Omitted Product Attributes in Discrete Choice Models

by
Amil Petrin
University of Chicago GSB and NBER, and
Kenneth Train
University of California, Berkeley

December 16, 2002

Abstract

We describe two methods for correcting an omitted variables problem in discrete choice models: a fixed effects approach and a control function approach. The control function approach is easier to implement and applicable in situations for which the fixed effects approach is not. We apply both methods to a cross-section of disaggregate data on customer's choice among television options including cable, satellite, and antenna. As theory predicts, the estimated price response rises substantially when either correction is applied. All of the estimated parameters and the implied price elasticities are very similar for both methods.

1 Introduction

Models of differentiated products are widely used for policy analysis (where impacts often depend on substitution patterns, such as whether the induced demand for new energy-efficient vehicles is drawn more from "gas-guzzlers" or current "gas-sippers"), merger analysis (where elasticities and cross-elasticities among similar products determine the welfare implications of a merger), design and forecasting of new products (where demand depends on the new product's similarity to other products, and where the issue of self-cannibalization of the firm's similar products is critical for profits), marketing analysis (where the demand for one product depends on the attributes of all similar products), and a host of other issues.

In aggregate (i.e., market-level) models of differentiated products, price is usually endogenous, determined by the interaction of demand and supply. Since the demand for differentiated products under heterogeneous preferences is inherently non-linear, the application of standard methods for correcting this endogeneity problem are not immediately applicable. Berry, Levinsohn, and Pakes (1995, henceforth BLP) developed and applied an approach using product-market fixed effects (as suggested by Berry, 1994) which provides consistent estimation in the face of omitted product attributes. The method has proven successful in numerous applications, including the demand for cable TV (Crawford, 2000), cereals (Nevo, 2001), and minivans (Petrin, 2002), to name only a few.

With disaggregate (i.e., customer-level) models of demand, price is not necessarily endogenous in the traditional sense, since the demand of the customer does not usually affect market price. However, the issues that give rise to the need for correction in aggregate models can also appear in disaggregate models. In particular, omitted product attributes can result in correlation between the price and the unobserved portion of utility: the market mechanism causes the price to be higher for products that display desirable attributes that are observed by consumers but not measured by the econometrician. Since these attributes affect demand at the customer level, price can be correlated with the error term even in disaggregate models. The BLP approach can be applied to disaggregate data, or a combination of aggregate and disaggregate data, as illustrated by Berry et al. (2001) and Goolsbee and Petrin (2002).

We propose an alternative approach based on control functions. The basic idea is to include extra variables in the estimation equation that condition out (*i.e.*, control for) the part of the error that is correlated with the regressors. The concept dates back at least to Heckman (1978) and Hausman (1978).[1] We implement it in a discrete choice environment where price endogeneity often raises econometric concerns.[2]

For general cases the assumptions under the control function approach differ from those for the fixed effects approach, with neither set of assumptions necessarily being nested within the other. For many questions the control function approach can be implemented in the most popular statistical packages, while the fixed effects approach usually requires the use of a lower-level programming language. Additionally, the control function approach is available in some situations for which the BLP approach is not. In the sections below we describe both approaches and apply them to disaggregate data on customers' choice among TV options. We find that the two approaches provide very similar estimates, including a substantial increase in the price response when either correction is applied.

2 Specification

Our specification is indexed by consumers (n), products (j), and markets (m). There are N consumers and J products. The price and attributes of the products vary over M markets. The price and some product attributes are observed by the econometrician; these are denoted p_{mj} and x_{mj}, respectively, for product j in market m. Some attributes are not observed by the econometrician but are known by consumers and affect their demand. The utility that customer n who lives in market m obtains from product j is decomposed into these observed and unobserved parts:

$$U_{nj} = V(p_{mj}, x_{mj}, s_n) + e_{nj} \tag{1}$$

where s_n denotes the observed characteristics of the customer, V is a calculable function up to parameters, and e_{nj} is defined as the difference that

[1] See also Madansky (1964).

[2] The term "control function" was introduced by Heckman and Robb (1985) in the context of selection models and has been applied to a Tobit model by Smith and Blundell (1986) and binary probit by Rivers and Vuong (1988). See also Trajtenberg (1989, 1990), Villas-Boas and Winer (1999), and Blundell and Powell (2001).

makes the equation an identity. Customer n buys in one of the markets; for simplicity, we say that customer n buys in market m without explicitly denoting the fact that m differs for different n.[3]

The choice probability is defined in the traditional way. Let e_n denote the vector $\langle e_{n1}, \ldots, e_{nJ} \rangle$, and let $\varphi(de_n)$ denote the density of e_n conditional on the observed variables. The choice probability for good i is then

$$P_{ni} = \int_{A_{ni}} \varphi(de_n) \tag{2}$$

where $A_{ni} = \{e_n \mid U_{ni} > U_{nj}\ \forall j \neq i\}$ is the set of e_n such that product i provides maximal utility.

Decomposing price into two components is an important aspect of both the fixed effects and the control function approach. Price in each market can be expressed as a function of instruments:

$$p_{mj} = g(j, z_m) + \mu_{mj}, \tag{3}$$

where z_m includes the observed attributes of the products and other observed variables that are independent of μ_{mj}. The error μ_{mj} incorporates factors like unobserved product attributes that affect price but are not captured by z_m.

The econometric difficulty arises when e_{nj} and p_{mj} are correlated (i.e. if e_{nj} and μ_{mj} are correlated). The standard choice models are derived under assumptions that do not incorporate this correlation. Logit and GEV models (e.g., McFadden, 1974, 1978) assume that the unobserved component of utility is independent of the observed variables. Mixed logit and probit (e.g., Brownstone and Train, 1999) allow the covariance of the unobserved component to depend on observed variables; however, the mean is assumed to be constant, which precludes correlation with price.

2.1 Fixed effects approach

The BLP approach includes alternative-specific constants for each product-market pair to address the price endogeneity problem. Utility is decomposed into a part that is the same for all customers in a market, labeled δ_{mj}, plus

[3] The more correct but more cumbersome notation would be $m(n)$ as the market in which n buys.

observed $\tilde{V}(\cdot)$ and unobserved $\tilde{\varepsilon}_{nj}$ parts that vary over customers in each market:

$$U_{nj} = \delta_{mj} + \tilde{V}(p_{mj}, x_{mj}, s_n) + \tilde{\varepsilon}_{nj} \tag{4}$$

Importantly, δ_{mj}, which is called the fixed effect for product j in market m, incorporates the average value of the omitted attributes along with the other components of utility that do not vary within a market.[4] It is typically assumed to be separable in price, observed attributes, and unobserved factors:

$$\delta_{mj} = \alpha p_{mj} + h(x_{mj}) + \xi_{mj}. \tag{5}$$

For a very general class of choice models Berry (1994) shows that such fixed effects exist and are unique.

