

## A Control Function Approach to Endogeneity in Consumer Choice Models

Amil Petrin

Associate Professor, Department of Economics  
University of Minnesota, Twin Cities  
4-101 Hanson Hall  
Minneapolis, MN 55455  
Voice: 612-625-0145  
Fax: 612-624-0209  
petrin@umn.edu

and

Kenneth Train\*

Adjunct Professor, Department of Economics  
University of California, Berkeley  
Berkeley CA 94720-3880  
Voice: 415-291-1023  
Fax: 415-291-1020  
train@econ.berkeley.edu

\* The authors thank the *JMR* AE and two reviewers for their useful comments, and gratefully acknowledge support from the National Bureau of Economic Research.

### Abstract

Endogeneity arises for numerous reasons in models of consumer choice. It leads to inconsistency with standard estimation methods that maintain independence between the model's error and the included variables. The authors describe a control function approach for handling endogeneity in choice models. Observed variables and economic theory are used to derive controls for the dependence between the endogenous variable and the demand error. The theory points to the relationships that contain information on the unobserved demand factor, like the pricing equation and the advertising equation. The authors' approach is an alternative to the commonly-used Berry, Levinsohn, Pakes (1995) product-market controls for unobserved quality. The authors apply both methods to examine households' choices among television options, including basic and premium cable packages, where unobserved attributes, such as quality of programming, are expected to be correlated with price. Without correcting for endogeneity, aggregate demand is estimated to be upward-sloping, suggesting omitted attributes are positively correlated with demand. Both the control function method and the product-market controls method produce downward-sloping demand estimates that are very similar.

Keywords: Customer choice, endogeneity, advertising, price effects, econometric models.

There are several discrete choice demand settings where researchers have shown that factors not included in the analysis are correlated with the included factors, violating the standard independence assumption for consistency (see e.g. Bass (1969) and Berry (1994)). In these cases the estimated impact of the observed factor on demand captures not only that factor's effect but also the effect of the unobserved factors that are correlated with it. For example, products with higher quality usually will have higher prices both because the attributes are costly to provide and because they raise demand. When some product attributes are either not observed by the researcher or are difficult to measure, such as stylishness of design, estimated price elasticities will be biased in the positive direction.<sup>1</sup>

The problem is often exacerbated by the difficulty of signing this bias. Consider estimated price elasticities with unobserved advertising. Optimizing firms maximize profits with respect to both price and advertising, so they cannot generally be independent. Firms might raise the price of their products when they advertise if they believe that it stimulates demand. Alternatively firms may lower price when they advertise, e.g. as a part of a sale. The possibility of either case makes the sign of the bias ambiguous.<sup>2</sup>

In this paper we propose a control function method for alleviating bias in discrete choice demand settings.<sup>3</sup> The approach includes extra variables in the empirical specification to condition out the variation in the unobserved factor that is not independent of the endogenous variable. We derive these controls using economic theory to point to alternative equations that contain information on the unobserved demand factor. While our empirical application focusses on the pricing equation, any equation that contains information on relevant unobserved factors may be available for use. For

example, we anticipate that researchers will explore the advertising equation, which is also impacted by the unobserved demand factor.

The most widely used bias-correction method in discrete choice demand settings is the “product-market” control approach developed by Berry (1994) and Berry, Levinsohn and Pakes (1995) (BLP) for market-level data, and then extended to consumer-level data (Berry, Levinsohn and Pakes, 2004; Goolsbee and Petrin, 2004). The approach appeals to the aggregate demand equations as a source of information on the unobserved demand factor, and has been applied to consumers' choice among TV options (Crawford, 2000; Goolsbee and Petrin (2004)), minivans (Petrin, 2002), and grocery goods (Nevo, 2001), Chintagunta, Dubé, and Goh, 2005), to name only a few.

Our control function approach provides a useful alternative to the BLP approach. The control function approach is both easier to estimate and is available in some situations in which the BLP estimator is not valid. For example, the BLP approach is not consistent in settings where there are zero, one, or just a small number of purchase observations per product because it requires that market shares be observed with relatively little sampling error (see Berry, Linton, and Pakes (2004)). The BLP approach is also not available for many recently developed empirical demand models, which either maintain assumptions that are not consistent with the BLP setting or are sufficiently complicated to preclude estimating the BLP controls (e.g Hendel and Nevo (2006), Bajari *et al.* (2007), and Fox (2008)). In contrast, our control function approach simply adds new regressors to the demand specification, making it available in all of these settings.

Either approach is applicable in our empirical application, and so we estimate both for comparison. We also provide more discussion relating the approaches in the BLP

estimation and results section.

Other methods relating to endogeneity in demand settings have been developed. Louviere *et al.* (2005) describe the various manifestations of endogeneity in marketing contexts and the implications for estimation. Kuksov and Villas-Boas (2007), building on earlier work by Villas-Boas and Winer (1999), describe methods for testing for endogeneity. A maximum likelihood approach has been developed by Villas-Boas and Winer (1999) and Gupta and Park (2008). Bayesian methods for handling endogeneity have been developed by Yang, Chen and Allenby (2003) and Jiang, Manchanda, and Rossi (2007).

In the following sections, the control function approach is described, example specifications are given, the relation to pricing behavior is discussed, and an application to households' choices among TV options is provided as illustration.

### *MODEL*

Consumer  $n$  chooses one of the  $J_n$  competing alternatives. The utility that the consumer obtains from alternative  $j$  is

$$(1) \quad U_{nj} = V(y_{nj}, x_{nj}, \beta_n) + \varepsilon_{nj}$$

where  $y_{nj}$  is the observed endogenous variable,  $x_{nj}$  is a vector of observed exogenous variables that affect the utility derived from choice  $j$ ,  $\beta_n$  are parameters that represent the tastes of consumer  $n$ , and the unobserved utility is denoted  $\varepsilon_{nj}$ .<sup>4</sup> The endogenous variable might be price, advertising, travel time, or whatever is relevant in the context. The econometric problem arises because  $\varepsilon_{nj}$  is not independent of  $y_{nj}$ , as maintained by

standard estimation techniques.

The idea behind the control function correction is to derive a proxy variable that conditions on the part of  $y_{nj}$  that depends on  $\varepsilon_{nj}$ . If this can be done then the remaining variation in the endogenous variable will be independent of the error and standard estimation approaches will again be consistent.

