta^2} X^D_{i,\beta^2} \\ \kappa_{\beta^3} X^D_{i,\beta^3} \\ \kappa_{\beta^4} X^D_{i,\beta^4} \\ \kappa_\gamma X^D_{i,\gamma} \end{pmatrix} + \Sigma^{\frac{1}{2}} \eta \qquad (13)$$

[9] From calibration Guadagni and Little (1983) get $\phi = 0.875$.

where $\eta \sim N(0, I_m)$ and $\Sigma^{\frac{1}{2}}$ is a Cholesky decomposition of a variance covariance matrix fitting the dimension of the random vector η. Simply assume $\Sigma = diag(\sigma_1^2, \ldots, \sigma_m^2)$. $X_{\cdot}^D$ is a vector of demographic variables and $\kappa_{\cdot}$ are parameters, that together allow as deterministic part of the random coefficients for observed heterogeneity in the form of component linear indices per random coefficient. Written like this, the number of demographics that enter each random coefficient can vary. We denote the vector of these functions as κ to shorten notation. This summarizes the form of the random coefficients as sum of a linear index on demographics and a random normal variable. In the first round I set $\kappa = 0$ and let only price have a random normal distribution.

Choice probabilities for consumer i picking brand j at time t take under the above assumptions the usual logit form conditional on realizations of η, that determine the stochastic part of the random coefficients. That stochastic part and the deterministic part of consumer heterogeneity are summarized in equations 16 and 17. Let $\theta = (\alpha, \beta_j^1, \ldots, \beta_j^1, \beta^2, \beta^3, \beta^4, \gamma, \kappa_{\cdot}, \Sigma, \phi)$ denote all parameters to be estimated. Thus, the individual choice probability for consumer i to pick brand j at time t takes the following form, conditional on the realization of η:

$$
\begin{aligned}
P_{ijt|\eta}(\theta) &= \frac{exp(\varpi_{ijt} + \nu_{ijt}^D + \nu_{ijt}^S)}{1 + \sum_{k=1}^{J} exp(\varpi_{ijt} + \nu_{ijt}^D + \nu_{ijt}^S)} && (14)\\
\varpi_{ijt} &= \beta_j^1 + \beta^2 X_{ijt}^A + \beta^3 X_{ijt}^S + \beta^4 X_{ijt}^V + \alpha p_{ijt} + \gamma a_{ijt} && (15)\\
\nu_{ijt}^D &= [1_j, X_{ijt}^A, X_{ijt}^S, X_{ijt}^V, p_{ijt}, a_{ijt}]\kappa && (16)\\
\nu_{ijt}^S &= [1_j, X_{ijt}^A, X_{ijt}^S, X_{ijt}^V, p_{ijt}, a_{ijt}]\Sigma^{\frac{1}{2}}\eta_i && (17)
\end{aligned}
$$

We have one price coefficient, as we do not state that promotional price reduction are different than long term price changes.[10] A possibility to capture this effect is to add a variable indicating extreme price changes, for example a dummy if prices cross a certain threshold.

Note that different to usual models estimated, all variables can vary across consumer, time and good. In the usual literature, the data come from one market, so that marketing variables do not vary anymore across i, TV advertising is unobserved and "variety information" is omitted so that

$$X_{ijt}^A = X_{jt}^A, p_{ijt} = p_{jt}, a_{ijt} = 0, X_{ijt}^V = 0 \forall i$$

[10]See the paper by Briesch, Chintagunta, and Matzkin (2002) that uses a nonseparable nonparametric function for price and promotional price changes to analyze the consumer reaction.

. Not having this simplification does not matter for the simple models 1 and 2, that are estimated as usual. It will make a difference for the models that conduct an endogeneity correction.

The endogeneity in a Berry market level model is usually unobserved product quality. If we assume that it is constant over time, we can control for it by including a constant. In model 1 and 2 it is $\beta_j^1, j = 1, \ldots, J$ that captures this effect. Note that model 2 is more general, in that it allows fluctuations on consumer level per brand. In model 1, the taste is fixed to the mean, whereas it is not in model 2. In Nevo (2001) no dummy is added for the time varying effect of brand over time (there ξ_{jt}). Instead it is subsumed into a composite error term with the iid extreme value error ϵ_{ijt}. Nevo had no other choice as given his data level he would have had too many parameters.

Now, as the individual choice probabilities are given, we can specify the probability that a consumer did a specific sequence of product choices conditional on η:

$$L_{i|\eta}(\theta) = \prod_{t=1}^{T_i} \prod_{j=1}^{J+1} P_{ijt|\eta}^{y_{ijt}}(\theta) \tag{18}$$

where y_{ijt} is a dummy, defined by $y_{ijt} = \mathbf{1}\{\text{purchase of } j \text{ by } i \text{ at time } t\}$. Then, we can specify the unconditional likelihood function for a purchase sequence of a household:

$$L_i(\theta) = \int L_{i|\eta}(\theta) f(\eta) d\eta \tag{19}$$

where $f(\eta)$ is the standard normal density.

Then the function to maximize given a sample of I consumers is:

$$LL(\theta) = \sum_{i=1}^{I} log[L_i(\theta)] \tag{20}$$

Estimation of this expression is conducted by Simulated Maximum Likelihood. See Train (2003) for details. For this, the integration is approximated by simulation methods. Conditional on parameters and a draw of η, the inner part of the integral for a household is calculated, this is done several times and the results are averaged to approximate the integral.

As I have a panel of consumer this is the standard panel RC Model, where β is not allowed to vary across each drawn observation of each consumer, but only one draw of β per each consumer's choice sequence is taken. In the estimation, I will estimate also the cross sectional version, where β is allowed to vary over consumers and time periods and compare both versions.[**to be completed...**]

2.3.2 Endogeneity Correction Model

The proposed correction follows closely the control function approach of Petrin and Train (2006). Recall equation 8, where now, if we make the unobservable visible as ξ_{ijt}:

$$U_{ijt} = \beta^1_{ij} + \beta^2_i X^A_{ijt} + \beta^3 X^S_{ijt} + \beta^4 X^V_{ijt} + \alpha_i p_{ijt} + \gamma_i a_{ijt} + \xi_{ijt} + \epsilon_{ijt} \quad (21)$$

$$U_{i(J+1)t} = \epsilon_{i(J+1)t} \quad (22)$$

Petrin and Train (in the following PT) set $a_{ijt} = X^S_{ijt} = X^V_{ijt} = 0$ in their application. For their approach, we need to define the control function equation. Using controls recovered from the control function equation in the above utility specification alleviates the endogeneity problem. See the appendix for their case, that is adapted to using the same data as Chintagunta, Dubé, and Goh. I see two obvious ways for defining the control function. The first alternative is:

$$p_{ijt} = E[p_{ij}|z_{ijt}] + \hat{\xi}_{ijt}, \quad j = 1, \ldots, 4, i = 1, \ldots, I \quad (23)$$

$$z_{ijt} = [1, p^\star_{ijt}] \quad (24)$$

Estimating the control function across t, I would only use the time variation per consumer to recover a per individual per product unobservable. As instruments PT use wholesale prices $p^\star_{ijt}$. From this then we get very "individual" unobservables that vary even per consumer and one would have $\hat{\xi}_{ijt}$ instead of $\hat{\xi}_{jt}$. In this setting, we can in principle "see" heterogeneity in the unobservables. Instead averaging over i gives an average that fits directly into the usual estimation routine of PT. The second alternative is:

$$p_{ijt} = E[p_j|z_{ijt}] + \hat{\xi}_{ijt}, \quad j = 1, \ldots, 4 \quad (25)$$

$$z_{ijt} = [1, p^\star_{ijt}] \quad (26)$$

Conduct a pooled regression over i and t. Then the same argument as in case one applies.