The econometric concern arises when ξ_{mj} is correlated with p_{mj}, that is, if omitted attributes affect market prices. Without the fixed effect, the error entering the likelihood is $\xi_{mj} + \tilde{\varepsilon}_{nj}$, and this error does not have a constant mean conditional on price. Typically, it is positively correlated with price, which biases the price sensitivity towards zero.

Conditional on the fixed effect, the unobserved portion of utility is given by $\tilde{\varepsilon}_{nj}$. The choice probability is then given by

$$P_{ni} = \int_{B_{ni}} \psi(d\tilde{\varepsilon}_n)$$

where $B_{ni} = \{\tilde{\varepsilon}_n \mid U_{ni} > U_{nj} \ \forall j \neq i\}$ and $\psi(d\tilde{\varepsilon}_n)$ is the density of $\tilde{\varepsilon}_n = \langle \tilde{\varepsilon}_{n1}, \ldots, \tilde{\varepsilon}_{nJ} \rangle$ conditional on the observed variables and fixed effects. This density is usually assumed to take one of the standard forms, since the component of utility that is assumed correlated with price is absorbed in δ_{mj}.

The fixed effects approach consists of three steps. First, the discrete choice model is estimated with alternative-specific constants for each product in each market. The model can include interactions of demographic variables with observed attributes, which constitute the elements of $\tilde{V}(\cdot)$. The form of this model, such as logit, GEV, probit or mixed logit, is determined by the distributional assumption that is placed on $\tilde{\varepsilon}_n$. Again,

[4] Our use of the term "fixed effects" is analogous to that in cross-section/time-series models, where the fixed effects capture the average in each geographical area around which time-specific values vary. Here the fixed effects capture the average in each market around which customer-specific values vary.

by including the fixed effects, consistent estimates of all parameters except those contained in the fixed effects obtain because ξ_{mj} is controlled for (i.e. the new error, $\tilde{\varepsilon}_{nj}$, is conditional on ξ_{mj}).

Since the first stage only provides estimates of parameters *not* included in the fixed effect, two more steps are undertaken to recover estimates of the remaining parameters. In the second step, the price equation (3) is estimated using exogenous instruments. The third step is the regression of fixed effects on characteristics and price (equation (5)) using predicted price instead of the actual price.[5] With estimates for all parameters available after this last step, any function of the parameters (like price elasticities) is straightforward to compute.

A computational issue arises because of the large number of parameters in the model. One fixed effect is estimated for each product (except one, for normalization) in each market, which constitutes $M \cdot (J-1)$ parameters in addition to those entering $\tilde{V}$ and those determining the covariance of $\tilde{\varepsilon}_n$. The $M \cdot (J-1)$ fixed effects can in principal be estimated by maximum likelihood along with the other parameters; however, doing so requires searching over a very high-dimensional parameter space.

BLP propose one approach to facilitate estimation. At each trial value of parameters entering $\tilde{V}$ and the covariance of $\tilde{\varepsilon}_n$, fixed effects are calculated that induce the forecasted shares in each market to equal the sample shares in that market, and the likelihood is calculated at these values. By concentrating the fixed effects out, the parameters entering $\tilde{V}$ and the covariance of $\tilde{\varepsilon}_n$ can be estimated using standard optimization methods.

BLP provide a computational device to aid in concentrating out the fixed effects. The fixed effects are calculated iteratively by repeated application of the formula:

$$\delta_{mj}^{t+1} = \delta_{mj}^{t} + ln(S_{mj}) - ln(F_{mj}^{t}).$$

where t denotes the iteration, S_{mj} is the sample share for product j in market m, and F_{mj}^t is the forecasted share for product j in market m calculated with $\delta_{mj}^t \ \forall j$. As mentioned above, this iteration for the fixed effects is performed for *each* trial value of the other parameters. It is both the computationally difficult and computationally costly part of the fixed effects estimation

[5] The second and third steps are actually performed simultaneously using two stage least squares (2SLS). We break 2SLS into its two component steps to facilitate discussion in section 2.3

routine, and it is entirely avoided in the control function approach.[6]

2.2 Control function approach

The basic logic of a control function approach is to add to the estimation equation an additional variable that conditions out the part of the error correlated with the regressors. In our framework the control function approach includes a new regressor to control for the part of price that varies with the omitted attributes.[7] As with the fixed effects approach, the price equation decomposes price into a part that is attributable to the instruments, including the observed attributes, and a residual that is attributable to other factors, including omitted attributes. The price residuals, $\mu_m = \langle \mu_{m1}, \ldots, \mu_{mJ} \rangle$, are called control variables. For each product, a control function is specified, denoted as $f_j(\mu_m)$ for product j. These control functions are new variables that, once calculated, can enter utility like any other observed variable.

With the new regressors $f_j(\mu_m)$ we write utility as

$$U_{nj} = \mathbf{V}(p_{mj}, x_{mj}, s_n, f_j(\mu_m)) + \varepsilon_{nj}, \tag{6}$$

where $\mathbf{V}$ is a calculable function of the observed variables, including the control functions, and ε_{nj} is the difference that makes this equation an identity. In principal, any function $\mathbf{V}$ can be specified, as well as any control function. Heterogeneity can be incorporated by interacting characteristics with demographics and/or giving them random coefficients. Similarly, the

[6] The procedure is called a contraction (BLP) or calibration (Train, 1986.) When utility contains an additive extreme value error, the procedure is guaranteed to converge; otherwise, it is not. Other approaches can be used instead; see Goolsbee and Petrin for an approach that is effective when the error is multivariate normal. Note that the fixed effects that are obtained by these methods are not the maximum likelihood estimates, except for a logit model. For probit and mixed logit, the MLE fixed effects do not in general equate sample shares with forecasted shares, so the procedure might be considered "near" MLE.