In this discrete choice context the approach posits that  $y_{nj}$  can be written as a function of all exogenous variables entering utility for any of the choices, denoted  $x_n$ , the variables  $z_n$  that do not enter utility directly but that do impact  $y_{nj}$  (typically the instruments), and a vector of  $J$  unobserved terms  $\mu_n$ :

$$(2) \quad y_{nj} = W(x_n, z_n, \mu_n).$$

The approach maintains that  $\mu_n$  and  $\varepsilon_{nj}$  are independent of  $x_n$  and  $z_n$  but are not independent of each other. This equation illustrates the source of the dependence between  $y_{nj}$  and  $\varepsilon_{nj}$ , as  $\mu_n$  impacts  $y_{nj}$  and is also not independent of  $\varepsilon_{nj}$ .

The key to the control function approach is to note that, under the maintained assumptions, *conditional on  $\mu_n$ ,  $\varepsilon_{nj}$  is independent of  $y_{nj}$* . The feasibility of the control function approach in any setting will be determined by whether the practitioner is able to recover  $\mu_n$  so it can be conditioned upon when the parameters are estimated.<sup>5</sup>

We analyze the control function case when  $y_{nj}$  is additive in its observed and unobserved covariates. A special case that is illustrative is when there is a single unobserved factor  $\mu_{nj}$  for each choice  $j$ :

$$(3) \quad y_{nj} = W(x_n, z_n; \gamma) + \mu_{nj}.$$

where we make explicit  $\gamma$  the parameters of this function. With additivity and the independence assumptions, the controls  $\mu_{nj}$  are straightforward to recover using any

standard estimator (like OLS). The question becomes how one enters the new controls into the utility function to condition out the dependence between  $y_{nj}$  and  $\varepsilon_{nj}$ .

One approach enters  $\mu_{nj}$  in a flexible manner so as to condition out any function of it. Decomposing  $\varepsilon_{nj}$  into the part that can be explained by a general function of  $\mu_{nj}$  and the residual yields:

$$(4) \quad \varepsilon_{nj} = CF(\mu_{nj}; \lambda) + \tilde{\varepsilon}_{nj}.$$

where  $CF(\mu_{nj}; \lambda)$  denotes the control function with parameters  $\lambda$ . The simplest approximation is to specify the control function as linear in  $\mu_{nj}$ , in which case the control function is  $CF(\mu_{nj}; \lambda) = \lambda \mu_{nj}$ ,  $\lambda$  is a scalar, and utility is given as

$$(5) \quad U_{nj} = V(y_{nj}, x_n, \beta_n) + \lambda \mu_{nj} + \tilde{\varepsilon}_{nj}$$

Alternatively one could allow for a polynomial approximation, adding higher-order terms of  $\mu_{nj}$  and the necessary additional parameters.

More generally, one might want to condition on the entire vector of controls  $\mu_n$  for any choice  $j$  when calculating the control function. In this case we have

$$(6) \quad \varepsilon_{nj} = CF(\mu_n; \lambda) + \tilde{\varepsilon}_{nj}.$$

which can be approximated to first order with a vector of parameters  $CF(\mu_n; \lambda) = \lambda' \mu_n$ . Again, higher-order terms are straightforward to add, although parameters increase rapidly in the number of alternative choices.

Given the researcher's chosen control function specification, we then have:

$$(7) \quad U_{nj} = V(y_{nj}, x_n, \beta_n) + CF(\mu_n, \lambda) + \tilde{\varepsilon}_{nj}.$$

Conditional on  $\mu_n$ , the probability that consumer  $n$  chooses alternative  $i$  is equal to

$$(8) \quad P_{ni} = \int I(U_{ni} > U_{nj} \quad \forall \quad j \neq i) f(\beta_n, \tilde{\varepsilon}_n) d\beta_n \quad d\tilde{\varepsilon}_n$$

where  $f(\cdot)$  is the joint density of  $\beta_n$  and  $\tilde{\varepsilon}_n$  and  $I(\cdot)$  is the indicator function. All that remains to complete the specification is a distributional assumption applied to  $f(\cdot)$ .

The usual approach is to choose specific functional forms for the distribution of  $\beta_n$  and  $\varepsilon_{nj}$  (e.g. normal or logit), although these are typically hard to motivate with economic theory. They are almost always chosen to be independent of each other, and we maintain that assumption here, which implies in our setting that  $\beta_n$  and  $\tilde{\varepsilon}_{nj}$  are independent, because conditioning on  $\mu_n$ , cannot induce dependence. In our application we use normal and logit, although researchers can use whatever assumptions they desire to suit their setup. As in any application, checking the robustness to distributional assumptions is important.

The model is estimated in two steps. First, the endogenous variable is regressed on observed choice characteristics and the instruments. The residuals of this regression are retained and used to calculate the control function. Second, the choice model is estimated with the control function entering as an extra variable or variables.

Since the second step uses an estimate of  $\mu_n$  from the first step, as opposed to the true  $\mu_n$ , the asymptotic sampling variance of the second-step estimator needs to take this extra source of variation into account. Either the bootstrap can be implemented, or the standard formulas for two-step estimators can be used (Murphy and Topel, 1985; Newey and McFadden, 1994). Karaca-Mandic and Train (2003) derive the specific form of these formulas that is applicable to the control function approach. As they note, the bootstrap and asymptotic formulas provide very similar standard errors for the application that we describe in our empirical results.



### PARAMETRIC FUNCTIONAL FORMS

We consider several parametric forms for the errors in both equations. These parametric forms lead to direct parametric forms for the control function itself and the distribution of the demand residuals conditional on the controls. While they need not be maintained, they do provide an alternative to entering  $\mu_n$  flexibly in utility and then choosing a distributional assumption for  $\tilde{\varepsilon}_{nj}$ .

*Example 1: Jointly normal errors, independent over j*

Suppose  $\mu_{nj}$  and  $\varepsilon_{nj}$  are jointly normal for each  $j$  and iid over  $j$ . Then

$$(9) \quad CF(\mu_n; \lambda) = E(\varepsilon_{nj} \mid \mu_n) = \lambda \mu_{nj}$$

for each  $j$  and the deviations  $\tilde{\varepsilon}_{nj} = \varepsilon_{nj} - CF(\mu_n; \lambda)$  are independent of  $\mu_{nj}$  and all other regressors. Thus, the control function for each alternative is the residual from the endogenous variable regression interacted with  $\lambda$ , the one coefficient to be estimated.