In my application I use the constructed instrument outlined in section 2.[**to be completed...**]

2.4 Model Discussion [incomplete]

3 Data and Basic Analysis

3.1 Data

The data I employ is an extensive househould-level panel collected from AC Nielsen, Germany. It is referred to under the name "Singlesource" and provides household, daily purchase and real-time media information over a period of 2 years from mid 2004 through mid 2006. The name Singlesource stems from the fact that the same household are followed just as in usual scanner datasets, but in addition, the real-time TV advertising history is recorded as well. [11] Companies like the GfK (Gesellschaft für Konsumforschung AG, Nürnberg, Germany) do not supply these data based on the same households. The dataset is collected nationwide throughout Germany. The datasets consists of two components: a household panel where purchases are followed and a subsample of the former where additionally all TV advertising contacts are recorded. As the data consist of several collected datasets from A.C. Nielsen, sample sizes of available information vary across topics of interest. See table 1 and 2 for details on the files and numbers of households for which relevant information is available.

File	Household (HH) appears if	Description
Cash	sampled	total value of purchases with time, store, plz
Wash	a purchase of detergent	detergent purchases with time, store, plz and product details (price, amount, characteristics)
demo	sampled	available socio-demographic variables
TV	with TV telemeter	TV representation factors

Table 1: Files that contain consumer data

The product categories under consideration are detergents and chocolate. For detergents, Erdem and Keane (1996) argue that (a) detergents are "fre-

[11]See the reviews about the A.C. Nielsen data collection procedure for Singlesource authored from various sampled consumers at the Website of Ciao GmbH in 2007.

Dataset	Criterion	No of Households
Cash	any purchase	16757
Cash	any purchase in "detergent" store	16737
Cash	above plus demographics	16737
Wash	any purchase of detergent	13455
Wash	above plus demographics	13455
Wash	TV coverage 2004	2953
Wash	TV coverage 2005	2630
Wash	TV coverage 2006	2571
Wash	TV coverage 04/05	2250
Wash	TV coverage 05/06	1993
Wash	TV coverage 04-06	1735

Table 2: Different No. of Households across datasets

	units	Mean	Median	Std.	Min	10%	90%	Max
Demograhics								
Income	HH	2566	2000	1930	750	1250	3500	10000
Urbanity	HH	5.6	6	1.55	1	3	7	7
No. persons	HH	2.43	2	1.20	1	1	5	10
No. kids	HH	0.53	0	0.88	0	0	2	8
MA Age	HH	47	45	14.8	18	28	68	93
MM Female	HH	0.76	1	0.43	0	0	1	1
self emp.	HH	0.09	0	0.28	0	0	0	1
white-collar h	HH	0.46	0	0.50	0	0	1	1
white-collar l	HH	0.23	0	0.42	0	0	1	1
blue-collar	HH	0.21	0	0.41	0	0	1	1
perm unemp	HH	0.01	0	0.1	0	0	0	1
Duration in sample	HH	320	278	255	0	$0^\star$	674	730
Duration in sample	HH	392	396	227	1	84	683	730
Detergents								
Price	PI	3.71	2.99	2.30	0.01	2.19	6.66	32.99
Size	PI							
Quantity	PI	1.11	1	0.44	1	1	1	16
Duration	PI	53	34	64.2	0	0	126	723
Duration1	HH	81.7	60.7	74.7	0	22.7	162	723
Duration2	HH	91.6	69.4	77.4	1	29.4	177	723
No. Brands bought	HH	2.4	2	1.7	1	1	5	14
Brand HHI (value)	HH	0.69	0.68	0.29	0.10	0.31	1	1
Brand HHI (volume)	HH	0.69	0.66	0.29	0.09	0.29	1	1
Store visits								
No. Stores visited	HH	2.34	2	1.49	1	1	4	13
Store HHI (value)	HH	0.69	0.64	0.28	0.07	0.32	1	1

Table 3: Summary Statistics of Household-level Data
$\star$: one time purchases, occur 2247 times

quently and regularly purchased products"; (b) brands are frequently introduced; (c) "firms heavily advertise in this category"; (d) detergents are "low in variety seeking". In addition to this, I guess detergents have an objective functionality that is to clean clothes which is required by many consumers. Detergents are of course storable products, which means that inventory considerations may be present. Chocolate is more an addictive product and does not provide a basic function as does detergents do. Advertising may have therefore very different effects on both categories. [**to be completed...**]

I observe daily visits to supermarkets and their amounts spent at each visit for two product categories: chocolate bars and detergent. Additionally, I know aggregate amounts spent per visit. I know the exact brand-size combinations bought, at which quantity and price. As usual, feature and display variables, are contained in the data as well. See table 3 for some descriptive statistics.[**to be completed...**]

The TV advertising variable contained in the data needs some explanation. It is possible to see if a household tuned to a channel with a product specific advert, we have the recorded time and channel and the topic/motive of the spot. We have to assume that this spot has an effect on the household, or the member that does the purchases, as we do not know which person of the household was tuned in to the programme.

3.2 Relation to other datasets

There are several datasets used in the literature, where each one has his comparative advantages for his application. Most important for the current paper are the differences in the quality and quantity of pricing and advertisement information. The data of Hendel and Nevo (2006) tracks households that purchase in one store and a complete store level dataset is available to deliver us all prices during purchase decisions of the consumer in a simple fashion, but it lacks the TV advertisement information of the dataset in this paper. Concerning prices, my data is thus of the same kind as Keane (1997), where he has only a few regional markets in the US in his dataset. He has to impute all data. I deal with this issue in a similar fashion to him as explained in the next section. Erdem and Keane (1996) have a dataset that daily tracks households in two stores from 1986 to 1988. In principle it has the same TV advertising data, but only for 51 weeks and for only 1800 households. Also it lacks the precise clip information we have to identify image and product specific clips. The product category of interest is also laundry detergents.[**to be completed...**]

3.3 Inferring Prices

There are several schemes to infer prices. I follow Erdem and Keane (1996) and Keane (1997). Firstly, our number of households will make it possible for us to construct a good alternative by filling in the gaps with consumers that shopped in the same store at the same time. Usually, all data is aggregated to weeks. Since we have daily data and if we are willing to assume that prices are constant in a week for a shop, we can use price information from other days within a week for the same store to get prices for most purchase occasions.