[7] Trajtenberg (1989, 1990) utilizes a similar approach of entering the price residual as a variable in a discrete choice model in the context of omitted attributes. However, he does not also enter the actual price, and thus does not estimate the price coefficient along with the rest of the model estimates. Instead he uses additional outside information to estimate the price coefficient (as he describes on pp. 463-465 of the 1989 paper). Bajari and Benkard (2001) provide the conditions under which the exact value of an unobserved attribute can be recovered from the price equation when a single market is under consideration.

control functions can be interacted with demographics and/or given random coefficients.

The choice probability is defined by the density of ε_n, which is conditional on the control functions and the specification of $\mathbf{V}$:

$$P_{ni} = \int_{C_{ni}} \phi(d\varepsilon_n)$$

where $C_{ni} = \{\varepsilon_n \mid U_{ni} > U_{nj} \ \forall j \neq i\}$ and ϕ is the density of ε. If ϕ takes a convenient form, then this choice probability can be used as the basis for parameter estimation. As mentioned above, ε_{nj} need not be independent of the control functions, price, or the other variables in $\mathbf{V}$, since probit and mixed logit allow its covariance to depend on observed variables. However, for all the standard models, ε_{nj} must have constant mean conditional on these variables. The purpose of the control function is to condition on the part of price related to omitted attributes so that remaining unobserved utility ε_{nj} has constant mean conditional on price.

Estimation is performed in two computationally simple steps. First, the price equations are estimated. Second, the choice model is estimated with the residuals from the price equations included as extra explanatory variables. The simplicity of the approach is an important part of its appeal; for many questions it can be implemented in the most popular statistical packages, while the fixed effects approach usually requires the use of a lower-level programming language.

The control function approach can be applied in some cases where the fixed effects approach is not feasible. For example, price might vary endogenously over every observation rather than just over groups of observations (e.g., over markets). In this case, the fixed effects are not separately identified from the other parameters in the model (because there is one fixed effect for each observation). However, as long as an appropriate control function can be specified, all of the parameters under the control function approach are identified.

Starting with the original specification from equation (1), some special cases are illustrative. A prominent example arises when e_{nj} consists of a component that varies over markets and a remainder, $e_{nj} = \xi_{mj} + \tilde{\varepsilon}_{nj}$, with ξ_{mj} and μ_{mj} distributed jointly normal. The conditional mean of e_{nj} is then $\Omega_{\xi\mu}\Omega_{\mu\mu}^{-1}\mu_{mj}$, where the Ω's are the appropriate covariance matrices. In this case, the control functions are linear in the price residuals and enter utility linearly.

Second, if the relevant covariances over products are zero (*i.e.*, $\Omega_{\xi\mu}$ and $\Omega_{\mu\mu}$ are diagonal), then the control function for each product is simply proportional to the price residual for that product: $f_j(\mu_m) = \lambda_j \mu_{mj}$. This is the specification used by Villas-Boas and Winer (1999).

More generally, any random variable can be divided into its (conditional) mean and a deviate from this mean. Thus, writing e_{nj} as its mean conditional on μ_m and a remainder, or

$$e_{nj} = f_j(\mu_m) + \varepsilon_{nj},$$

yields the expression for utility

$$U_{nj} = V(p_{mj}, x_{mj}, s_n) + f_j(\mu_m) + \varepsilon_{nj},$$

which is a control function model. We now turn to a comparison of the control function and the fixed effects approaches.

2.3 Comparison under standard specifications

In the most common specifications, both approaches can be seen as including the price residuals but differing in how they enter the likelihood. Using this common feature of both approaches, we now show when parameter estimates will differ if the fixed effects model is the "correct" model and a simple control function approach is used to approximate it.

We explore the most popular specification for the fixed effects approach, i.e., when the fixed effects are assumed to be linear in price, observed and unobserved attributes. Substituting $\hat{p}_{mj} + \hat{\mu}_{mj}$ for p_{mj} in (5) yields

$$\delta_{mj} = \alpha \hat{p}_{mj} + \beta x_{mj} + (\alpha \hat{\mu}_{mj} + \xi_{mj}). \qquad (7)$$

Estimates of α and β are obtained from this estimating equation in the third step of the fixed effects routine, where the fixed effects are regressed on instrumented price $\hat{p}_{mj}$ and other regressors, with the error term in parentheses. Consistent estimates result because $\hat{p}_{mj}$ and x_{mj} are, by construction, uncorrelated with $\hat{\mu}_{mj}$ and, by assumption, uncorrelated with ξ_{mj}.

By defining $\tilde{\xi}_{mj}$ as

$$\tilde{\xi}_{mj} = \xi_{mj} - \lambda_j \hat{\mu}_{mj}, \qquad (8)$$

the remaining part of the error ξ_{mj} after μ_{mj} is accounted for, (5) (and (7)) can be reexpressed as

$$\delta_{mj} = \alpha p_{mj} + \beta x_{mj} + \lambda_j \hat{\mu}_{mj} + \tilde{\xi}_{mj}. \tag{9}$$

This equation is the equation that would be estimated if $\hat{\mu}_{mj}$ were directly included in the fixed effects regression with *actual* prices (and other attributes). As Hausman (1978) notes, OLS on (9) yields identical estimates for α and β as OLS on (7)[8], which in turn is identical to IV/2SLS on (5). The fixed effects approach is therefore equivalent to entering the price residuals into the fixed effects regression. The price residuals control for the correlation between price and the unobserved portion of utility.

The control function approach uses the price residuals for the same purpose but enters them directly into the choice model. The analogous control function model is obtained by substituting (9) into (4), or

$$U_{nj} = \alpha p_{mj} + \beta x_{mj} + \tilde{V} + \lambda_j \hat{\mu}_{mj} + (\tilde{\xi}_{mj} + \tilde{\varepsilon}_{nj}). \tag{10}$$

The question arises: does the additional error term $\tilde{\xi}_{mj} = \xi_{mj} - \lambda_j \hat{\mu}_{mj}$, which is part of unobserved utility in the control function approach that is controlled for in the fixed effects approach, lead to inconsistent estimates in the discrete choice model? Stated alternatively: Suppose we could estimate a discrete choice model with $\tilde{\xi}_{mj}$ entering explicitly as a variable. Would the estimated coefficients of the other variables and their implications for elasticities be different from the control function approach?