Utility is

$$(10) \quad U_{nj} = V(y_{nj}, x_{nj}, \beta_n) + \lambda \mu_{nj} + \tilde{\varepsilon}_{nj}$$

where  $\tilde{\varepsilon}_{nj}$  is iid normal with zero mean.

If  $\beta_n$  is fixed, then the model is an independent probit with the residual entering as an extra variable. If  $\beta_n$  is random, then the model is a mixed independent probit, mixed over the density of  $\beta_n$  (Train, 2003, Chs. 5 and 6 on probit and mixed logit.) It is

important to note, however, that the scale of the estimated model differs from that of the original model. In particular,  $\text{Var}(\tilde{\varepsilon}_{nj}) < \text{Var}(\varepsilon_{nj})$ , such that normalizing by setting  $\text{Var}(\tilde{\varepsilon}_{nj}) = 1$  raises the magnitude of coefficients relative to the normalization  $\text{Var}(\varepsilon_{nj}) = 1$ .

*Example 2: Extreme value and joint normal error components, independent over j*

The previous example can be modified to generate a mixed logit, which has the same normalization for scale in the original and estimated model. This is one of the specifications utilized by Villas-Boas and Winer (1999, section 2). Let  $\varepsilon_{nj} = \varepsilon_{nj}^1 + \varepsilon_{nj}^2$  where  $\varepsilon_{nj}^1$  and  $\mu_{nj}$  are jointly normal, and  $\varepsilon_{nj}^2$  is iid extreme value for all j. Then utility with the control function is

$$(11) \quad U_{nj} = V(y_{nj}, x_{nj}, \beta_n) + \lambda \mu_{nj} + \sigma \eta_{nj} + \varepsilon_{nj}^2$$

where  $\eta_{nj}$  is iid standard normal. The model is a mixed logit, with mixing over the error components  $\eta_{nj}$ , whose standard deviation  $\sigma$  is estimated, as well as over the random elements of  $\beta_n$ . The scale in the original utility is normalized by setting the scale of the extreme value distribution for  $\varepsilon_{nj}^2$ .

*Example 3: Extreme value and joint normal error components, correlation over j*

The generalization is straightforward conceptually, but increases the number of parameters considerably. Let  $\varepsilon_n^1$  and  $\mu_n$  be jointly normal with zero mean and covariance

$\Omega$ . This covariance matrix is  $2J \times 2J$  and is composed of submatrices labeled  $\Omega_{\mu\mu}$ ,  $\Omega_{\mu\epsilon}$ ,  $\Omega_{\epsilon\epsilon}$

Then  $CF(\mu_n, \lambda) = E(\epsilon_n^1 | \mu_n) = \Lambda\mu_n$ , where the elements of matrix  $\Lambda$  are related to the elements of  $\Omega$ :  $\Lambda = \Omega_{\mu\epsilon} \Omega_{\mu\mu}^{-1}$ . Stacked utilities become:

$$(12) \quad U_n = V(y_n, x_n, \beta_n) + \Lambda\mu_n + \Gamma\eta_n + \epsilon_n^2$$

where  $\eta_n$  is now a vector of  $J$  iid standard normal deviates and  $\Gamma$  is the lower-triangular Choleski matrix of  $\Omega_{\epsilon\epsilon} - \Omega_{\mu\epsilon} \Omega_{\mu\mu}^{-1} \Omega_{\mu\epsilon}$ . In this case, the residuals for each alternative enter the utility of all alternatives, and the mixing is over a set of  $J$  normal error components. Villas-Boas and Winer (1999, section 3) generalize this specification further by allowing  $\epsilon_n^2$  to be correlated over alternatives, specifying it to be normally distributed instead of extreme value to accommodate this correlation.

### *PRICING BEHAVIOR AND THE CONTROL FUNCTION APPROACH*

Consider consumers' choice among products where the endogenous variable  $y_{nj}$  is price  $p_{nj}$ . We investigate the control function approach using some variant of the controls suggested above with both marginal cost and monopoly pricing. The utility that consumer  $n$  obtains from product  $j$  is specified as in example 2:

$$(13) \quad U_{nj} = V(p_{nj}, x_{nj}, \beta_n) + \epsilon_{nj}^1 + \epsilon_{nj}^2$$

where  $\epsilon_{nj}^1$  is correlated with price and  $\epsilon_{nj}^2$  is iid extreme value. Here  $\epsilon_{nj}^1$  might represent unobserved attributes of the product that are not independent of price. Typically, prices vary over people because different people are in different markets.

The marginal cost of product  $j$  in consumer  $n$ 's choice set is denoted  $MC(z_{nj}, v_{nj})$ ,

where are  $z_{nj}$  exogenous variables observed by the analyst and  $v_{nj}$  is unobserved. The observed variables  $z_{nj}$  will typically overlap with  $x_{nj}$  insofar as observed attributes of the product affect both demand and cost.

### *Marginal cost pricing*

Consider marginal cost (MC) pricing and assume MC is separable in the unobserved term. The pricing equation becomes:

$$(14) \quad p_{nj} = MC(z_{nj}, v_{nj}) = W(z_{nj}, \gamma) + v_{nj}$$

where  $\gamma$  are parameters to be estimated. Following example 2, assume that  $\varepsilon_{nj}^1$  and  $v_{nj}$  are jointly normal, iid over  $j$ . Correlation may arise, e.g., because unobserved attributes affect utility as well as costs, thereby entering both  $\varepsilon_{nj}^1$  and  $v_{nj}$ . Given the separability assumption on the unobserved term in the pricing equation, utility becomes:

$$(15) \quad U_{nj} = V(p_{nj}, x_{nj}, \beta_n) + \lambda v_{nj} + \sigma \eta_{nj} + \varepsilon_{nj}^2$$

where  $\eta_{nj}$  is iid standard normal. The same specification is appropriate when there is a constant markup over cost in the determination of prices.