Secondly, a subgroup of retailers have nationwide targets, so that filling the gaps can be done by using the information from other branches of that retailer.

I could perform several steps to see how much of our sample is then filled up for estimation, but up to now I restrict myself only to the first step but line out also possible later ones. The reason is that I want to avoid measurement errors in prices.

1. Collect all purchases per brand in a calendar week in a chain in a "Landkreis" area. Fill up missing prices within this group.

2. If this fails, widen the time interval from weekly to two adjacent weeks and use the average price of the weeks based on the days with sales.

3. For the nationwide chains with national price targets, collect all purchases per brand on a week in a chain. Fill up prices.

4. If this fails, widen the time interval from one week to the two adjacent weeks.

Looking at the number of prices per postal code ("PLZ"), I find that the price information is too scarce to do the imputation across PLZs. There are about 15K PLZ codes and 434 counties, i.e. "Landkreise". The fewer number of Landkreise make it much more likely to find two households in the same area than when we use PLZ codes.

Different from the US, in Germany the practice of issuing price coupons that reduce the retail price in store are not common, so that I do not need to correct the collected prices. Keane (1997) notes that existence of the US coupon-redemption system leads to exaggerated price elasticities from models that do not account for this.**[to be completed...]**

3.4 Basic Analysis

[to be completed...]

3.5 Estimation sample

To get results that are free from outliers or data artifacts, we select households/ observations for the estimation of the models from section 2 according to the following criteria: (a) in the choice situation, the consumers choice set had at least 4 elements. (b) all shopping trips, that resulted into no detergent purchase and had a total shopping volume below 5EUR are dropped. (c) all shopping trips, that resulted into no detergent purchase are dropped, which were on the same day as another shopping trip with a detergent purchase (d) all shopping trips, that resulted into detergent purchases are dropped, which were on the same day as other shopping trips with a detergent purchase (e) all households that had no detergent purchases were dropped.

As not all information is available for all the households, we must estimate some of the specification for different samples. We estimated models along the following 5 samples. Due to missing values in some variables, the effective sample sizes in the estimations sample can differ within each of the five samples depending on the specific model being estimated.

1. S1: sample consists of all observations we have.
2. S2: sample consists only of the purchase occasions at the 15 biggest key accounts, i.e. chains.
3. S3: sample consists only of the purchase occasions at the 13 biggest key accounts, excluding the two biggest chains from S2 that have only products that are not advertised on TV
4. S4: sample consists of all observations of sample 2, for which we could construct instruments
5. S5: sample consists of all observations of sample 3, for which we could construct instruments

4 Results

4.1 Estimation Results

Here come the results of the estimated models from section 2. The table headline specifies the estimated model, the sample according to the codes in the previous section, and the way prices enter the specification. As the literature does not quite agree whether to enter nominal price or efficiency price defined as price divided by contents in liters or kg of the detergent package, I estimate both versions. Tables with efficiency prices have this information in the headline. The control function enters as variable u the specifications. Variable *duration* controls for inventory effects. Variables name *inh* refer to the amount of detergent (in kg or liters) in the package.

The tables 4 to 7 show the results for the simple standard logits. Price effects are significant, and most other variables in the simple specifications too, especially duration to control the inventory effects and the retail activity variables. TV advertising shows not significance. Duration is negative as expected from section 2, retail activity variables positive.

Table 4: Logit Model - S1 - Part A

	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
purchase						
price	−.634***	−.527***	−.564***	−.447***	−.431***	.017
	(.01)	(.02)	(.02)	(.02)	(.02)	(.02)
inh	−.953***	−.887***	−.624***			
	(.02)	(.03)	(.03)			
duration		−.048***	−.042***	−.039***	−.072***	−.007*
		(.00)	(.00)	(.00)	(.00)	(.00)
liquid			−.946***	−2.951***	−2.886***	.428***
			(.02)	(.08)	(.08)	(.08)
handbill			.169**	.105	.072	.093
			(.06)	(.06)	(.06)	(.06)
feature			.213***	.188***	.151**	.493***
			(.06)	(.06)	(.06)	(.06)
display			.578***	.577***	.578***	.169**
			(.06)	(.06)	(.06)	(.06)
prceflag			−.118*	−.207***	−.230***	−.017
			(.05)	(.05)	(.05)	(.05)
inhp				−.862***	−.826***	−.029
				(.03)	(.03)	(.02)
inhl				.474***	.469***	−.194***
				(.05)	(.05)	(.05)
duration2					.001***	−.000*
					(.00)	(.00)
brand dummies	No	No	No	No	No	Yes
No. of obs	353911	299365	299365	299365	299365	299365
log-likelihood	−34466.12	−28032.13	−27170.26	−26878.06	−25957.85	−22872.37
Akaike-IC	68936.24	56070.26	54356.51	53774.12	51937.69	45796.75
Schwarz-IC	68957.79	56102.09	54441.39	53869.61	52054.40	46072.59

Note: preliminary results

Table 5: Logit Model - S1, S4 - Part B

	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
purchase						
price	–.391***	–.389***	–.382***	–.014	–.015	–.007
inhp	–.891***	–.893***	–.903***	–.048	–.047	–.044
inhl	.596***	.593***	.580***	–.035	–.036	–.156
duration	–.075***	–.075***	–.075***	–.013	–.013	.005
duration2	.001***	.001***	.001***	–.000	–.000	–.000**
liquid	–3.210***	–3.208***	–3.202***	.032	.036	.260
handbill	–.028	–.027	–.027	–.019	–.020	–.035
feature	.211	.212	.212	.555***	.553***	.566***
display	.520***	.518***	.520***	.172	.172	.146
prceflag	–.210*	–.211*	–.208*	–.035	–.035	.016
countc7ad	.070					
countc14ad	.098					
countc21ad	–.101					
countc28ad	.027					
timec7ad		.002		.002		
timec14ad		.004		.004		
timec21ad		–.005		–.006		
timec28ad		.002		.002		
powc7ad			.383		.386	1.279
powc14ad			–.306		–.356	–1.005
powc21ad			.205		.209	.521
powc28ad			–.114		–.117	–.253
GLdum						21.490
brand dummies	No	No	No	Yes	Yes	Yes
No. of obs	78431	78431	78431	78431	78431	73374
log-likelihood	–6575.91	–6576.23	–6577.01	–5842.90	–5842.97	–3742.18
Akaike-IC	13181.82	13182.46	13184.02	11745.80	11745.94	7546.36
Schwarz-IC	13320.87	13321.51	13323.07	12023.90	12024.04	7831.67