By the assumptions of the fixed effects approach, $\tilde{\xi}_{mj}$ has *constant mean* conditional on p_{mj}, x_{mj}, and $\hat{\mu}_{mj}$. Therefore, its inclusion/exclusion will affect the estimated parameters in this non-linear environment if: (1) it is correlated with the customer-level demographics in $\tilde{V}$, or, (2) its deviations around its mean are not represented correctly in the analyst's specification of the error distribution. As in any choice model, misspecification can be directly addressed by appropriate modification of the specification. Heterogeneity can be incorporated by interacting the control functions with

[8] In (9), substitute $\hat{p}_{mj} + \hat{\mu}_{mj}$ for p_{mj}: $\delta_{mj} = \alpha \hat{p}_{mj} + \beta x_{mj} + (\lambda_j + \alpha)\hat{\mu}_{mj} + \tilde{\xi}_{mj}$. OLS on this equation gives the same estimates for α and β and OLS on (9). OLS on this equation also gives the same estimates for α and β as OLS on (7), since $\hat{\mu}_{mj}$ is orthogonal by construction to $\hat{p}_{mj}$ and x_{mj}. OLS on (9) therefore gives the same estimates as OLS on (7).

demographics. Heteroskedasticity is addressed through alternative-specific error components in a mixed logit, unequal variances in a probit, and/or by giving the control functions random coefficients. Finally, testing and semi- and non-parametric methods, as described by Blundell and Powell (2001), can be used to assist in determining the appropriate specifications, as we now illustrate.

3 Application

We apply the methods to households' choice of television reception options. The specification and data are similar to those of Goolsbee and Petrin (2002). Four alternatives are considered available to households: (1) antenna only, (2) cable with basic or extended service, (3) cable with a premium service added, such as HBO, and (4) satellite dish. Basic and extended cable are combined because the data do not differentiate which of these options the households chose. Goolsbee and Petrin describe the market for cable and satellite TV, emphasizing the importance of accounting for omitted attributes, such as the quality of programming, in demand estimation. We apply both the fixed effects and the control function approach to data from 2001.

Our sample consists of 11,810 households in 172 geographically distinct markets. Each market contains one cable franchise that offers basic, extended, and premium packages. There are a number of multiple system operators like AT+T and Time-Warner which own many cable franchises throughout the country (thus serving several markets). The price and other attributes of the cable options vary over markets, even for markets served by the same multiple system operator. Satellite prices do not vary geographically, and the price of antenna-only is assumed to be zero. The price variation that is needed to estimate price impacts arises from the cable alternatives. Details of the data are given in the appendix.

For the control function approach, utility is specified (after extensive testing to be described below) as:

$$U_{nj} = \alpha p_{mj} + \sum_{g=2}^{5} \theta_g p_{mj} d_{gn} + \beta x_{mj} + k_j s_n + \lambda_j \mu_{mj} + (\sigma \nu_n c_j + \epsilon_{nj}). \quad (11)$$

The price effect is specified to differ by income group. Five income groups are identified, with the lowest income group taken as the base. The dummy d_{gn} identifies whether household n is in income group g. The price coefficient for a household in the lowest income group is α while that for a household in group $g > 1$ is $\alpha + \theta_g$. The alternative-specific constant for alternative j is k_j. These constants are interacted with demographic variables as well as entering directly. The variable μ_{mj} is the residual from the first-stage price regression, for j representing either extended-basic cable and premium cable. No such residuals are included for antenna-only and satellite since these prices do not vary geographically. These residuals are the control functions that are included to account for omitted attributes; we discuss their construction and alternative specifications below.

An error component is included to allow for correlation in unobserved utility over the three non-antenna alternatives. In particular, $c_j = 1$ if j is one of the three non-antenna alternatives and $c_j = 0$ otherwise, and ν_n is an iid standard normal deviate. The coefficient σ is the standard deviation of the error component, reflecting the degree of correlation among the non-antenna alternatives.

The final error term, ϵ_{nj}, is assumed to be iid extreme value, conditional on the explanatory variables including the control functions.[9] The choice probability therefore takes the form of a mixed logit (Train, 1998; Brownstone and Train, 1999), with the mixing over the distribution of ν_n:

$$P_{ni} = \int \frac{e^{V_{ni}+\sigma\nu c_i}}{\sum_{j=1}^{4} e^{V_{nj}+\sigma\nu c_j}} h(\nu) d\nu$$

where $h(\cdot)$ is the standard normal density and $V_{nj} = \alpha p_{mj} + \sum_{g=2}^{5} \theta_g p_{mj} d_{gn} + \beta x_{mj} + k_j s_n + \lambda_j \mu_{mj}$. The integral is approximated through simulation: a value of ν is drawn from the standard normal density, the logit formula is calculated for this value of ν, the process is repeated for numerous draws, and the results are averaged.[10]

Table 1 gives the estimated parameters. The first column gives the model without any correction for the correlation between price and omitted

[9]Note that ε_{nj} in (6) is the sum of the two error terms, $\sigma\nu_n c_j + \epsilon_{nj}$, in (11).

[10]To increase accuracy, Halton (1960) draws are used instead of independent random draws. Bhat(2001) found that 100 Halton draws perform better than 1000 independent random draws, a result that has been confirmed on other datasets by Train (2000, 2003), Hensher (2001), and Munizaga and Alvarez-Daziano (2001).

attributes; utility is the same as specified above except that the residuals, μ_{mj}, are not included. The second column applies the control function approach by including the residuals.[11] Without correction, the base price coefficient α is small, sufficiently so that the price coefficient $\alpha+\theta_g$ is positive for three of the five income groups, rendering the model implausible and unusable for policy analysis. Inclusion of the control functions raises the magnitude of the estimated base price coefficient, as expected. A negative price coefficient is obtained for all incomes groups. The magnitude decreases as income rises, with the highest income group obtaining a price coefficient that is about thirty percent smaller than that of the lowest income group.

