Journal of Agricultural and Resource Economics, 18(2): 141–146

On the Estimation of Separable Demand Models

David L. Edgerton

Alternative stochastic specifications of conditional demand models are considered. The results of LaFrance concerning the inconsistency of least squares are supported, but the class of models that allow standard instrumental variable estimation is broadened considerably.

Key words: conditional demand models, instrumental variables, weak separability.

Introduction

The problems connected with the estimation of conditional demand models have been addressed by a number of authors, with somewhat contradictory results. The block recursivity of separable systems has sometimes been cited as an argument for using least squares (LS)[1] (see, for example, Bieri and de Janvry, p. 21). Other authors (e.g., Deaton, p. 167) have stressed the interrelations between the conditional and unconditional error terms as a reason for expecting LS to be inconsistent.[2]

In an important paper, LaFrance clarified many of the issues that had previously caused confusion in the literature. Starting from a plausible stochastic generalization of usual demand theory, he showed, among other things, (*a*) that LS estimation of the conditional demand model will in general be inconsistent, and (*b*) that standard instrumental variable (IV) techniques will also yield inconsistent results in nonlinear models. He suggested using an iterative estimation method proposed by Anderson to obtain consistency.

In this article we will show that the stochastic generalization suggested by LaFrance is not unique. While it seems unlikely that an alternative specification would invalidate LaFrance's results concerning LS, it is easy to show that another plausible specification will enable standard IV methods to yield consistent estimates in certain nonlinear models. It is thus important to formally incorporate the stochastic specification and estimation methodology into the modeling process.

Conditional and Unconditional Demand Systems

Consider a vector of commodities of interest $\boldsymbol{x}$ with corresponding prices $\boldsymbol{p}$. The goods $\boldsymbol{x}$ are assumed to be weakly separable from all other goods $\boldsymbol{z}$, whose prices are given by $\boldsymbol{q}$. The unconditional demand for $\boldsymbol{x}$ will be a function of all prices and total expenditure (y), which can be expressed in nonstochastic form as

(1) $$\boldsymbol{x} = \boldsymbol{h}(\boldsymbol{p}, \boldsymbol{q}, y).$$

Defining group expenditure as

(2) $$y_x = \boldsymbol{p}'\boldsymbol{x} = \xi(\boldsymbol{p}, \boldsymbol{q}, y),$$

The author is an associate professor in the Department of Economics, Lund University, Lund, Sweden.

This research is a part of the Nordic Food Demand Study, supported by the Swedish Agricultural and Forestry Research Council.

The comments of an anonymous reviewer are greatly appreciated. Any errors or omissions remain the responsibility of the author.

we can utilize separability, and write the conditional demand function as

(3) $$\boldsymbol{x} = \tilde{\boldsymbol{h}}(\boldsymbol{p}, y_x).$$

Separability thus implies, and is implied by, the relation $\boldsymbol{h}(\boldsymbol{p}, \boldsymbol{q}, y) = \tilde{\boldsymbol{h}}(\boldsymbol{p}, \xi(\boldsymbol{p}, \boldsymbol{q}, y))$.

Since economic theory tells us little about the stochastic nature of the econometric model, LaFrance suggests expressing a stochastic demand model in terms of the conditional expectations of $\boldsymbol{x}$ and y_x, i.e., replacing (1) through (3) by

(4) $$\bar{\boldsymbol{x}} = E(\boldsymbol{x} | \boldsymbol{p}, \boldsymbol{q}, y) = \boldsymbol{h}(\boldsymbol{p}, \boldsymbol{q}, y),$$

(5) $$\bar{y}_x = \boldsymbol{p}'\bar{\boldsymbol{x}} = \xi(\boldsymbol{p}, \boldsymbol{q}, y),$$

and

(6) $$\bar{\boldsymbol{x}} = \tilde{\boldsymbol{h}}(\boldsymbol{p}, \bar{y}_x),$$

where (6) follows from (4) and (5). If we write $\boldsymbol{x} = \bar{\boldsymbol{x}} + \epsilon_x$ and $y_x = \bar{y}_x + \upsilon_x$, then it follows that $E(\epsilon_x | \boldsymbol{p}, \boldsymbol{q}, y) = \boldsymbol{0}$, $E(\upsilon_x | \boldsymbol{p}, \boldsymbol{q}, y) = 0$, and $\boldsymbol{p}'\epsilon_x = \upsilon_x$. The conditional demand function can now be expressed with the observable level of group expenditure on the right-hand side, that is

