

Alternative Estimation Methods of Nonlinear Demand Systems

Oral Capps, Jr.

Several contemporary models of consumer demand comprise complete sets of nonlinear demand functions. Estimation methods should take into account parameter nonlinearity, cross-equation correlation, variance-covariance singularity of the disturbance terms, and various parameter restrictions. This paper presents a theoretical discussion and some empirical results using the maximum likelihood (ML) method and the iterative version of Zellner's seemingly unrelated regression (IZEF) method in the estimation of a nonlinear system of demand equations (the linear expenditure system) when the disturbance terms are both contemporaneously and serially correlated. On the basis of the evaluation of parameter estimates and their asymptotic standard errors as well as the cost of computation effort, the IZEF technique is preferred over the ML technique in this empirical problem.

A major topic in the literature of theoretical and applied econometrics is the estimation of systems of nonlinear equations. In particular, several basic contemporary models of consumer demand comprise sets of nonlinear demand functions.¹ Estimation methods should take into account essential parameter nonlinearity, cross-equation correlation, variance-covariance singularity of the disturbance terms, and various parameter restrictions. Consequently, practitioners are continually faced with theoretically satisfactory models imbued with practical difficulties.

Two viable methods for estimating nonlinear demand systems are the maximum likelihood (ML) procedure and the iterative version of Zellner's seemingly unrelated regression (IZEF) procedure.

Oral Capps, Jr. is an Assistant Professor, Department of Agricultural Economics and Statistics, Virginia Polytechnic Institute and State University, Blacksburg.

This work was sponsored by the S-165 Regional Committee on U.S. Food Demand and Consumption Behavior under Virginia Tech Contract No. 2124270.

¹ For a concise survey of this literature, see Barten (1977), Brown and Deaton, Goldberger, Powell, and Theil (1975).

This paper describes these estimation techniques in some detail under the assumption that the disturbance terms are both serially and contemporaneously correlated. In addition, this paper presents empirical results using these methods in the estimation of parameters for a particular system of demand functions—the linear expenditure system (LES)—using U.S. personal consumption data.

The organization of the paper is as follows. The next section provides a description of the ML procedure and the IZEF procedure. The third section deals with the LES specification, and the fourth section concerns the data. The fifth section contains the empirical results, and concluding comments follow in the sixth section.

Development of the ML and IZEF Procedures

To formalize the development of the ML procedure and the IZEF procedure, assume that each of m nonlinear regression equations may be written in a convenient vector form:

$$\begin{aligned} y_i &= f(\mathbf{x}_i, \theta) + \epsilon_i \\ i &= 1, \dots, m \end{aligned} \quad (1)$$

where $\theta = (\theta_1, \dots, \theta_r)'$ denotes the $r \times 1$ vector of unknown parameters in the system of equations, $\mathbf{x}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{ip_i})'$ denotes the $1 \times p_i$ vector of independent variables for the i^{th} equation, \mathbf{y}_i denotes the $n \times 1$ dependent variable vector for the i^{th} equation, and ϵ_i denotes the $n \times 1$ vector of disturbances for the i^{th} equation. n refers to the number of observations in the sample, and in addition, \mathbf{x}_{ij} , $j = 1, 2, \dots, p_i$ is of dimension $n \times 1$. The number of observations must exceed the maximum number of parameters that occur in any equation. Then $\epsilon_i = \mathbf{y}_i - f(\mathbf{x}_i, \theta)$, and ϵ_i is a function of the known \mathbf{x}_i and \mathbf{y}_i as well as the unknown θ . With \mathbf{x}_i and \mathbf{y}_i given, the set of ϵ_i becomes a function of θ alone, say $G(\theta)$.