### *Monopoly pricing*

Consider monopoly pricing where price depends on the elasticity of demand as well as marginal cost. The pricing equation for a monopolist is:<sup>6</sup>

$$(16) \quad p_{nj} = (p_{nj} / |e(\varepsilon_n^1)|) + MC(z_{nj}, v_{nj})$$

where  $e(\varepsilon_n^1)$  is the elasticity of demand. This elasticity depends on all factors that affect demand, including attributes that are not observed by the analyst. The elasticity is written as a function of  $\varepsilon_n^1$  in order to explicitly denote this dependence.

Now suppose that the analyst estimates the price equation:

$$(17) \quad p_{nj} = W(x_{nj}, z_{nj}, \gamma) + \mu_{nj}$$

$\varepsilon_n^1$  enters the pricing equation in a non-separable manner, suggesting the additively separable  $\mu_{nj}$  will not fully condition out the entire dependence of  $p_{nj}$  on  $\varepsilon_{nj}^1$ . Any remaining dependence will bias the estimated price elasticity.

In this situation Villas-Boas (2007) has suggested working in the reverse direction. Instead of specifying the joint distribution of  $\langle v_n, \varepsilon_n^1 \rangle$ , and then deriving the implications for the distribution of  $\langle \mu_n, \varepsilon_n^1 \rangle$ , Villas-Boas has shown that, if the price of each product is strictly monotonic in its marginal cost, then there exists a distribution of  $\langle v_n, \varepsilon_n^1 \rangle$  and a marginal cost function that is consistent with any given distribution of  $\langle \mu_n, \varepsilon_n^1 \rangle$ . This result implies that the analyst can specify a distribution of  $\varepsilon_n^1$  conditional  $\mu_n$ , as needed for the control function approach, and know that there is some distribution of  $v_n$  and  $\varepsilon_n^1$  that gives rise to it.

#### *EMPIRICAL MODEL AND DATA*

As illustration we apply the control function approach to households' choice of television reception options. The specification and data are similar to those of Goolsbee and Petrin (2004), who applied the BLP approach. By utilizing a situation where both approaches can be applied, we are able to compare results.

Households are considered to have four alternatives for TV: (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 endogeneity of price, which arises because unobserved attributes of cable TV like the quality of programming are not independent of price.

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.

Table 1 provides information about the sampled households and the service options that are available to them. Nearly 85 percent of the sample lives in single family dwellings, and average income is about \$62,000. The most popular TV option is basic and extended cable, which is chosen by 45 percent of the households. Less than a quarter of the households have antenna reception only. The average price for basic and extended cable is about \$28 per month, with this price ranging from \$16 to \$45 (not shown in the table). The additional fee for premium cable is \$40 on average, ranging from \$26 to \$56. More details of the data are given in the Web Appendix.

Since the attributes of the TV alternatives are the same for all households in a geographic market, we add a subscript for markets. Let  $U_{njm}$  be the utility that household  $n$  who lives in market  $m$  obtains from alternative  $j$ . The price of alternative  $j$  in market  $m$

is  $p_{mj}$ , which is not subscripted by  $n$  since it is the same for all households in the market  $m$ . Price is zero for antenna TV, and the price of satellite TV does not vary over markets or households. The price of the two cable options varies over geographic markets, and unobserved attributes of cable service (such as quality of programming) are expected to be correlated with price. The utility of the two cable options ( $j = 2,3$ ) is specified as in example 2:

$$(18) \quad U_{njm} = V(p_{mj}, x_{mj}, \beta_n) + \varepsilon_{nj}^1 + \varepsilon_{nj}^2$$

where  $\varepsilon_{nj}^1$  is correlated with price,  $\varepsilon_{nj}^2$  is iid extreme value, and  $x_{nj}$  captures exogenous observed attributes. Utility for the two options with constant price ( $j = 1,4$ ) is the same but without the correlated error component  $\varepsilon_{nj}^1$ . Price for the cable options is specified as linear in instruments plus a separable error:

$$(19) \quad p_{mj} = \gamma z_{mj} + \mu_{mj}$$

We specify  $\mu_{mj}$  and  $\varepsilon_{nj}^1$  for  $j = 2,3$  to be jointly normal, independent over  $j$ . Utility with the control function for alternative  $j = 2,3$  is then:

$$(20) \quad U_{njm} = V(p_{mj}, x_{mj}, \beta_n) + \lambda \mu_{mj} + \sigma_j \eta_{nj} + \varepsilon_{nj}^2$$

where  $\eta_{nj}$  is standard normal.

To complete the model, we specify  $V(\cdot)$  as:

$$(21) \quad V(p_{mj}, x_{mj}, \beta_n) = \alpha p_{mj} + \sum_{g=2-5} \theta_g p_{mj} d_{gn} + \tau x_{mj} + \delta_j k_j + \kappa k_j s_n + \varphi \omega_n c_j.$$

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  $\alpha$  while that for a household in group  $g > 1$  is  $\alpha + \theta_g$ . The nonprice attributes  $x_{mj}$  enter with fixed coefficients. The alternative-specific constant for

alternative  $j$  is  $k_j$ . These constants are entered directly and also interacted with demographic variables,  $s_n$ .

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  $\omega_n$  is an iid standard normal deviate. The coefficient  $\phi$  is the standard deviation of this error component, reflecting the degree of correlation among the non-antenna alternatives.

The choice probability therefore takes the form of a mixed logit (Train, 1998; Brownstone and Train, 1999), with the mixing over the distribution of the error components:

$$(22) \quad P_{ni} = \int \frac{e^{v_{ni}(\eta_2, \eta_3, \omega)}}{\sum_{j=1}^4 e^{v_{nj}(\eta_2, \eta_3, \omega)}} \phi(\eta_2)\phi(\eta_3)\phi(\omega) d\omega d\eta_3 d\eta_2$$

where  $\phi(\cdot)$  is the standard normal density and

$$(23) \quad V(\eta_j, \omega) = \alpha p_{mj} + \sum_{g=2-5} \theta_g p_{mj} d_{gn} + \tau x_{mj} + \delta_j k_j + \kappa k_j s_n + \phi \omega c_j + \lambda \mu_{mj} + \sigma_j \eta_j .$$

The integral is approximated through simulation: a value of  $\eta_2$ ,  $\eta_3$ , and  $\omega$  is drawn from their standard normal densities, the logit formula is calculated for this draw, the process is repeated for numerous draws, and the results are averaged. 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).