(7) $$\boldsymbol{x} = \tilde{\boldsymbol{h}}(\boldsymbol{p}, y_x) + \tilde{\epsilon}_x,$$

where

(8) $$\tilde{\epsilon}_x = \epsilon_x + \tilde{\boldsymbol{h}}(\boldsymbol{p}, \bar{y}_x) - \tilde{\boldsymbol{h}}(\boldsymbol{p}, \bar{y}_x + \upsilon_x).$$

LaFrance uses these relations to prove, among other things, the following:

Lemma 1: Given that $E(\epsilon_x) = \boldsymbol{0}$ and $E(\upsilon_x) = 0$, then $E(\tilde{\epsilon}_x) = \boldsymbol{0}$ if and only if *either* $\boldsymbol{p}'\epsilon_x = 0$ *or* $\tilde{\boldsymbol{h}}$ is linear in y_x (LaFrance, lemmas 1 and 2).

The first condition is equivalent to $|\boldsymbol{\Sigma}_{xx}| = 0$, where $\boldsymbol{\Sigma}_{xx}$ is the covariance matrix of the unconditional errors ϵ_x. Note that this matrix is in general nonsingular, since adding-up merely implies that $\boldsymbol{p}'\epsilon_x + \boldsymbol{q}'\epsilon_z = 0$ and $\boldsymbol{p}'\tilde{\epsilon}_x = 0$. A second result is that

(9) $$\mathrm{Cov}(y_x, \tilde{\epsilon}_x) = \boldsymbol{\Sigma}_{xx}\boldsymbol{p} - E[\tilde{\boldsymbol{h}}(\boldsymbol{p}, \bar{y}_x + \boldsymbol{p}'\epsilon_x)\epsilon_x'\boldsymbol{p}],$$

from which the following can be obtained:

Lemma 2: $\mathrm{Cov}(y_x, \tilde{\epsilon}_x) = \boldsymbol{0}$ if and only if (*a*) $|\boldsymbol{\Sigma}_{xx}| = 0$; (*b*) $\boldsymbol{\Sigma}_{xx}$ is constant and $\tilde{\boldsymbol{h}}$ is linear in y_x, with its coefficient equal to a specific function of $\boldsymbol{p}$ and $\boldsymbol{\Sigma}_{xx}$; or (*c*) $\tilde{\boldsymbol{h}}$ is linear in y_x, and the unconditional errors satisfy the generalized rational random errors hypothesis (LaFrance, lemmas 1, 3, and 4).

The above equality holds approximately in cases (*b*) or (*c*) when $\tilde{\boldsymbol{h}}$ is nonlinear in y_x, as long as the first-order Taylor approximation of $\tilde{\boldsymbol{h}}$ satisfies the given restrictions.

These results are very strong. LaFrance interprets the second result as showing that LS will only yield consistent estimates of (7) under very restrictive conditions. The first result is interpreted as implying that standard IV methods (see Amemiya, or Bowden and Turkington) will only be consistent when $\tilde{\boldsymbol{h}}$ is linear in y_x. The estimation method suggested by LaFrance is an iterated generalized least squares method proposed by Anderson.

All the above results are dependent, however, on the stochastic generalization of the deterministic demand system given by (4) and (5). The definitions given there seem both natural and plausible, but an obvious question is whether there exist *other* plausible definitions that lead to different results.

Consider, for example, Deaton and Muellbauer's Almost Ideal Demand System. The conditional demand function for this model explains the within group budget shares[3] ($w_i = x_i p_i / y_x$) as a linear function of prices and the logarithm of group expenditure ($\eta_x = \ln y_x$). It is thus reasonable to assume that w_i and η_x are the variables of primary interest in the model, rather than x and y_x. In this case it is more convenient to assume a stochastic specification based on the conditional expectations $\bar{\boldsymbol{w}} = E(\boldsymbol{w} | \boldsymbol{p}, \boldsymbol{q}, y)$ and $\bar{\eta}_x = E(\eta_x | \boldsymbol{p}, \boldsymbol{q}, y)$, instead of (4) and (5).