Note: preliminary results

Table 6: Logit Model - Eff. Prices - S1 - Part A

	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
purchase						
price	–1.279***	–1.131***	–1.062***	–1.028***	–1.015***	.020
	(.02)	(.02)	(.02)	(.02)	(.02)	(.02)
inh	–.760***	–.731***	–.712***			
	(.02)	(.02)	(.02)			
duration		–.033***	–.030***	–.030***	–.050***	–.007*
		(.00)	(.00)	(.00)	(.00)	(.00)
liquid			–.595***	–.992***	–.990***	.408***
			(.03)	(.09)	(.10)	(.08)
handbill			.166**	.152**	.116	.091
			(.06)	(.06)	(.06)	(.06)
feature			.329***	.325***	.288***	.492***
			(.06)	(.06)	(.06)	(.06)
display			.549***	.553***	.555***	.168**
			(.06)	(.06)	(.06)	(.06)
prceflag			.231***	.206***	.179***	–.020
			(.05)	(.05)	(.05)	(.05)
inhp				–.742***	–.720***	–.007
				(.02)	(.02)	(.01)
inhl				–.499***	–.481***	–.160***
				(.05)	(.05)	(.04)
duration2					.000***	–.000*
					(.00)	(.00)
brand dummies	No	No	No	No	No	Yes
No. of obs	353911	299365	299365	299365	299365	299365
log-likelihood	–31489.81	–26008.20	–25582.72	–25573.86	–24733.13	–22872.12
Akaike-IC	62983.63	52022.40	51181.44	51165.73	49488.26	45796.25
Schwarz-IC	63005.18	52054.23	51266.31	51261.21	49604.96	46072.09

Note: price in efficiency units preliminary results

Table 7: Logit Model - Eff. Prices - S1,S4 - Part B

	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
purchase						
price	–.934***	–.930***	–.913***	–.014	–.014	.127
inhp	–.775***	–.777***	–.786***	–.065*	–.065*	–.021
inhl	–.236**	–.239**	–.237**	–.061	–.063	–.135
duration	–.054***	–.054***	–.054***	–.013	–.013	.005
duration2	.000***	.000***	.000***	–.000	–.000	–.000**
liquid	–1.536***	–1.532***	–1.553***	.047	.051	.298
handbill	.060	.062	.067	–.017	–.018	–.021
feature	.326**	.326**	.326**	.557***	.555***	.586***
display	.481***	.483***	.487***	.173	.173	.151
prceflag	.145	.144	.151	–.031	–.031	.047
countc7ad	.087					
countc14ad	.096					
countc21ad	–.126					
countc28ad	.076					
timec7ad		.002		.002		
timec14ad		.004		.004		
timec21ad		–.006		–.006		
timec28ad		.004		.002		
powc7ad			.373		.385	1.296
powc14ad			–.277		–.356	–1.017
powc21ad			.186		.209	.519
powc28ad			–.081		–.118	–.253
GLdum						21.971
brand dummies	No	No	No	Yes	Yes	Yes
No. of obs	78431	78431	78431	78431	78431	73374
log-likelihood	–6288.44	–6289.91	–6294.64	–5842.93	–5843.00	–3741.34
Akaike-IC	12606.87	12609.83	12619.28	11745.86	11746.01	7544.68
Schwarz-IC	12745.92	12748.88	12758.33	12023.96	12024.11	7829.98

Note: preliminary results

The tables 8 to 11 show the pooled and panel random coefficient logit model. The results look qualitatively similar to the simple logit case. Note of course that the scale is very different due to the norming of the price variable in the case of efficiency prices.

Whereas advertising seems unimportant in most models, in the panel mixed logit model with efficiency prices of table 11 we get weak significance [fix: stars on coefficient missing, but some values are at 10% level significant]. Especially interesting is the effect of the brand dummies on the price coefficient in the same table: the price coefficients differ by factor 3 depending on the introduction of brand dummies!

The following tables 12 to 14 show results for the logit model with the "simple" endogeneity correction.

For the simple models in the IV Logit case of table 12, the endogeneity correction in fact behaves as the literature predicts. The absolute value of the price effect increases if we add u to the specification. If we add brand dummies as in table 13, interestingly the model gets messed up, as the dummies soak up all the information and hardly any other variable is informative.

If we keep out the brand dummies, and add TV advertising variables (when the variables with the suffix "ad" are introduced), price effects diminish (are half in magnitude compared to the simple specifications), but all TV advertising variables are insignificant. Note that this is another sample, as the TV information is not available for all households. At this point I have not checked, whether the price effects do reappear if we use the sample of 9977 obs in table 13 and keep both advertising and brand dummies out.

Check that in table 14 compared to table 13 the results are qualitatively similar.

Qualitatively, the results also stay the same doing the following pairwise comparisons (tables are not reported): (i) samples S4 vs. S5 (ii) with vs. without GL dummy.

The tables 15 to 18 show results for the mixed logit model with the "simple" endogeneity correction. As is visible in tables 15 and 16, the endogeneity correction does not its predicted job as in the IV Logit case for the simple specifications: adding u has no significant effect on the price coefficient. But we see, that the choice whether to use panel or pooled mixed logit is influential, in that it changes the level of the price coefficient quite dramatically. Note that duration has a clear impact on the price coefficient in the pooled case, but none of meaningful magnitude in the panel version.

When looking at the tables with efficiency prices, we see in table 17 that the inclusion of u reduces the price coefficient towards zero, which contradicts the previous literature. If we allow for nonlinear duration effects ($duration2$), the model is statistically better (LL-Ratio Test on last two columns).

Table 8: Mixed Logit Models, panel - S1

	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Mean					
inh	–.455***	–.459***	–.330***		
	(.03)	(.03)	(.03)		
price	–1.058***	–.981***	–.926***	–.876***	–.867***
	(.02)	(.02)	(.02)	(.02)	(.02)
duration		–.030***	–.022***	–.020***	–.040***
		(.00)	(.00)	(.00)	(.00)
liquid			–.869***	–1.937***	–1.936***
			(.03)	(.08)	(.08)
handbill			–.072	–.113	–.148*
			(.06)	(.06)	(.06)
feature			.364***	.338***	.295***
			(.06)	(.06)	(.06)
display			.491***	.502***	.495***
			(.06)	(.06)	(.06)
prceflag			–.117*	–.144**	–.178**
			(.05)	(.05)	(.06)
inhp				–.419***	–.405***
				(.03)	(.03)
inhl				.282***	.295***
				(.05)	(.05)
duration2					.000***
					(.00)
SD					
price	.639***	.617***	.552***	.537***	.532***
	(.01)	(.01)	(.01)	(.01)	(.01)
No. of obs	353911	299365	299365	299365	299365
log-likelihood	–31041.30	–25549.48	–24911.13	–24803.93	–24015.64
Akaike-IC	62088.60	51106.97	49840.25	49627.86	48055.29
Schwarz-IC	62120.93	51149.40	49935.74	49733.95	48182.60