## *RESULTS*

The first step of the approach is to estimate the pricing functions to recover the residuals entering the control functions in the choice model. The price in each market was regressed against the product attributes listed in Table 2 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  $m$ .<sup>7</sup> 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 demand shocks (like unobserved advertising). A separate instrument is created for the price of extended-basic cable and the price of premium cable and separate regressions were run for extended-basic price and premium price using all instruments in each equation.<sup>8</sup>

The residuals from these regressions enter without transformation in the mixed logit model; that is, the control functions are a coefficient times the product-market residual, which is the first and simplest specification proposed from the model section. Specifically, the residual from the extended-basic cable price regression enters the extended-basic cable alternatives, and similarly for the premium cable.

Table 2 gives the estimated parameters. The variables are listed in three groups: those that vary over markets but not over consumers in each market, those that vary over consumers in each market, and the extra variables that are included to correct for endogeneity. The first column gives the model without any correction for the correlation between price and omitted attributes; utility is the same as specified above except that the

residuals,  $\hat{\mu}_{mj}$ , and induced error components,  $\eta_j$ , are not included. The second column applies the control function approach by including the residuals and error components.

Without correction, the base price coefficient  $\alpha$  is estimated to be -0.0202. As stated above, the price coefficient is allowed to differ by income group; the estimated price coefficient differential by income group,  $\theta_g$ , are given in the second panel of the table since these variables vary over households in each market. For the second income group, the estimated price coefficient is the base of -0.0202 plus the differential of .0149, for a estimated price coefficient of -.0053. Note however that for income groups 3-5, which comprise the majority of households, the estimated differential exceeds the base in magnitude, such that the estimated price coefficients are positive. This result contradicts the expectation of downward-sloping demand and renders the model implausible for predictive purposes and useable for welfare analysis (since welfare analysis assumes a negative price coefficient.)

Inclusion of the control functions adjusts the estimated price coefficients in the expected way. A significantly negative price coefficient is obtained for all income groups as the base coefficient estimate increases almost fivefold to -0.10. Price elasticities decrease 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.

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.

Neither of the error components is statistically significant, and the hypothesis that both have zero standard deviations cannot be rejected at any meaningful level of confidence. This result might imply that the residuals capture the market-specific unobserved attributed of cable service completely or perhaps reflects the empirical difficulty, discussed above, of estimating alternative-specific normal error components when each alternative also has an iid extreme value error component.

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, perhaps because 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 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.

Fewer coefficients are significant in the model with correction for endogeneity than uncorrected. This result is expected, since the correction for endogeneity is attempting to obtain more information from the data (namely, the relation of unobserved factors to price, and well as the relation of observed factors to demand.) Stated alternatively, the uncorrected model gives a false sense of precision by assuming that price is independent of unobserved factors, when in fact price is related to these factors. Interestingly, all of the coefficients that become insignificant with correction, when they were significant without correction, are for variables that vary over markets but not over consumers in each market. This pattern reflects the fact that unobserved attributes that are correlated with price vary over markets but not consumers within each market, since price itself only varies over markets.

The appropriate control function and distribution for  $\tilde{\varepsilon}_{jn}$  is a specification issue. We tried other specifications, including: both residuals entering in each cable alternative (to allow for correlation across alternatives as in example 3); a series expansion, both signed and unsigned, of the residuals (to allow for the conditional mean not being exactly as given by a joint normal); correlated rather than independent error components; exclusion of one or both of the error components (since they are not significant). These alternative specifications all provided very similar results.

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 are widely used but controversial (Bresnahan, 1997; Hausman, 1997b). With disaggregate demand models, the need for additional instruments is not as stringent as in models with just aggregate data because aggregate demographics do not enter the disaggregate models but do affect market price. They can therefore serve as the extra instruments that are needed for demand estimation.<sup>9</sup>

We re-estimated the model without using the prices in other areas as instruments but including the aggregate demographics. With the control function approach, the estimated price coefficient rose when the Hausman-type prices were removed as instruments. This is the direction of change that would be expected if the prices in other markets incorporated the impact of unobserved demand shocks. The other coefficients were not affected under either approach.

Given the widespread use of the Berry, Levinsohn, and Pakes (1995) approach, we provide a brief comparison with the control function approach, and then we discuss BLP model estimation and results for the same data.

The BLP approach uses the aggregate demand equations to recover the unobserved demand factors by matching observed market shares to those predicted by the model. In contrast, our control function approach is based on looking to different equations for information on the unobserved demand factor, like the pricing or advertising equation.

In most applications the control function approach will be easier to implement than the BLP approach. Often the first step is just a regression and the second is maximum likelihood, so the approach can be estimated with standard software packages such as STATA, SAS (which now has a mixed logit and probit routine), LIMDEP, and Biogeme.<sup>10</sup> The two-step estimator requires one to account for the estimated regressors (as discussed earlier), and the sampling covariance can be estimated by bootstrap with these packages.

One must incorporate a contraction procedure into the estimation routine to implement the BLP estimator. It iteratively calculates the constants that equate predicted and actual shares *at each trial value of the parameters*. This computation is not trivial, especially when consumer-level data is being used in the estimated specification. Because of this computational burden the BLP procedure is to our knowledge still not available in any of the common statistical packages.

Since the BLP approach matches observed to predicted shares in a non-linear

setting, it turns out to be very sensitive to sampling error in market shares, as shown in Berry, Linton, and Pakes (2004). It is not consistent in settings where there are zero, one, or just a small number of purchase observations per product relative to the number of consumers, as is true in some data sets.<sup>11</sup> It also requires all goods to be strict substitutes, something not required by the control function setup.

The BLP approach includes a constant  $\delta_{mj}$  for each alternative in each market. All of the elements of utility that do not vary within a market are subsumed into these constants. The utility specification given above becomes:

$$(24) \quad U_{njm} = \delta_{mj} + \sum_{g=2-5} \theta_g p_{mj} d_{gn} + \kappa k_j s_n + \varphi \omega_n c_j + \varepsilon_{nj}^2$$

The constants are expressed as a function of price and other observed attributes:

$$(25) \quad \delta_{mj} = \alpha p_{mj} + \tau x_{mj} + \delta_j k_j + \varepsilon_{mj}^1.$$

Assuming  $\varepsilon_{nj}^2$  and  $\omega_n$  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 constants for each product-market alternative and without the extra error components that are induced by the control function. The equation for the constants is estimated by instrumental variables since utility is assumed to be linear in  $\varepsilon_{mj}^1$  and  $\varepsilon_{mj}^1$  is correlated with price.