An Alternative Stochastic Specification

Let us now consider quite general transformations of x and y_x,

$$w = f(x, p, y_x), \tag{10}$$

and

$$\eta_x = g(y_x), \tag{11}$$

where the inverse functions of g and f (for given p and y_x) both exist. Using the deterministic demand functions (1) through (3), we can write

$$w = f(h(p, q, y), p, \xi(p, q, y)) = k(p, q, y), \tag{12}$$

$$\eta_x = g(p'x) = g(\xi(p, q, y)) = \zeta(p, q, y), \tag{13}$$

and

$$w = f(\tilde{h}(p, y_x), p, y_x) = f(\tilde{h}(p, g^{-1}(\eta_x)), p, g^{-1}(\eta_x)) = \tilde{k}(p, \eta_x), \tag{14}$$

where $k(p, q, y) = \tilde{k}(p, \zeta(p, q, y))$, and[4] $p'x = p'f^{-1}(w, p, g^{-1}(\eta_x)) = g^{-1}(\eta_x)$. Note that we can also write $w = f(x, p, p'x) = \tilde{f}(x, p)$, but that $\tilde{f}$ need not be invertable for given p.

A stochastic generalization of the deterministic model, in terms of the conditional expectations of w and η_x, can now be constructed. Equations (4) through (6) will become

$$\bar{w} = E(w \mid p, q, y) = k(p, q, y), \tag{15}$$

$$\bar{\eta}_x = \zeta(p, q, y), \tag{16}$$

and

$$\bar{w} = \tilde{k}(p, \bar{\eta}_x). \tag{17}$$

If we now write $w = \bar{w} + \epsilon_w$ and $\eta_x = \bar{\eta}_x + v_w$, then ϵ_w and v_w have conditional expectations equal to zero. Defining $\bar{x} = f^{-1}(\bar{w}, p, g^{-1}(\bar{\eta}_x))$ implies that $\bar{\eta}_x = g(p'\bar{x})$, but $\bar{x}$ is no longer the conditional expectation of x if f^{-1} is nonlinear in w or η_x. Note that there no longer exists any simple relationship between v_w and ϵ_w.

Equations (15) through (17) and (4) through (6), respectively, comprise mutually exclusive alternatives when either g or f is nonlinear. They thus represent two distinct stochastic models. Both are quite plausible generalizations of the deterministic model, but they have somewhat different implications. Estimation of LaFrance's specification (if it is correct) yields unbiased predictions of x and y_x, while the specification given here yields (if correct) unbiased predictions of w and η_x. This is quite a normal situation when comparing budget-share and expenditure models, and is not usually a cause for concern.

We can continue our analysis by developing the conditional demand function as for (7), namely

$$w = \tilde{k}(p, \eta_x) + \tilde{\epsilon}_w, \tag{18}$$

where

$$\tilde{\epsilon}_w = \epsilon_w + \tilde{k}(p, \bar{\eta}_x) - \tilde{k}(p, \bar{\eta}_x + v_w). \tag{19}$$

The following result is now directly comparable to lemma 1:

> *Lemma 3:* Given that $E(\epsilon_w) = \mathbf{0}$ and $E(v_w) = 0$, then $E(\tilde{\epsilon}_w) = \mathbf{0}$ if and only if *either* $p'(x - \bar{x}) = 0$ *or* $\tilde{k}$ is linear in η_x.

Thus if $\tilde{k}$ is linear in η_x, then usual nonlinear IV estimation will yield consistent estimates.[5]

The first condition in lemma 3 has a similar interpretation to that in lemma 1, namely that the vector of unconditional demand errors, $\epsilon_x = x - \bar{x}$, has a singular distribution (although ϵ_x does not have zero conditional expectation here). Note also that the conditional budget share errors, $\tilde{\epsilon}_w$, has a singular distribution, due to adding up, whether or not $p'(x - \bar{x}) = 0$.[6]

As previously, LS will not be consistent, since

$$\text{(20)} \qquad \mathrm{Cov}(\eta_x, \tilde{\epsilon}_w) = E(v_x\epsilon_w) - E[\tilde{\boldsymbol{k}}(\boldsymbol{p}, \bar{\eta}_x + v_x)v_x]$$

can only be zero under very special conditions, analogous to those in lemma 2.