Note: preliminary results

Table 9: Mixed Logit Models, pooled - Eff Prices - S1

	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Mean					
inh	.007	−.011	−.023		
	(.01)	(.01)	(.01)		
price	−6.260***	−5.341***	−5.477***	−5.483***	−5.590
	(.14)	(.14)	(.16)	(.16)	(.00)
duration		−.035***	−.037***	−.037***	−.039
		(.00)	(.00)	(.00)	(.00)
liquid			.108**	.222**	.196
			(.04)	(.08)	.
handbill			.025	.034	−.019
			(.08)	(.08)	(.00)
feature			.463***	.465***	.397
			(.08)	(.08)	.
display			.224**	.219**	.212
			(.08)	(.08)	.
prceflag			.156*	.164*	.131
			(.07)	(.07)	.
inhp				−.020	−.012
				(.01)	(.00)
inhl				−.086	−.069
				(.04)	(.00)
duration2					.000
					.
SD					
price	3.689***	3.178***	3.244***	3.244***	3.277
	(.10)	(.09)	(.10)	(.10)	.
No. of obs	353911	299365	299365	299365	299365
log-likelihood	−28602.51	−23944.54	−23900.05	−23898.88	.00
Akaike-IC	57211.02	47897.08	47818.10	47817.76	.00
Schwarz-IC	57243.35	47939.51	47913.59	47923.85	.00

Note: preliminary results

Table 10: Mixed Logit Model, panel - Eff. Prices - S1 - Part A

	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Mean					
inh	−.413***	−.436***	−.474***		
	(.01)	(.01)	(.01)		
price	−1.825***	−1.809***	−1.656***	−1.641***	−1.671***
	(.02)	(.03)	(.03)	(.03)	(.03)
duration		−.008***	−.006***	−.006***	−.006
		(.00)	(.00)	(.00)	(.00)
liquid			−.520***	−.710***	−.739***
			(.03)	(.09)	(.09)
handbill			.001	−.009	−.047
			(.06)	(.06)	(.06)
feature			.394***	.392***	.354***
			(.06)	(.06)	(.06)
display			.441***	.444***	.434***
			(.06)	(.06)	(.06)
prceflag			.142**	.133*	.100
			(.05)	(.05)	(.06)
inhp				−.485***	−.480***
				(.02)	(.02)
inhl				−.370***	−.356***
				(.05)	(.05)
duration2					−.000
					(.00)
SD					
price	.935***	.941***	.890***	.888***	.896***
	(.02)	(.02)	(.02)	(.02)	(.02)
No. of obs	353911	299365	299365	299365	299365
log-likelihood	−29086.17	−24255.91	−24000.78	−23998.12	−23232.39
Akaike-IC	58178.33	48519.82	48019.57	48016.25	46488.79
Schwarz-IC	58210.66	48562.26	48115.05	48122.34	46616.10

Note: preliminary results

Table 11: Mixed Logit Models, panel - Eff. Prices - S1 - Part B

	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
Mean						
inhp	–.006	–.527***	–.530***	–.527***	–.081**	–.080**
inhl	–.123**	–.163*	–.163*	–.166*	–.037	–.036
duration	.035***	–.008	–.008	–.008	.032***	.032***
duration2	–.001***	–.000	–.000	–.000	–.001***	–.001***
liquid	.332***	–1.198***	–1.204***	–1.190***	–.054	–.054
handbill	–.030	–.135	–.130	–.131	–.152	–.151
feature	.432***	.375**	.375**	.372**	.513***	.512***
display	.126*	.412***	.408***	.413***	.126	.125
prceflag	.006	.071	.070	.070	–.044	–.044
bdum10	–3.286***				–2.759***	–2.763***
price	–.447***	–1.557***	–1.548***	–1.557***	–.532***	–.530***
countc7ad		.047			.036	
countc14ad		.172			.158	
countc21ad		–.118			–.123	
countc28ad		.047			.021	
relc7ad			1.011			
relc14ad			–.105			
relc21ad			–1.122			
relc28ad			1.759			
timec7ad				.000		.000
timec14ad				.007		.006
timec21ad				–.006		–.006
timec28ad				.003		.002
SD						
price	.686***	.850***	.851***	.852***	.656***	.656***
brand dummies	Yes	No	No	No	Yes	Yes
No. of obs	299365	78431	78431	78431	78431	78431
log-likelihood	–21978.17	–5908.76	–5913.78	–5909.97	–5615.16	–5615.24
Akaike-IC	44010.33	11849.52	11859.56	11851.94	11292.32	11292.48
Schwarz-IC	44296.79	11997.84	12007.88	12000.26	11579.69	11579.85

Note: preliminary results

Table 12: IV Logit Model - S4 - Part A

	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
purchase						
price	−.617***	−.703***	−.656***	−.616***	−.805***	−.692***
	(.06)	(.07)	(.07)	(.07)	(.08)	(.09)
inh	−1.489***	−1.331***	−1.221***	−1.174***	−.818***	
	(.11)	(.12)	(.13)	(.14)	(.14)	
u		.457***	.580***		.834***	.804***
		(.12)	(.14)		(.14)	(.14)
duration			−.030***	−.030***	−.028***	−.027***
			(.01)	(.01)	(.01)	(.01)
liquid				−.522***	−.660***	−2.044***
				(.11)	(.11)	(.51)
handbill				.292	.405	.408*
				(.21)	(.21)	(.21)
feature				.522*	.444*	.486*
				(.22)	(.22)	(.22)
display				.439	.589*	.583*
				(.24)	(.25)	(.25)
prceflag				−.103	.029	−.038
				(.18)	(.18)	(.18)
inhp						−1.046***
						(.16)
inhl						−.103
						(.29)
No. of obs	46064	46064	39213	39213	39213	39213
log-likelihood	−2226.10	−2218.91	−1872.18	−1860.37	−1841.72	−1837.50
Akaike-IC	4456.20	4443.82	3752.35	3736.74	3701.43	3695.01
Schwarz-IC	4473.68	4470.04	3786.66	3805.36	3778.62	3780.78