Estimation is performed in two stages, with the first stage being the computationally burdensome one. First the mixed logit model is estimated with constants for each alternative and each market. These constants are recovered by solving for the values that match observed to predicted market shares in each market and for each product at every set of parameter values until the minimum is located. Then these estimated constants are regressed against the product attributes using 3SLS.<sup>12</sup> A separate

equation is used for the extended-basic cable, premium cable, and satellite constants, with the coefficients of the product attributes constrained across equations as in the control function setup (and every characteristics-based setup of which we are aware).

The results are given in Table 3. 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 constants 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 estimated price coefficients. 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  $-0.1003$  obtained with the control function approach. The estimates of  $\theta_g$ , 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.

The demographic coefficients in Table 3 provide similar conclusions as those from the control function approach. Education induces households to buy less TV reception. Larger households tend not to buy extended-basic cable. 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.

Differences appear not to be statistically significant.

Table 4 gives price elasticities from the models for each approach. The two methods give similar elasticities. For example, the same-price elasticity for basic and extended cable is -1.08 with the control function approach and -0.97 under the BLP approach.

### *CONCLUSION*

The concern that price, advertising, or other variables are endogenous has proven to be important in many applications. In this paper we propose a control function approach for handling endogeneity in choice models. It uses observed variables and economic theory to derive controls for the part of the unobserved demand factor that is not independent of the endogenous variable. We use the pricing equation in this paper to derive controls, but we believe there are many other possible equations - like the advertising equation - which also contain information on unobserved demand factors.

The approach provides an alternative to the commonly-used Berry, Levinsohn, Pakes (1995) product-market controls for unobserved quality, which is sensitive to sampling error in market shares, more difficult to estimate - especially with consumer-level data - and not applicable for many recently proposed discrete choice demand estimators.

We apply both methods to examine households' choices among television options, including basic and premium cable packages, where unobserved attributes, such as quality of programming, are expected to be correlated with price. Without correcting for

endogeneity, aggregate demand for each TV option is upward-sloping. The corrected estimates from both the control function method and the product-market controls method produce similar and much more realistic demand elasticities.

## REFERENCES

Bass, Frank (1969), "A Simultaneous Equation Regression Study of Advertising and Sales of Cigarettes," *Journal of Marketing Research*, 6 (3), 291–300.

Bajari, Patrick, Jeremy Fox, Kyoo Il Kim, and Stephen Ryan (2007), "A Simple Nonparametric Estimator for the Distribution of Random Coefficients in Discrete Choice Models," working paper, Department of Economics, University of Minnesota.

Berry, Steve (1994), "Estimating Discrete Choice Models of Product Differentiation," *RAND Journal of Economics*, 25 (2), 242–62.

———, James Levinsohn, and Ariel Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 63 (4), 841–89.

———, ———, and ——— (2004), "Differentiated Products Demand System from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 112 (1), 68–105.

———, Oliver Linton, and Ariel Pakes (2004), "Limits Theorems for Estimating Parameters of Differentiated Product Demand Systems," *Review of Economic Studies*, 71 (3), 613–54.

Bhat, Chandra (2001), "Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model," *Transportation Research B*, 35 (7), 677–93.

Blundell, Richard and James Powell (2004), "Endogeneity in Semiparametric Binary Response Models," *Review of Economic Studies*, 71 (3), 655–79.

Bresnahan, Tim (1997), "The Apple-Cinnamon Cheerios War: Valuing New Goods, Identifying Market Power, and Economic Measurement," working paper, Stanford University.

Brownstone, David and Kenneth Train (1999), "Forecasting New Product Penetration with Flexible Substitution Patterns," *Journal of Econometrics*, 89 (1-2), 109–129.

Chintagunta, Pradeep, Jean-Pierre Dubé, and Khim Yong Goh (2005), "Beyond the Endogeneity Bias: The Effect of Unmeasured Brand Characteristics on Household Level Brand Choice Models," *Management Science*, 51 (5), 832–49.

Crawford, Gregory (2000), "The Impact of the 1992 Cable Act on Household Demand and Welfare," *RAND Journal of Economics*, 31 (3), 422–49.

Fox, Jeremy (2008), "Semiparametric Estimation of Multinomial Discrete Choice Models Using a Subset of Choices," *RAND Journal of Economics*, 38 (4), 1002–1019.

Goolsbee, Austan and Amil Petrin (2004), "The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV," *Econometrica*, 72 (2), 351–82.

Gupta, Sachin and Sungho Park (2008), "A Simulated Maximum Likelihood Estimator for the Random Coefficient Logit Model Using Aggregate Data," *Journal of Marketing Research*, forthcoming.

Halton, John (1960), "On the Efficiency of Evaluating Certain Quasi-Random Sequences of Points in Evaluating Multi-dimensional Integrals," *Numerische Mathematik*, 2 (1), 84–90.

Hausman, Jerry (1978), "Specification Tests in Econometrics," *Econometrica*, 46 (6), 1251–72.

——— (1997a), "Valuation of New Goods Under Perfect and Imperfect Competition," in *The Economics of New Goods*, R. Gordon and T. Bresnahan, eds. Chicago: University of Chicago Press.

——— (1997b), "Reply to Professor Bresnahan," working paper, Department of Economics, MIT.

Heckman, Jerry (1978), "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica*, 46 (4), 931–59.

——— and Richard Robb (1985), “Alternative Methods for Evaluating the Impacts of Interventions: An Overview,” *Journal of Econometrics*, 30 (1-2), 239–67.

Hendel, Igal and Aviv Nevo (2006), “Measuring the Implications of Sales and Consumer Inventory Behavior,” *Econometrica*, 74 (6), 1636–74.

Hensher, David (2001), “The Valuation of Commuter Travel Time Savings for Car Drivers in New Zealand: Evaluating Alternative Model Specifications,” *Transportation*, 28 (2), 101–118.

Imbens, Guido, and Whitney Newey (2008), “Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity,” working paper, Department of Economics, Harvard University.