An Example and Some Implications

The Almost Ideal Demand System is usually written in determinstic form as

$$\text{(21)} \qquad w_i = \alpha_i + \Sigma_j\gamma_{ij} \ln p_j + \beta_i(\ln y_x - \ln P_I),$$

where $w_i = x_ip_i/y_x$, and $\ln P_I$ is a price index to be estimated (as a nonlinear function of $\boldsymbol{p}$ and the α and γ parameters) or approximated (e.g., as Stone's price index, $\Sigma_k w_k \ln p_k$). The group expenditure function for y_x can be given by

$$\text{(22)} \qquad y_x = \xi(\boldsymbol{p}, \boldsymbol{q}, y; \boldsymbol{\theta}),$$

which may, or may not, be in Almost Ideal form, and where $\boldsymbol{p}$ and $\boldsymbol{q}$ may, or may not, form group price indices.

Using LaFrance's specification of the stochastic Almost Ideal model, we obtain the following conditional demand system expressed in terms of observed expenditures:

$$\text{(23)} \qquad x_i = \alpha_i(y_xp_i^{-1}) + \Sigma_j\gamma_{ij}(y_xp_i^{-1})\ln p_j + \beta_i(y_xp_i^{-1})(\ln y_x - \ln P_I) + \tilde{\epsilon}_{xi}.$$

Since $E(\tilde{\epsilon}_x) \neq \mathbf{0}$ in general, the model to estimate would be

$$\text{(24)} \qquad x_i = \alpha_i(\bar{y}_xp_i^{-1}) + \Sigma_j\gamma_{ij}(\bar{y}_xp_i^{-1})\ln p_j + \beta_i(\bar{y}_xp_i^{-1})(\ln \bar{y}_x - \ln P_I) + \epsilon_{xi}$$

$$\text{(25)} \qquad y_x = \xi(\boldsymbol{p}, \boldsymbol{q}, y; \theta) + \nu_x,$$

and Anderson's method (or double regression nonlinear TSLS) will yield consistent estimates. Note that the function ξ has to be specified.

The alternative stochastic specification used in this article yields, for $\eta_x = \ln y_x$,

$$\text{(26)} \qquad w_i = \alpha_i + \Sigma_j\gamma_{ij} \ln p_j + \beta_i(\eta_x - \ln P_I) + \tilde{\epsilon}_{wi},$$

which, being linear in η_x, implies that $E(\tilde{\epsilon}_w) = \mathbf{0}$. Using y or $\ln y$ in the instrument set for η_x will yield consistent estimates. Note that we do not need to specify the group expenditure function when applying this method. If, however, we estimate

$$\text{(27)} \qquad \eta_x = \zeta(\boldsymbol{p}, \boldsymbol{q}, y; \boldsymbol{\theta}) + v_w,$$

then this can be used explicitly to form instruments (this is called the method of internal instruments by Bowden and Turkington, p. 166). LS will not be consistent, except in exceptional circumstances. The bias introduced by using LS may be small, however, and the usual Hausman–Wu test could be used to detect this.

The specification (26) seems more convenient than (24)–(25), but the choice between them must be an empirical matter. Some form of nonnested test seems to be called for, although the necessity of different estimation methods in the two models will cause complications. An alternative approach could be to test for specification error in each of the two models, using some IV generalization of Ramsey's RESET test. Note also that the two models given above are not the only possible stochastic generalizations of the Almost Ideal Demand System. The following model,

$$\text{(28)} \qquad w_i = \alpha_i + \Sigma_j\gamma_{ij} \ln p_j + \beta_i(\ln \bar{y}_x - \ln P_I) + \epsilon_{wi},$$

together with (25), will have to be estimated by Anderson's method, but will yield different estimates than those given by (24). Note that reformulating (28), so that the right-hand side is expressed in terms of y_x, will imply

$$\text{(29)} \qquad w_i = \alpha_i + \Sigma_j\gamma_{ij} \ln p_j + \beta_i(\ln y_x - \ln P_I) + \tilde{\epsilon}^*_{wi},$$

which only differs from (26) as regards the formulation of the error terms. That is,

(30) $$\tilde{\epsilon}_w^* = \epsilon_w + \beta(\ln \bar{y}_x - \ln(\bar{y}_x + v_x)),$$

while

(31) $$\tilde{\epsilon}_w = \epsilon_w + \beta v_w.$$

Conclusions

The importance of the stochastic specification of demand systems is often overlooked. The implications of separability on the specification of conditional demand models have only recently been investigated, and the article by LaFrance concerning exogeneity is of crucial importance in empirical analysis. Some of the conclusions drawn in that article are, however, dependent on the specific stochastic form that was assumed. Other specifications, which are at least equally plausible, lead to other conclusions.