Several product attributes are included in the model. In the model without correction, one of these attributes enters with an implausible sign: number of cable channels. With correction, all of the product attributes enter with expected signs. The magnitudes are generally reasonable. An extra premium channel is valued more than an extra cable (non-premium) channel. An extra over-the-air channel is also valued more than an extra non-premium cable channel, presumably because there are fewer over-the-air channels such that each one becomes more valuable. One interpretation is that the proliferation of cable channels with low programming content makes the value of extra cable channels relatively low. The option to obtain pay-per-view is valued highly. Note that this attribute, unlike the others, is not on a per-channel basis; its coefficient represents the value of the option to purchase pay-per-view events. The point estimates imply that households

[11] Since $\hat{\mu}$ are used to approximate μ in the estimation routine, the standard errors from the traditional formulas (and output by standard estimation routines) are biased downward. To correct for the additional source of variance in $\hat{\mu}$, we add a new term to the estimated variance of the parameters obtained from treating $\hat{\mu}$ as the true μ. This new component comes from bootstrapping the price regressions. That is, we repeatedly estimate the price regressions with bootstrapped samples, calculate the residuals, and re-estimate the mixed logit model with the new residuals. The variance in parameter estimates over the bootstrapped price samples is added to the variance estimates that obtain from the traditional formulas (which are appropriate when μ is observed without error). These total standard errors are given in the table. The adjustment is important for the standard errors of the base price coefficient, the coefficients for the residuals, and the coefficients of the product attributes, which increase between 50-100%, more closely approximating the standard errors from of the fixed effect approach. Karaca-Mandic and Train (2002) provide a formula for the asymptotic standard errors in this type of two-step estimation; they find that in our application the formula gives standard errors that are very similar to those obtained with the bootstrap procedure.

Table 1: Mixed Logit Model of TV Reception Choice
Control Function Approach

Alternatives: 1. Antenna only, 2. Basic and extended cable, 3. Premium cable, 4. Satellite
Variables enter alternatives in parentheses and zero in other alternatives.

Explanatory variable	Uncorrected	With control functions
	(Standard errors in parentheses)	
Price, in dollars per month (1-4)	-.0202 (.0047)	-.0969 (.0400)
Price for income group 2 (1-4)	.0149 (.0024)	.0150 (.0025)
Price for income group 3 (1-4)	.0246 (.0030)	.0247 (.0031)
Price for income group 4 (1-4)	.0269 (.0034)	.0269 (.0035)
Price for income group 5 (1-4)	.0308 (.0036)	.0308 (.0038)
Number of cable channels (2,3)	-.0023 (.0011)	.0026 (.0029)
Number of premium channels (3)	.0375 (.0163)	.0448 (.0233)
Number of over-the-air channels (1)	.0265 (.0090)	.0222 (.0111)
Whether pay per view is offered (2,3)	.4315 (.0666)	.5813 (.1104)
Indicator: ATT is cable company (2)	.1279 (.0946)	-.1949 (.1845)
Indicator: ATT is cable company (3)	.0993 (.1195)	-.2370 (.1944)
Indicator: Adelphia Comm is cable company (2)	.3304 (.1224)	.3425 (.1898)
Indicator: Adelphia Comm is cable company (3)	.2817 (.1511)	.2392 (.2246)
Indicator: Cablevision is cable company (2)	.6923 (.2243)	.1342 (.3677)
Indicator: Cablevision is cable company (3)	1.328 (.2448)	.7350 (.3856)
Indicator: Charter Comm is cable company (2)	.0279 (.1010)	-.0580 (.1441)
Indicator: Charter Comm is cable company (3)	-.0618 (.1310)	-.1757 (.1825)
Indicator: Comcast is cable company (2)	.2325 (.1107)	-.0938 (.2072)
Indicator: Comcast is cable company (3)	.5010 (.1325)	.1656 (.2262)
Indicator: Cox Comm is cable company (2)	.2907 (.1386)	-.0577 (.2496)
Indicator: Cox Comm is cable company (3)	.5258 (.1637)	.0874 (.2954)
Indicator: Time-Warner is cable company (2)	.1393 (.0974)	-.0817 (.1507)
Indicator: Time-Warner cable company (3)	.2294 (.1242)	-.0689 (.1891)
Education level of household (2)	-.0644 (.0220)	-.0619 (.0221)
Education level of household (3)	-.1137 (.0278)	-.1123 (.0280)
Education level of household (4)	-.1965 (.0369)	-.1967 (.0372)
Household size (2)	-.0494 (.0240)	-.0518 (.0241)
Household size (3)	.0160 (.0286)	.0134 (.0287)
Household size (4)	.0044 (.0357)	.0050 (.0358)
Household rents dwelling (2-3)	-.2471 (.0867)	-.2436 (.0886)
Household rents dwelling (4)	-.2129 (.1562)	-.2149 (.1569)
Single family dwelling (4)	.7622 (.1523)	.7649 (.1523)
Residual for extended-basic cable price (2)		.0805 (.0416)
Residual for premium cable price (4)		.0873 (.0418)
Alternative specific constant (2)	1.119 (.2668)	2.972 (1.057)
Alternative specific constant (3)	.1683 (.3158)	2.903 (1.487)
Alternative specific constant (4)	-.2213 (.4102)	4.218 (2.386)
Error components, standard deviation (2-4)	..5087 (.6789)	.5553 (.8567)
Log likelihood at convergence	-14660.84	-14635.47
Number of observations: 11810		

are willing to pay \$6.00 to \$8.88 per month for this option, depending on their income.

Several demographic variables enter the model. Their estimated coefficients are fairly similar in the corrected and uncorrected models. The estimates suggest that households with higher education tend to purchase less TV reception: the education coefficients are progressively more highly negative for antenna-only (which is zero by normalization), extended-basic cable, premium cable, and satellite. Larger households tend not to buy extended-basic cable as readily as smaller households. Differences by household size with respect to the other alternatives are highly insignificant. A dummy for whether the household rents its dwelling is included in the two cable alternatives and separately in the satellite alternative. These variables account for the fact that renters are perhaps less able to install a cable hookup and less willing to incur the capital cost of a satellite dish than a household that owns its dwelling. The estimated coefficients are negative, confirming these expectations. Finally, a dummy for whether the household lives in a single-family dwelling enters the satellite alternative, to account for the fact that it is relatively difficult to install a satellite dish on a multi-family dwelling. As expected, the estimated coefficient is positive.

The residuals from the first-stage price regressions enter the model to account for the omitted attributes. These control functions are created as follows. The price in each market was regressed against the product attributes listed in Table 1 plus Hausman (1997a)-type price instruments. The price instrument for market m is calculated as the average price in other markets that are served by the same multiple system operator as market m. A separate instrument is created for the price of extended-basic cable and the price of premium cable. Separate regressions were run for extended-basic price and premium price, using all of the instruments in each equation (the instruments are discussed later). The residuals were calculated from the estimated regressions. These residuals enter without transformation in the mixed logit model. For the main specification the residual from the extended-basic cable price regression enters the extended-basic cable alternatives, and similarly for the premium cable.