Louviere, Jordan, et al. (2005), “Recent Progress on Endogeneity in Choice Modeling,” *Marketing Letters*, 16 (3-4), 255–65.

Jiang, Renna, Puneet Manchanda, and Peter Rossi (2007), “Bayesian Analysis of Random Coefficient Logit Models Using Aggregate Data,” working paper, Graduate School of Business, University of Chicago.

Karaca-Mandic, Pinar and Kenneth. Train (2003), “Standard Error Correction in

Two-Step Estimation with Nested Samples,” *Econometrics Journal*, 6 (2), 401–407.

Kuksov, Dmitri and J. Miguel Villas-Boas (2007), “Endogeneity and Individual Customer Choice,” *Journal of Marketing Research*, forthcoming.

Martin, Leslie (2008), “Consumer Demand for Compact Fluorescent Light Bulbs,” working paper, Department of Agricultural and Resource Economics, University of California, Berkeley.

Munizaga, Marcela and Ricardo Alvarez-Daziano (2001), “Mixed Logit Versus Nested Logit and Probit,” working paper, Departamento de Ingeniera Civil, Universidad de Chile.

Murphy, Kevin, and Robert Topel (1985), “Estimation and Inference in Two Step Econometric Models,” *Journal of Business and Economic Statistics*, 3 (4), 370–79.

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

Newey, Whitney and Daniel McFadden (1994), “Large Sample Estimation and Hypothesis Testing,” in *The Handbook of Econometrics*, Volume 4, Z. Griliches and M. Intriligator, eds. Amsterdam: Elsevier Science Publishers, 2111–45.

Petrin, Amil (2002), “Quantifying the Benefits of New Products; The Case of the

Minivan,” *Journal of Political Economy*, 110 (4), 705–729.

Rivers, Douglas and Quang Vuong (1988), “Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models,” *Journal of Econometrics*, 39 (3), 347–66.

Smith, Richard and Richard Blundell (1986), “An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply,” *Econometrica*, 54 (3), 679–86.

Train, Kenneth (1998), “Recreation Demand Models with Taste Differences over People,” *Land Economics*, 74 (2), 230-239.

——— (2000), “Halton Sequences for Mixed Logit,” working paper no. E00-278, Department of Economics, University of California, Berkeley.

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

Villas-Boas, J. Miguel (2007), “A Note on Limited Versus Full Information Estimation in Non-Linear Models,” working paper, Haas School of Business, University of California, Berkeley.



——— and Russell. Winer (1999), “Endogeneity in Brand Choice Models,”

*Management Science*, 45 (10), 1324–38.

Yang, Sha, Yuxin Chen, and Greg Allenby (2003), “Bayesian Analysis of Simultaneous Demand and Supply,” *Quantitative Marketing and Economics*, 1 (3), 251–75.

*FOOTNOTES*

1. Another common bias is that arising when estimating the elasticity of demand for recreational, shopping, theater, etc. activities with respect to travel-time. If individuals with strong tastes for these types of activities select into living close to these activities, then travel time to the desired activity will be negatively correlated with unobserved taste, leading to a negative bias in the elasticity of demand with respect to travel time.
2. Variation in unobserved variables, including unmeasured attributes and advertising, may itself be caused by changes in demand conditions, such as shifts in tastes. In this case, a model that describes the variation in demand and its relation to the unobserved variables would more fully represent the situation.
3. The term “control function” was introduced by Heckman and Robb (1985) in the context of selection models, but the concepts date back at least to Heckman (1978) and Hausman (1978). The method has been applied to a tobit model by Smith and Blundell (1986) and binary probit by Rivers and Vuong (1988). Blundell and Powell (2004) include it in their discussion of semiparametric methods for binary choice.
4. “Observed” and “unobserved” are from the perspective of the practitioner. All terms are observed by the consumer when making the decision.
5. As shown in Imbens and Newey (2008) it is sufficient to condition on any one-to-one function of  $\mu_n$ .
6. The more common form of this equation is  $(p-MC)/p = 1/|e|$ .
7. To our knowledge, there has been no earlier application of the control function

- approach with cross-sectional market data; Villas-Boas and Winer use lagged prices as instruments in their time-series model.
8. The use of alternative instruments is discussed in the robustness section.
  9. 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 is presumably attributable to the difference in aggregate demographics. This part of the price difference provides variation in price over households that can be used for estimation of price response.
  10. Matlab and gauss codes for mixed logit are also available (free) from Train's website at <http://elsa.berkeley.edu/~train/software.html>.
  11. E.g., in many housing datasets, houses are only observed to be purchased zero or one time, violating the consistency condition of the BLP estimator. Another example is Martin's (2008) study of customer's choice between incandescent and compact fluorescent light bulbs (CFLs), where advertising and promotions (such as discount coupons) occurred on a weekly basis and varied over stores, and yet it was common for a store not to sell any CFLs in a given week. With the market defined as a store-week, Martin reports that 65 percent of the market shares were zero for CFLs.
  12. The negative of the number of over-the-air channels enters these equations, since this attribute enters the antenna-only alternative in the model of Table 2 whereas it is now entering the constants for the non-antenna alternatives.

Table 1: Demographic variables and service attributes	
Average income	62,368.
Income groups:	Share:
below \$25,000	19.60
25,000 – 49,999	24.48
50,000 – 74,999	24.39
75,000 – 99,999	17.44
100,000 and above	14.09
Unmarried	31.93
Single family dwelling	84.58
Rent	16.34
Household size:	
1 person	18.88
2 people	39.40
3 people	16.76
4 people	15.39
5 or more people	9.56
Chosen TV option:	
Antenna only	23.14
Basic and extended cable	44.79
Premium cable	20.72
Satellite	11.36
Attributes of service in HH's area	Average
Over-the-air channels	10.7
Basic/extended cable channels	62.9
Additional premium cable channels	5.8
Price for basic and extended cable	27.96