Note: preliminary results

Table 13: IV Logit Model S4 - Part B

	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
purchase						
price	.112	−.034	−.240	−.410*	−.364*	.332
	(.15)	(.20)	(.15)	(.17)	(.17)	(.45)
inhp	−.144	.001	−1.953***	−1.631***	−1.713***	−1.887**
	(.23)	(.26)	(.32)	(.34)	(.34)	(.70)
inhl	−.290	−.258	.029	.123	.073	−.155
	(.40)	(.40)	(.64)	(.61)	(.62)	(.72)
duration	.027	.027	−.026	−.020	−.020	−.009
	(.02)	(.02)	(.03)	(.03)	(.03)	(.03)
duration2	−.001**	−.001**	−.000	−.000	−.000	−.001
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
liquid	.481	.592	−3.236**	−3.037**	−3.052**	−2.658
	(.70)	(.70)	(1.12)	(1.08)	(1.09)	(1.47)
handbill	.602**	.617**	.549	.637	.617	.817*
	(.23)	(.23)	(.38)	(.38)	(.38)	(.38)
feature	.518*	.497*	.903*	.873*	.889*	.770
	(.22)	(.22)	(.43)	(.43)	(.44)	(.44)
display	.133	.162	−.427	−.370	−.355	−.481
	(.25)	(.25)	(.55)	(.57)	(.57)	(.56)
prceflag	−.022	−.019	−.359	−.233	−.243	−.027
	(.21)	(.21)	(.36)	(.37)	(.37)	(.40)
u		.198		.568*	.541*	.092
		(.18)		(.26)	(.27)	(.32)
countc7ad			.115	.108		.141
			(.31)	(.31)		(.32)
countc14ad			.140	.136		.026
			(.34)	(.33)		(.36)
countc21ad			−.276	−.265		−.141
			(.43)	(.43)		(.43)
countc28ad			.157	.153		.072
			(.26)	(.26)		(.26)
liqc7ad					−1.167	
					(2066.86)	
liqc14ad					−7.749	
					(1420.93)	
liqc21ad					7.054	
					(1147.76)	
liqc28ad					−10.477	
					(999.25)	
brand dummies	Yes	Yes	No	No	No	Yes
No. of obs	39213	39213	9977	9977	9977	9977
log-likelihood	−1682.85	−1682.23	−458.01	−455.57	−455.64	−446.63
Akaike-IC	3415.71	3416.45	946.02	943.15	943.29	953.26
Schwarz-IC	3630.13	3639.45	1054.14	1058.48	1058.62	1169.50

Note: preliminary results

Table 14: IV Logit Model - Eff. Prices - S5 - Part B

	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
purchase						
price	.552**	.546*	–.217	–.368	–.322	–.005
	(.20)	(.25)	(.17)	(.19)	(.20)	(.60)
inhp	.214	.211	–2.217***	–2.002***	–2.061***	–1.142
	(.13)	(.14)	(.32)	(.33)	(.34)	(.89)
inhl	.450	.445	.421	.340	.317	1.014
	(.47)	(.48)	(.84)	(.83)	(.83)	(1.13)
duration	.050*	.050*	–.023	–.016	–.017	–.011
	(.02)	(.02)	(.04)	(.04)	(.04)	(.04)
duration2	–.002**	–.002**	–.001	–.001	–.001	–.001
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
liquid	–.135	–.134	–4.427**	–4.122**	–4.126**	–3.789*
	(.75)	(.75)	(1.53)	(1.52)	(1.53)	(1.67)
handbill	.728**	.730**	.545	.682	.670	.682
	(.25)	(.25)	(.43)	(.44)	(.43)	(.45)
feature	.920**	.921**	.689	.836	.859	.493
	(.30)	(.30)	(.66)	(.68)	(.68)	(.69)
display	.262	.263	–.199	–.111	–.075	–.413
	(.27)	(.27)	(.66)	(.67)	(.67)	(.69)
prceflag	–.025	–.026	–.168	–.099	–.115	–.046
	(.22)	(.22)	(.39)	(.40)	(.40)	(.43)
u		.008		.539*	.524	.190
		(.17)		(.27)	(.28)	(.36)
countc7ad			.158	.153		.140
			(.32)	(.32)		(.33)
countc14ad			.123	.116		.046
			(.35)	(.34)		(.36)
countc21ad			–.309	–.300		–.198
			(.42)	(.42)		(.43)
countc28ad			.195	.189		.111
			(.25)	(.25)		(.26)
liqc7ad					–1.160	
					(1678.19)	
liqc14ad					–7.218	
					(1180.07)	
liqc21ad					6.930	
					(1099.09)	
liqc28ad					–10.231	
					(967.48)	
brand dummies	Yes	Yes	No	No	No	Yes
No. of obs	25600	25600	7000	7000	7000	7000
log-likelihood	–1088.67	–1088.67	–310.26	–308.30	–308.61	–302.38
Akaike-IC	2227.34	2229.33	650.52	648.60	649.22	664.76
Schwarz-IC	2431.10	2441.24	753.32	758.26	758.88	870.37

Note: preliminary results

Table 15: IV Mixed Logit Models, pooled - S4

	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Mean							
inh	−.928***	−.841***	−.567**	−.614***	−.407*		
	(.16)	(.16)	(.18)	(.18)	(.19)		
price	−2.234***	−2.216***	−2.772***	−2.695***	−2.592***	−2.553***	−2.593***
	(.27)	(.25)	(.30)	(.30)	(.29)	(.29)	(.33)
u		.290	.397		.606**	.599**	.741***
		(.16)	(.21)		(.21)	(.21)	(.22)
duration			−.045***	−.046***	−.042***	−.042***	.007
			(.01)	(.01)	(.01)	(.01)	(.02)
liquid				−.205	−.294	−.763	−.936
				(.16)	(.16)	(.73)	(.74)
handbill				.602	.734*	.730*	.553
				(.35)	(.34)	(.34)	(.35)
feature				.890*	.755*	.760*	.693
				(.37)	(.35)	(.35)	(.36)
display				.238	.295	.302	.444
				(.38)	(.37)	(.37)	(.39)
prceflag				.019	.123	.103	.004
				(.29)	(.28)	(.28)	(.29)
inhp						−.465*	−.475*
						(.21)	(.22)
inhl						−.160	−.101
						(.42)	(.42)
duration2							−.001*
							(.00)
SD							
price	1.168***	1.128***	1.454***	1.427***	1.299***	1.295***	1.268***
	(.14)	(.13)	(.16)	(.16)	(.15)	(.15)	(.16)
No. of obs	46064	46064	39213	39213	39213	39213	39213
log-likelihood	−2146.09	−2144.53	−1799.21	−1792.40	−1788.34	−1788.12	−1712.93
Akaike-IC	4298.19	4297.05	3608.42	3602.80	3596.67	3598.24	3451.87
Schwarz-IC	4324.40	4332.00	3651.31	3679.99	3682.44	3692.58	3563.37