The residuals enter significantly and with the expected sign. In particular, a positive residual occurs when the price of the product is higher than can be explained by observed attributes and other observed factors. A positive residual suggests that the product possesses desirable attributes that are not included in the analysis. The residual entering the demand model with a positive coefficient is consistent with this interpretation.

As stated above, the appropriate control function to include is a specification issue. We experimented with other specifications, including the use of both residuals in each alternative, a series expansion on the residuals (both signed and unsigned), random coefficients on the residuals, and the residuals interacted with other variables. In all cases, the extra generality did not affect the estimated parameters, and the hypothesis that the simpler specification is correct could not be rejected at any reasonable level of confidence.[12]

We turn now to the fixed effect approach. All of the elements of utility that do not vary within a market are subsumed into the fixed effects. The fixed effects are expressed as a function of price and other observed attributes:

$$\delta_{mj} = \alpha p_{mj} + \beta x_{mj} + \xi_{mj}.$$

The utility specification given above becomes:

$$U_{nj} = \delta_{mj} + \sum_{g=2}^{5} \theta_g p_{mj} d_{gn} + k_j s_n + \sigma \nu_n c_j + \tilde{\epsilon}_{nj}.$$

Assuming $\tilde{\epsilon}_{nj}$ and ν_n are iid extreme value and standard normal respectively leads to a mixed logit of the same form as for the control function approach except with fixed effects for each alternative and market.

Estimation is performed in two stages. First the mixed logit model is estimated, using the contraction procedure described above for the fixed effects. Then the fixed effects are regressed against the product attributes using 3SLS. A separate equation is used for the extended-basic cable, premium cable, and satellite fixed effects, with the coefficients of the product attributes constrained across equations so as to be consistent with the usual BLP approach (and the model in Table 1). The *negative* of the number of over-the-air channels enters these equations, since this attribute enters the

[12]We also tested for specification issues unrelated to the control function, including random coefficients of other variables, and other types of error components. The simpler specification could not be rejected, and the parameters estimates were essentially unchanged.

antenna-only alternative in the model of Table 1 whereas it is now entering the fixed effects of the non-antenna alternatives.

The results are given in Table 2. The bottom part of the table gives the estimates of the demographic coefficients in the mixed logit model. The top part of the table gives the results of the regression of fixed effects on product attributes. The first column at the top gives the OLS results, which do not account for omitted attributes, and the second column gives the 3SLS results.

As with the control function approach, the correction for omitted variables raises the price coefficient. Without correction, three of the five income groups receive a positive estimated price coefficient. With correction, all groups obtain a significantly negative price coefficient.

The estimated base price coefficient is -.0922, compared to the -0.0969 obtained with the control function approach.[13] The difference is not statistically significant at any reasonable confidence level. The estimates of θ_g, the incremental price coefficient for higher income groups, are very similar under the two approaches. As in the control function approach, the number of cable channels obtains a negative coefficient when endogeneity is ignored and becomes positive as expected when the endogeneity is corrected. All of the product attributes obtain similar values as with the control function approach. We tested the hypothesis that the coefficients of the product attributes and the base price coefficient are the same as the point estimates from the control function approach (i.e., as in Table 1.) The test statistic for a Wald test is 0.88, which with five degrees of freedom has a P-value of 0.9717, indicating that the hypothesis of equality cannot be rejected at any meaningful level of confidence.[14]

The demographic coefficients in Table 2 provide similar conclusions as

[13]Since the paper was first circulated, our finding of similar estimates under the two approaches has been reported on two other datasets by different authors. J-P Dubé reestimated the fixed effects model in Chintagunta, Dubé, and Goh (2002) for margarine demand using a control function approach. He reported that the estimates and implied elasticities are very similar under the two methods, and different from those estimates obtained without either correction. Karaca-Mandic (2002) reported the same finding for her demand model for DVDs.

[14]This test does not account for the variation in the estimates from the control function approach; however, the P value for a test that takes this variation into account would be even higher.

Table 2: Mixed Logit Model of TV Reception Choice
Fixed Effects Approach

Alternatives: 1. Antenna only, 2. Basic and extended cable, 3. Premium cable, 4. Satellite Variable enters alternatives in parentheses and is zero in other modes.		
Explanatory variable	OLS	3SLS
	(Standard errors in parentheses)	
Price, in dollars per month (1-4)	-.0245 (.0091)	-.0922 (.0409)
Number of cable channels (2,3)	-.0024 (.0027)	.0017 (.0042)
Number of premium channels (3)	.0132 (.0502)	.0463 (.0329)
Number of over-the-air channels (neg.) (1)	.0168 (.0132)	.0196 (.0186)
Whether pay per view is offered (2,3)	.5872 (.1326)	.7144 (.1814)
Indicator: ATT is cable company (2)	-.3458 (.2127)	-.2934 (.2353)
Indicator: ATT is cable company (3)	.0158 (.2262)	-.0017 (.2541)
Indicator: Adelphia Comm is cable company (2)	.4883 (.2943)	.3837 (.2733)
Indicator: Adelphia Comm is cable company (3)	.6111 (.3121)	.5219 (.3065)
Indicator: Cablevision is cable company (2)	.1905 (.5368)	-.1912 (.5596)
Indicator: Cablevision is cable company (3)	1.215 (.5829)	.7400 (.6193)
Indicator: Charter Comm is cable company (2)	-.1807 (.2387)	-.1871 (.2196)
Indicator: Charter Comm is cable company (3)	-.0408 (.2539)	-.0685 (.2488)
Indicator: Comcast is cable company (2)	-.4097 (.2601)	-.4034 (.2755)
Indicator: Comcast is cable company (3)	.1427 (.2755)	.0989 (.3002)
Indicator: Cox Comm is cable company (2)	-.6419 (.4302)	-.6336 (.4225)
Indicator: Cox Comm is cable company (3)	-.0398 (.4564)	-.1563 (.4827)
Indicator: Time-Warner is cable company (2)	-.3756 (.2335)	-.3439 (.2281)
Indicator: Time-Warner cable company (3)	.0527 (.2503)	-.0009 (.2597)
Alternative specific constant (2)	1.659 (.3486)	3.185 (1.007)
Alternative specific constant (3)	.6462 (.4725)	2.819 (1.480)
Alternative specific constant (4)	.6583 (.1733)	4.635 (.2193)
Price for income group 2 (1-4)	.0156 (.0021)	
Price for income group 3 (1-4)	.0273 (.0023)	
Price for income group 4 (1-4)	.0299 (.0027)	
Price for income group 5 (1-4)	.0353 (.0029)	
Education level of household (2)	-.0521 (.0173)	
Education level of household (3)	-.1385 (.0203)	
Education level of household (4)	-.2525 (.0308)	
Household size (2)	-.0984 (.0240)	
Household size (3)	-.0155 (.0277)	
Household size (4)	-.0235 (.0363)	
Household rents dwelling (2-3)	-.1494 (.0772)	
Household rents dwelling (4)	-.5470 (.1349)	
Single family dwelling (4)	.1967 (.1023)	
Error components, standard deviation (2-4)	.7775 (.1664)	
Log likelihood at convergence	-13927.40	
Number of observations: 11810		