Price for premium cable	39.58
-------------------------	-------

Table 2: 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 CF
	(Standard errors in parentheses)	
<u>Variables that vary over markets but are constant over consumers in each market</u>		
Price, in dollars per month (1-4)	<b>-.0202</b> (.0047)	<b>-.1003</b> (.0471)
Number of cable channels (2,3)	<b>-.0023</b> (.0011)	.0026 (.0039)
Number of premium channels (3)	<b>.0375</b> (.0163)	.0559 (.0382)
Number of over-the-air channels (1)	<b>.0265</b> (.0090)	.0232 (.0152)
Whether pay per view is offered (2,3)	<b>.4315</b> (.0666)	<b>.5992</b> (.1792)
Indicator: ATT is cable company (2)	.1279 (.0946)	-.2072 (.2437)
Indicator: ATT is cable company (3)	.0993 (.1195)	-.2559 (.2737)
Indicator: Adelphia is cable company (2)	<b>.3304</b> (.1224)	.3443 (.2930)
Indicator: Adelphia is cable company (3)	.2817 (.1511)	.2504 (.3400)
Indicator: Cablevision is cable company (2)	<b>.6923</b> (.2243)	.1031 (.3749)
Indicator: Cablevision is cable company (3)	<b>1.328</b> (.2448)	1.015 (.5412)
Indicator: Charter is cable company (2)	.0279 (.1010)	-.0587 (.2259)
Indicator: Charter is cable company (3)	-.0618 (.1310)	-.2171 (.2139)
Indicator: Comcast is cable company (2)	<b>.2325</b> (.1107)	-.1111 (.3694)
Indicator: Comcast is cable company (3)	<b>.5010</b> (.1325)	.2619 (.3210)
Indicator: Cox is cable company (2)	<b>.2907</b> (.1386)	-.0720 (.3314)
Indicator: Cox is cable company (3)	<b>.5258</b> (.1637)	.1678 (.5065)
Indicator: Time-Warner cable company (2)	.1393 (.0974)	-.0902 (.2213)
Indicator: Time-Warner cable company (3)	.2294 (.1242)	-.0462 (.2254)
Alternative specific constant (2)	<b>1.119</b> (.2668)	<b>3.060</b> (1.054)

Alternative specific constant (3)	.1683 (.3158)	2.439 (1.542)
Alternative specific constant (4)	-.2213 (.4102)	4.386 (2.690)
<u>Variables that vary over consumers in each market</u>		
Price for income group 2 (1-4)	<b>.0149</b> (.0024)	<b>.0154</b> (.0026)
Price for income group 3 (1-4)	<b>.0246</b> (.0030)	<b>.0253</b> (.0038)
Price for income group 4 (1-4)	<b>.0269</b> (.0034)	<b>.0271</b> (.0042)
Price for income group 5 (1-4)	<b>.0308</b> (.0036)	<b>.0311</b> (.0040)
Education level of household (2)	<b>-.0644</b> (.0220)	<b>-.0640</b> (.0254)
Education level of household (3)	<b>-.1137</b> (.0278)	<b>-.1129</b> (.0371)
Education level of household (4)	<b>-.1965</b> (.0369)	<b>-.1987</b> (.0384)
Household size (2)	<b>-.0494</b> (.0240)	<b>-.0556</b> (.0283)
Household size (3)	.0160 (.0286)	.0303 (.0421)
Household size (4)	.0044 (.0357)	.0023 (.0447)
Household rents dwelling (2-3)	<b>-.2471</b> (.0867)	<b>-.2719</b> (.0891)
Household rents dwelling (4)	-.2129 (.1562)	-.2008 (.1329)
Single family dwelling (4)	<b>.7622</b> (.1523)	<b>.7790</b> (.2071)
Error component for non-antenna alts, SD (2-4)	.5087 (.6789)	.4994 (.7344)
<u>Terms to correct for endogeneity</u>		
Residual for extended-basic cable price (2)		.0833 (.0481)
Residual for premium cable price (3)		<b>.0929</b> (.0499)
Error component for basic and extended cable, SD (2)		.0488 (1.423)
Error component for premium cable, SD (3)		1.425 (1.142)
Log likelihood at convergence	-14660.84	-14645.21
Number of observations:	11810	11810
Note: Estimates that are statistically significant ( $p < .05$ ) are in bold.		



Table 3: Mixed Logit Model of TV Reception Choice: BLP 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)	
<u>Variables that vary over markets but are constant over consumers in each market</u>		
Price, in dollars per month (1-4)	<b>-.0245</b> (.0091)	<b>-.0922</b> (.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)	<b>.5872</b> (.1326)	<b>.7144</b> (.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)	<b>1.215</b> (.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)	<b>1.659</b> (.3486)	<b>3.185</b> (1.007)
Alternative specific constant (3)	.6462 (.4725)	2.819 (1.480)



Alternative specific constant (4)	<b>.6583</b> (.1733)	<b>4.635</b> (.2193)
<u>Variables that vary over consumers in each market</u>		
Price for income group 2 (1-4)		<b>.0156</b> (.0021)
Price for income group 3 (1-4)		<b>.0273</b> (.0023)
Price for income group 4 (1-4)		<b>.0299</b> (.0027)
Price for income group 5 (1-4)		<b>.0353</b> (.0029)
Education level of household (2)		<b>-.0521</b> (.0173)
Education level of household (3)		<b>-.1385</b> (.0203)
Education level of household (4)		<b>-.2525</b> (.0308)
Household size (2)		<b>-.0984</b> (.0240)
Household size (3)		-.0155 (.0277)
Household size (4)		-.0235 (.0363)
Household rents dwelling (2-3)		<b>-.1494</b> (.0772)
Household rents dwelling (4)		-.5470 (.1349)
Single family dwelling (4)		.1967 (.1023)
Error comp for non-antenna alts, SD (2-4)		<b>.7775</b> (.1664)
Log likelihood at convergence		-13927.40
Number of observations		11810
Note: Estimates that are statistically significant ( $p < .05$ ) are in bold.		

Table 4: Estimated Elasticities

	CF	BLP
Price of extended-basic cable		
Antenna-only share	.97	.79
Extended-basic cable share	-1.08	-.97
Premium cable share	.76	.88
Satellite share	1.02	.87
Price of premium cable		
Antenna-only share	.48	.52
Extended-basic cable share	.50	.57
Premium cable share	-1.83	-2.04
Satellite share	.48	.58
Price of satellite		
Antenna-only share	.50	.42
Extended-basic cable share	.40	.43
Premium cable share	.37	.45
Satellite share	-3.77	-3.59

**WEB APPENDIX****A Control Function Approach to Endogeneity  
in Consumer Choice Models**

Amil Petrin and Kenneth Train

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 constants 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.