In particular, conditional demand systems that are linear in *some* (linear or nonlinear) function of the group expenditures *can* be specified so that usual IV methods are consistent. Other specifications, which demand a double regression interpretation of TSLS, are also always possible. The choice of which specification to use will be an empirical matter, which generally can be difficult to separate from other misspecification problems such as choice of funcational form, etc.

The *automatic* use of LS in conditional demand systems cannot be justified. The matter of finding a consistent estimation method to use as a yardstick, seems, unfortunately, still to be unresolved.

[*Received April 1993; final revision received July 1993.*]

Notes

[1] We use the notation LS here to denote all estimation methods that are based on a theoretical zero correlation between regressors and errors, e.g., ordinary LS, nonlinear LS, seemingly unrelated regressions (SUR), iterated SUR, etc.

[2] Deaton argues that the asymptotic bias in using LS should be small, however.

[3] LaFrance mentions budget-share models in his footnote 2, but in his case, he considers the unconditional budget shares $x_i p_i / y$.

[4] The notation $\boldsymbol{f}^{-1}(\boldsymbol{w}, \boldsymbol{p}, y_x)$ is used loosely here to mean the inverse of $\boldsymbol{w} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{p}, y_x)$ for given $\boldsymbol{p}$ and y_x.

[5] Anderson's method is equivalent to iterated generalized two-stage least squares performed as a double regression method. Both Kelejian and Edgerton (1972) have shown that in nonlinear sitautions, the double regression and IV interpretations of TSLS cannot both be consistent for the same stochastic specification. In linear models the two forms are, of course, numerically identical. Edgerton (1973) has shown how a nonlinear form of Wold's fix-point method (which is equivalent to iterated double regression TSLS) can be applied if, in the present notation, $E(\tilde{\epsilon}_w) = \mathbf{0}$.

[6] In the case of simple functions, such as expenditures, budget shares, or uniform transformations of quantities demanded, the results of Lau are of interest. Note also that if we consider uniform simple functions of quantities, expenditures, or budget shares, then lemmas 1 and 3 restrict us to rank 2 conditional demand models in order that the error terms should have the requisite properties.

References

Amemiya, T. *Advanced Econometrics.* Cambridge MA: Harvard University Press, 1985.

Anderson, R. W. "Perfect Price Aggregation and Empirical Demand Analysis." *Econometrica* 47(1979):1209–30.

Bieri, J., and A. de Janvry. *Empirical Analysis of Demand Under Consumer Budgeting.* Monograph No. 30, Giannini Foundation, University of California, Berkeley, 1972.

Bowden, R. J., and D. A. Turkington. *Instrumental Variables.* Cambridge: Cambridge University Press, 1984.

Deaton, A. *Models and Projections of Demand in Post-War Britain.* London: Chapman and Hall, 1975.

Deaton, A., and J. Muellbauer. "An Almost Ideal Demand System." *Amer. Econ. Rev.* 70(1980):312–26.

Edgerton, D. L. "The Combined Use of Predictors and Taylor Series in Nonlinear Interdependent Systems." *Economie Appliquée* 26(1973):537–99.

———. "Some Properties of Two-Stage Least Squares as Applied to Nonlinear Models." *Internat. Econ. Rev.* 13(1972):26–32.

Kelejian, H. "Two-Stage Least Squares and Econometric Systems Linear in Parameters but Nonlinear in Endogenous Variables." *J. Amer. Statis. Assoc.* 66(1971):373–74.

LaFrance, J. T. "When Is Expenditure 'Exogenous' in Separable Demand Models?" *West. J. Agr. Econ.* 16(1991): 49–62.

Lau, L. J. "Complete Systems of Consumer Demand Functions through Duality." In *Frontiers of Quantitative Economics,* Vol. III, ed., M. D. Inriligator, chapter 2b. New York: Academic Press, 1976.