those from the control function approach. Education induces households to buy less TV reception. Larger households tend not to buy extended-basic cable, and other differences are not significant. Renters tend not to buy cable and satellite as readily as owners. And single-family dwellers tend to purchase satellite reception more readily than households who live in multi-family dwellings.

Table 3 gives price elasticities from the models for each approach. Given that the price coefficients are nearly the same from the two methods, similar elasticities would be expected, except for one issue. In particular, the two methods calculate elasticities at different probabilities for each household. The fixed effects approach calculates elasticities at the probabilities that arise when fixed effects are included in the model, such that forecasted shares equal sample shares in each market. For the control function approach, forecasted shares need not exactly equal the sample shares in each market.[15] As Table 3 indicates, this potential difference does not cause the elasticities to differ, as the two methods provide very similar estimates.

Since we estimate both approaches, we can undertake a test of specification that tells us whether omitting $\tilde{\xi}_{mj} - \lambda_j \hat{\mu}_{mj}$ from the control function specification affects estimates of parameters. The point estimates are almost identical in both cases, and tests for differences in parameter estimates with and without $\tilde{\xi}_{mj} - \lambda_j \hat{\mu}_{mj}$ as an additional regressor fail to reject "no differences" at any reasonable level of significance.

As always with endogeneity, the selection of instruments is an issue. As stated above, we used the product attributes and Hausman-type prices as instruments, which follows the practice adopted in Goolsbee and Petrin. The use of Hausman-type price instruments can be controversial (Bresnahan, 1997; Hausman, 1997b). In our context, these instruments are appropriate if the prices of the same multiple system operator in other markets reflect common costs of the multiple system operator but not common unobserved attributes.

With disaggregate data for several markets, market-level averages of the demographic variables may be valid instruments if they affect market price

[15]The model under the control function approach could be calibrated to each market prior to forecasting, such that market shares equal sample shares under this model also. This is a hybrid approach where fixed effects are not used in estimation but are calculated for forecasting.

Table 3:Estimated Elasticities

	Control Function	Fixed Effects
Price of extended-basic cable		
Antenna-only share	0.96	0.79
Extended-basic cable share	-1.18	-0.97
Premium cable share	0.99	0.88
Satellite share	0.95	0.87
Price of premium cable		
Antenna-only share	0.60	0.52
Extended-basic cable share	0.65	0.57
Premium cable share	-2.36	-2.04
Satellite share	0.64	0.58
Price of satellite		
Antenna-only share	0.43	0.42
Extended-basic cable share	0.48	0.43
Premium cable share	0.48	0.45
Satellite share	-3.79	-3.59

and are uncorrelated with the remaining error term. For example, consider two households that have the same demographics but live in areas where the aggregate demographics are different. Part of the price difference between the two areas may be attributable to the difference in aggregate demographics. If this part is not correlated with unobserved attributes, aggregate demographics are valid instruments.

We re-estimated the models with the demographic variables included as extra instruments. The base price coefficient under the fixed effects approach dropped from -.0922 to -.0739, which, while noticeable, is not a statistically significant difference. With the control function approach, the base price coefficient was very similar with or without the demographics as instruments. We also re-estimated both models without using the Hausman-type prices as instruments, that is, just using aggregate demographics as instruments. In every case these estimates were similar across the models, and tests for differences cannot reject the “no differences” hypothesis.

4 Conclusion

With disaggregate (i.e., customer-level) models of demand, price is not necessarily endogenous in the traditional sense, since the demand of the customer does not usually affect market price. However, omitted product attributes can result in correlation between the price and the unobserved portion of utility: the market mechanism causes the price to be higher for products that display desirable attributes that are observed by consumers but not measured by the econometrician. Since these attributes affect demand at the customer level, price can be correlated with the error term even in disaggregate models.

One popular solution to this econometric problem, proposed by BLP, is to include product-market fixed effects. We suggest an alternative approach based on control functions. The basic idea is to include extra variables in the estimation equation that condition out (*i.e.*, control for) the part of the error that is correlated with the regressors. We implement it in a discrete choice environment where price endogeneity often raises econometric concerns.

For general cases the assumptions under the control function approach differ from those for the fixed effects approach, with neither set of assumptions necessarily being nested within the other. For many questions the control function approach can be implemented in the most popular statistical packages, while the fixed effects approach usually requires the use of a lower-level programming language. Additionally, the control function approach is available in some situations for which the BLP approach is not. In this paper we describe both approaches and apply them to disaggregate data on customers' choice among TV options. We find that the two approaches provide very similar estimates, including a substantial increase in the price response when either correction is applied.

Appendix

We obtained information on households' television choices, the characteristics of households, and the prices and attributes of the cable franchise serving the household's geographic area. This information comes from two sources, the Forrester Technographics 2001 survey and Warren Publishing's 2001 Television and Cable Factbook. The Forrester survey was designed to be a nationally representative sample of households. It asks respondents about their ownership and use of various electronic and computer-related goods. To these data we match information about cable franchises from Warren Publishing's 2001 Factbook, which is the most comprehensive reference for cable system attributes and prices in the industry.