Note: preliminary results

Table 16: IV Mixed Logit Models, panel - S4

	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Mean							
inh	−1.122***	−1.001***	−.928***	−.906***	−.604***		
	(.13)	(.14)	(.15)	(.15)	(.16)		
price	−1.240***	−1.299***	−1.279***	−1.225***	−1.365***	−1.250***	−1.365***
	(.10)	(.11)	(.12)	(.11)	(.12)	(.13)	(.14)
u		.333*	.488**		.689***	.659***	.747***
		(.13)	(.15)		(.15)	(.15)	(.16)
duration			−.013*	−.012*	−.011	−.011	.054**
			(.01)	(.01)	(.01)	(.01)	(.02)
liquid				−.463***	−.560***	−1.774**	−2.015***
				(.11)	(.11)	(.57)	(.59)
handbill				.342	.431	.429	.302
				(.23)	(.23)	(.23)	(.25)
feature				.578*	.525*	.543*	.542*
				(.24)	(.24)	(.24)	(.26)
display				.431	.559*	.555*	.603*
				(.27)	(.27)	(.27)	(.29)
prceflag				−.115	−.019	−.072	−.191
				(.20)	(.20)	(.20)	(.22)
inhp						−.810***	−.883***
						(.19)	(.20)
inhl						.006	.058
						(.32)	(.33)
duration2							−.002***
							(.00)
SD							
price	.744***	.736***	.700***	.690***	.664***	.657***	.695***
	(.05)	(.05)	(.05)	(.05)	(.05)	(.05)	(.06)
No. of obs	46064	46064	39213	39213	39213	39213	39213
log-likelihood	−2083.63	−2080.49	−1776.10	−1765.39	−1754.54	−1752.08	−1666.83
Akaike-IC	4173.26	4168.98	3562.20	3548.78	3529.09	3526.15	3359.66
Schwarz-IC	4199.48	4203.93	3605.08	3625.97	3614.85	3620.50	3471.16

Note: preliminary results

Table 17: IV Mixed Logit Models, pooled - Eff. Prices - S4

	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Mean							
inh	−.998***	−1.000***	−.666***	−.709***	−.764***		
	(.13)	(.13)	(.18)	(.17)	(.17)		
price	−2.852***	−2.689***	−3.451***	−3.259***	−2.768***	−2.897***	−2.561***
	(.47)	(.44)	(.70)	(.64)	(.58)	(.60)	(.53)
u		.352*	.471**		.613***	.618***	.706***
		(.15)	(.18)		(.17)	(.17)	(.17)
duration			−.035***	−.035***	−.031***	−.032***	.015
			(.01)	(.01)	(.01)	(.01)	(.02)
liquid				−.212	−.247	.371	.033
				(.14)	(.13)	(.76)	(.75)
handbill				.638*	.748*	.761*	.555
				(.32)	(.30)	(.31)	(.30)
feature				.684*	.559*	.554	.505
				(.31)	(.28)	(.29)	(.28)
display				.215	.287	.261	.390
				(.34)	(.32)	(.33)	(.32)
prceflag				.138	.199	.232	.115
				(.26)	(.24)	(.25)	(.25)
inhp						−.699***	−.900***
						(.18)	(.18)
inhl						−1.086**	−1.098**
						(.42)	(.41)
duration2							−.001*
							(.00)
SD							
price	1.461***	1.359***	1.762***	1.668***	1.383***	1.427***	1.239***
	(.24)	(.22)	(.35)	(.32)	(.29)	(.29)	(.26)
No. of obs	46064	46064	39213	39213	39213	39213	39213
log-likelihood	−2152.44	−2149.66	−1808.26	−1801.76	−1795.73	−1795.38	−1719.77
Akaike-IC	4310.89	4307.32	3626.52	3621.52	3611.46	3612.76	3465.54
Schwarz-IC	4337.10	4342.27	3669.40	3698.71	3697.23	3707.11	3577.03

Note: preliminary results

Table 18: IV Mixed Logit Models, panel - Eff. Prices - S4

	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
Mean							
inh	−1.248***	−1.207***	−1.165***	−1.166***	−1.085***		
	(.09)	(.09)	(.10)	(.10)	(.10)		
price	−1.696***	−1.716***	−1.643***	−1.605***	−1.649***	−1.636***	−1.811***
	(.13)	(.13)	(.14)	(.14)	(.14)	(.15)	(.17)
u		.332*	.528***		.631***	.629***	.708***
		(.13)	(.15)		(.15)	(.15)	(.15)
duration			−.011	−.011	−.010	−.010	.064***
			(.01)	(.01)	(.01)	(.01)	(.02)
liquid				−.329**	−.357**	−.510	−.732
				(.11)	(.11)	(.65)	(.67)
handbill				.441	.534*	.533*	.408
				(.23)	(.23)	(.23)	(.25)
feature				.544*	.488*	.490*	.486*
				(.23)	(.23)	(.23)	(.25)
display				.362	.463	.466	.507
				(.26)	(.26)	(.26)	(.28)
prceflag				.001	.073	.066	−.051
				(.20)	(.20)	(.21)	(.22)
inhp						−1.098***	−1.219***
						(.11)	(.12)
inhl						−1.000**	−1.004**
						(.37)	(.38)
duration2							−.002***
							(.00)
SD							
price	.986***	.972***	.888***	.901***	.861***	.859***	.941***
	(.07)	(.07)	(.08)	(.08)	(.08)	(.08)	(.08)
No. of obs	46064	46064	39213	39213	39213	39213	39213
log-likelihood	−2064.94	−2061.69	−1759.28	−1752.50	−1742.82	−1742.79	−1657.44
Akaike-IC	4135.87	4131.38	3528.56	3523.01	3505.64	3507.58	3340.87
Schwarz-IC	4162.09	4166.33	3571.44	3600.20	3591.41	3601.93	3452.37

Note: preliminary results

Finally, the results of table 18 suggest, that price effects are very stable, and do hardly react to in-/exclusion of other variables. Even without additional controls the role of the price endogeneity correction is marginalized, thus showing that in the most flexible specification estimated so far it plays no role for the magnitude of the coefficients.

The estimates with brand dummies and TV advertising are still to come **[to be continued...]**

Erdem and Keane (1996) also estimate a comparable, reduced form model. They get no significant effect of advertising. In their structural model however, advertising matters. As they model advertising to have an informational effect, they show that the informational content of advertising is lower than that of experience by purchasing. Although dissatisfying, our results concerning the significance of the TV advertising variable fits their result. My results suggest, that perhaps the role of TV advertising is hard to measure, as it may contribute in the long run to the effect of the brand dummies through image / reputation effects. Perhaps as mentioned earlier the absence of price coupons in Germany is a problem for the endogeneity correction, perhaps in the US data it was this kind of measurement error, that it was correcting.

4.2 Model Fit and predictive ability

[to be completed...]

4.3 Estimation issues

[to be completed...]

5 Conclusion

We have estimated several discrete choice specifications to assess the effect of the endogeneity correction for prices suggested by Petrin and Train (2006). I do this controlling for five causes that are partly ignored in previous work when facing price endogeneity. The dataset employed is rich in detail and we are able to control the sources of the endogeneity away, such that the explanations given so far in the literature would not explain, if the price endogeneity correction would still have an effect. Although we have seen that only a few thousand observations remain if we use the full level of detail that is in the data, the price endogeneity correction loses its bite. I conclude that in situations with good price data, but otherwise crude information, the price endogeneity correction is useful. But as soon as the data gets richer in detail, the relevance of the endogeneity correction diminishes, which contrasts the results of the literature. Furthermore, the dominance of brand fixed effects in the estimated model pinpoints to a different understanding of TV advertising. TV advertising is insignificant in the short run to influence consumer purchase decisions, but may play a prominent role in the formation of brand fixed effects in the long run.

A Model Review

A.1 Model of Berry (1994)

Berry wants to estimate demand parameters from a discrete choice specification in the presence of an unobservable causing price endogeneity. In the basic setup, J products and I markets are observed because only aggregate data is observed. Observed characteristics are given by x_j where this is a scalar for simplicity and ξ_j is an unobserved characteristic. Note that x_j is assumed exogenous, ξ_j independent across markets, and ξ_j is mean independent of x_j: $E[\xi_j|x_j] = E[\xi_j]$. Consumer model is utility is given by

$$U(x_j, \xi_j, p_j, \nu_i; \theta_d) \tag{27}$$

where ν_i are unoberved consumer characteristics that require a parametric assumption and θ_d are demand parameters. Need to fix U and $f(\nu)$. Example is RC Model with deterministic and random part: $u_{ij} = \tilde{\xi}_j + \nu_{ij}$. Consumer chooses in market product j given x, ξ, p if:

$$U(x_j, \xi_j, p_j, \nu_i; \theta_d) > U(x_k, \xi_k, p_k, \nu_i; \theta_d) \quad \forall j \neq k \tag{28}$$

ν_i is only remaining RV. Collect the ν_i that lead to choice of j:

$$A_j(\tilde{\xi}) = \{\nu_i | \tilde{\xi}_j + \nu_{ij} > \tilde{\xi}_k + \nu_{ik}, \quad \forall j \neq k\} \tag{29}$$

Then market share of good j is $P(\nu_i \in A_j(\tilde{\xi}))$. Then need to assume a parametric form for ν, get cdf $F(\nu|x, \tilde{\xi})$ and density $f(\nu|x, \tilde{\xi})$. Then I have market share given as

$$S_j(\tilde{\xi}(x, p, \xi, \theta_d), x, \theta_d) = \int\limits_{A_j(\tilde{\xi})} f(\nu|x, \tilde{\xi}) \tag{30}$$

Suppose given data, parameters θ_d and known distribution of ν, market share is a function of $\tilde{\xi}$. At the true values of s and $\tilde{\xi}$ it must hold:

$$s_j = S_j(\tilde{\xi}) \longleftrightarrow \tilde{\xi} = S_j^{-1}(s_j) \tag{31}$$

Berry shows that this inversion works if $S_j(\tilde{\xi})$ is continuously differentiable, and the following restrictions on derivatives holds: $\frac{\partial S_j}{\partial \tilde{\xi}_j} > 0$ and $\frac{\partial S_j}{\partial \tilde{\xi}_k} < 0, \forall j \neq k$.

This leads to the following procedure for estimation:

1. Calculate $\widehat{\tilde{\xi}} = S_j^{-1}(s_j)$ from observed data of market shares.

2. Regress (IV) mean utility level on prices and controls: $\widehat{\tilde{\xi}} = x_j\beta - \alpha p_j + \xi_j$. Specification depends on parametric form of $U()$.

This works under two sets of assumptions: (a) independence across markets: ξ_{js} indp. ξ_{jr}, but ξ_{jr} not indp. ξ_{kr} (b) independence across firms in one market: ξ_j indp. ξ_s.

If ν not given, Berry suggests to specify a parametric family for ν. Then $f(\nu|x,\sigma)$ and $S_j = S_j(\tilde{\xi},\sigma)$ and $\tilde{\xi} = \tilde{\xi}(s,\sigma)$. Berry, Levinsohn, and Pakes (1995) study this case further.

A.2 Model of Chintagunta, Dubé and Goh (2005)

[to be completed...]

A.3 Model of Petrin and Train (2006)

[to be completed...]

References

Ackerberg, D. A. (2001): “Empirically Distinguishing Informative and Prestige Effects of Advertising,” *The RAND Journal of Economics*, 32(2), 316–333.

——— (2003): “Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination,” *International Economic Review*, 44(3), 1007–1040.

Bagwell, K. (2005): "The Economic Analysis of Advertising," forthcoming in Mark Armstrong and Rob Porter (eds.), Handbook of Industrial Organization, Vol. 3, North-Holland: Amsterdam.

Berry, S. (1994): "Estimating discrete-choice models of product differentiation," *Rand Journal of Economics*, 25, 242–262.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–889.

Briesch, R. A., P. K. Chintagunta, and R. L. Matzkin (2002): "Semiparametric Estimation of Brand Choice Behavior," *Journal of the American Statistical Association*, 97(460), 973–982.

Chintagunta, P., J. P. Dubé, and K. Y. Goh (2005): "Beyond the Endogeneity Bias: The Effect of Unmeasured Brand Characteristics on Household-Level Brand Choice Models," *Management Science*, 51(5), 832–849.

Ciao (2007): "ACNielsen Werbeforschungsunternehmen > Erfahrungsberichte," Ciao GmbH, Munich, Germany, Website, viewed 14 January 2007, http://www.ciao.de/ACNielsen_Werbeforschungsunternehmen__942530.

Erdem, T., and M. Keane (1996): "Decision Making under Uncertainty: Capturing Dynamic Brand Choice Porcesses in Turbulent Consumer Goods Markets," *Marketing Science*, 15(1), 1–20.

Guadagni, P., and J. Little (1983): "A logit model of brand choice calibrated on scanner data," *Marketing Science*, 2, 203–238.

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

Horsky, D., S. Misra, and P. Nelson (2006): "Observed and Unobserved Preference Heterogeneity in Brand-Choice Models," *Marketing Science*, 25(4), 322–335.

Kaul, A., and D. R. Wittink (1995): "Empirical Generalizations about the Impact of Advertising on Price Sensitivity and Price," *Marketing Science*, 14(3), 151–160.

Keane, M. (1997): "Modeling heterogeneity and state dependence in consumer choice behvior," *Journal of Business and Economic Statistics*, 15(3), 310–327.

Lauga, D. O. (2008): "Persuasive Advertising with Sophisticated but Impressionable Consumers," Working paper.

McFadden, D. (1974): "Conditional logit analysis of qualitative choice behavior," in *Frontiers in Econometrics*, ed. by P. Zarembka, chap. 4, pp. 105–142. New York: Academic Press.

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

Petrin, A., and K. Train (2006): "Control Function Corrections for Unobserved Factors in Differentiated Product Models," mimeo.

Train, K. (2003): *Discrete Choice Methods with Simulation*. Cambridge University Press.