To minimize sampling error in market shares, we restricted our analysis to markets where there are at least 30 respondents in the Forrester survey. This screen yields 300 cable franchise markets with a total of almost 30,000 households. We randomly choose 172 of these 300 markets, so as to reduce the number of fixed effects that needed to be estimated. From these 172 markets, we randomly selected 11810 households, oversampling those households from smaller markets (again, to minimize sampling error). These 11810 households are used in the estimation with weights equal to the inverse of their probability of being sampled.

As stated in the body of the paper, the alternatives in the discrete choice model are: expanded basic cable, premium cable (which can only be purchased bundled with expanded basic), Direct Broadcast Satellite, and no multi-channel video (i.e., local antenna reception only). In the Forrester survey, respondents reported whether they have cable or satellite, and the amount they spend on premium television. We classified respondents as having premium if they reported that they have cable and spend more than \$10 per month on premium viewing, which is the average price of the most popular premium channel, HBO. We classified respondents as choosing expanded basic if they reported that they have cable and they spend less than \$10 per month on premium viewing.

The survey provides various demographic characteristics, including family income, household size, education, and type of living accommodations. It also includes an identifier for the household's television market, which can be used to link households to their cable franchise provider.

The cable franchise market of each surveyed household was matched to cable system information from Warren Publishing's 2001 Television and Cable Factbook. The attributes we include are the channel capacity of a cable system, the number of pay channels available, whether pay per view is available from that cable franchise, the price of basic plus expanded basic service, and the price of premium service. We also obtain from the Factbook the number of over-the-air channels available in the franchise market. Finally, for the price of satellite, we use \$50 per month plus an annual \$100 installation and equipment cost.

References

Bajari, P. and C. Benkard (2001), 'Demand estimation with hetergeneous consumers and unobserved product characteristics: A hedonic approach', Working Paper, Department of Economics, Stanford University.

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.

Berry, S., J. Levinsohn and A. Pakes (2001), 'Differentiated products demand system from a combination of micro and macro data: The new car market', Working Paper, National Bureau of Economic Research.

Bhat, C. (2001), 'Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model', *Transportation Research B* **35**, 677–693.

Blundell, R. and J. Powell (2001), 'Endogeneity in semiparametric binary response models', Working Paper, Department of Economics, University College London.

Bresnahan, T. (1997), 'The apple-cinnamon cheerios war: Valuing new goods, identifying market power, and economic measurement', Working Paper, Stanford University.

Brownstone, D. and K. Train (1999), 'Forecasting new product penetration with flexible substitution patterns', *Journal of Econometrics* **89**, 109–129.

Chintagunta, P., J. Dubé and K. Goh (2002), 'Targeted pricing and the estimation of consumer choice models in the presence of unmeasured product characteristics', Working Paper, Graduate School of Business, University of Chicago.

Crawford, G. (2000), 'The impact of the 1992 cable act on household demand and welfare', *RAND Journal of Economics* **31**, 422–449.

Goolsbee, A. and A. Petrin (2002), 'The consumer gains from direct broadcast satellites and the competition with cable tv', Working Paper, Graduate School of Business, University of Chicago.

Halton, J. (1960), 'On the efficiency of evaluating certain quasi-random sequences of points in evaluating multi-dimensional integrals', *Numerische Mathematik* **2**, 84–90.

Hausman, J. (1978), 'Specification tests in econometrics', *Econometrica* **46**, 1251–1272.

Hausman, J. (1997*a*), 'Reply to professor bresnahan', Working Paper, MIT.

Hausman, J. (1997*b*), Valuation of new goods under perfect and imperfect competition, *in* R.Gordon and T.Bresnahan, eds, 'The Economics of New Goods', University of Chicago Press, Chicago.

Heckman, J. (1978), 'Dummy endogenous variables in a simultaneous equation system', *Econometrica* **46**, 931–959.

Heckman, J. and R. Robb (1985), 'Alternative methods for evaluating the impacts of interventions: An overview', *Journal of Econometrics* **30**, 239–267.

Hensher, D. (2001), 'The valuation of commter travel time savings for car drivers in new zealand: Evaluating alternative model specifications', *Transportation* **28**, 101–118.

Karaca-Mandic, P. (2002), 'Complementarities in technology adoption: The case of consumer adoption of dvd console players', Working Paper, Department of Economics, University of California, Berkeley.

Karaca-Mandic, P. and K. Train (2002), 'Standard error correction in two-step estimation with nested samples', Working Paper, Department of Economics, University of California, Berkeley.

Madansky, A. (1964), 'Instrumental variables in factor analysis', *Psychometrika* **29**(2), 105–113.

McFadden, D. (1974), Conditional logit analysis of qualitative choice behavior, *in* P.Zarembka, ed., 'Frontiers in Econometrics', Academic Press, New York, pp. 105–42.

McFadden, D. (1978), Modeling the choice of residential location, *in* A.Karlqvist, L.Lundqvist, F.Snickars and J.Weibull, eds, 'Spatial Interaction Theory and Planning Models', North-Holland, Amsterdam, pp. 75–96.

Munizaga, M. and R. Alvarez-Daziano (2001), 'Mixed logit versus nested logit and probit', Working Paper, Departmento de Ingeniera Civil, Universidad de Chile.

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

Petrin, A. (2002), 'Quantifying the benefits of new products; the case of the minivan', *Journal of Political Economy* **110**, 705–729.

Rivers, D. and Q. Vuong (1988), 'Limited information estimators and exogeneity tests for simultaneous probit models', *Journal of Econometrics* **39**, 347–366.

Smith, R. and R. Blundell (1986), 'An exogeneity test for a simultaneous equation tobit model with an application to labor supply', *Econometrica* **54**, 679–686.

Train, K. (1986), *Qualitative Choice Analysis*, MIT Press, Cambridge, MA.

Train, K. (1998), 'Recreation demand models with taste variation', *Land Economics* **74**, 230–239.

Train, K. (2000), 'Halton sequences for mixed logit', Working Paper No. E00-278, Department of Economics, University of California, Berkeley.

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

Trajtenberg, M. (1989), 'The welfare analysis of product innovations, with an application to computed tomography scanners', *Journal of Political Economy* **97**, 444–479.

Trajtenberg, M. (1990), *Economic Analysis of Product Innovation: The Case of CT Scanners.*, Harvard University Press.

Villas-Boas, J. and R. Winer (1999), 'Endogeneity in brand choice models', *Management Science* **45**, 1324–1338.