The ML Procedure

The ML method is perhaps the best-known and most well-established method of estimation to deal with systems of nonlinear equations (Barten, 1969; Berndt, Hall, Hall, and Hausman; Goldfeld and Quandt, 1972, 1976; Green, Hassan, and Johnson; and Parks, 1971). Assume that the disturbance terms are random variables which possess a joint probability density function (pdf) $p(\epsilon)$ of known mathematical form, where ϵ denotes the complete vector of disturbance terms (nm in all). According to the principle of ML estimation, values of θ are sought to maximize the likelihood function p or, more conveniently, the logarithm of p . Thus, $G_{ML}(\theta) = \log p(\epsilon)$, and ML estimators of θ are such that $G_{ML}(\hat{\theta}_{ML}) = \sup G_{ML}(\theta)$.

Typically, most practitioners use this type of computation routine under the following assumptions on the probability distribution of the disturbance terms:

- (i) the disturbance terms have zero means:

$$E(\epsilon_{ij}) = 0$$

$$i = 1, 2, \dots, m; j = 1, 2, \dots, n$$

- (ii) The disturbance terms are not serially correlated (no autocorrelation):

$$E[\epsilon_{ij}\epsilon_{ts}] = 0 \text{ for } j \neq s \text{ and all } i, t;$$

$$i, t = 1, 2, \dots, m; j, s = 1, \dots, n$$

- (iii) the disturbance terms are contemporaneously correlated:

$$E[\epsilon_{ij}\epsilon_{it}] = \sigma_{it} \text{ for all } j = 1, 2, \dots, n$$

$$\text{and all } i, t = 1, 2, \dots, m$$

- (iv) the disturbance terms $\epsilon_j = (\epsilon_{1j}, \epsilon_{2j}, \dots, \epsilon_{mj})'$, $j = 1, 2, \dots, n$ are independent and follow the same distribution, generally a normal distribution, with positive definite symmetric variance-covariance matrix Σ_{max} . For simplicity, $E(\epsilon_j\epsilon_j') = \Sigma$ for all $j = 1, 2, \dots, n$. Thus, the complete random disturbance vector ϵ has $E(\epsilon) = 0$ and $E(\epsilon\epsilon') = \Omega = \Sigma \otimes I$, where I is the $n \times n$ identity matrix and \otimes denotes the direct- or Kronecker-product operation.

More recently, however, some researchers have taken note of evidence of autocorrelation in the estimation of systems of nonlinear equations (Berndt and Savin; Green, Hassan, and Johnson). Consequently, procedures which take into account autocorrelation are necessary. Consider the case for which the disturbances in each equation satisfy $E(\epsilon_{ij}\epsilon_{ts}) \neq 0$ for $j \neq s$ and for all i, t and for which the disturbances follow a first-order autoregressive scheme. The structure for ϵ_{ij} in this case is simply $\epsilon_{ij} = \rho_i \epsilon_{i,j-1} + u_{ij}$, $j = 2, \dots, n$, $i = 1, 2, \dots, m$, where the $\mathbf{u} = (u_{1j}, u_{2j}, \dots, u_{mj})'$ are independently, identically distributed normal random vectors with mean vector zero and contemporaneous variance-covariance matrix $\Sigma = [\sigma_{it}]$. With this formulation (Parks, 1967),

$$E[\epsilon\epsilon']$$

$$= \begin{bmatrix} \sigma_{11}P_1P_1' & \sigma_{12}P_1P_2' & \dots & \sigma_{1m}P_1P_m' \\ \sigma_{21}P_2P_1' & \sigma_{22}P_2P_2' & \dots & \sigma_{2m}P_2P_m' \\ \sigma_{m1}P_mP_1' & \sigma_{m2}P_mP_2' & \dots & \sigma_{mm}P_mP_m' \end{bmatrix}$$

$$= \Omega^* = P(\Sigma \otimes I)P' \tag{2}$$

where P is the block diagonal matrix [P_i], and

$$P_i = \begin{bmatrix} (1 - \rho_i^2)^{-1/2} & 0 & 0 & \dots & 0 \\ \rho_i(1 - \rho_i^2)^{-1/2} & 1 & 0 & \dots & 0 \\ \rho_i^2(1 - \rho_i^2)^{-1/2} & \rho_i & 1 & \dots & 0 \\ \rho_i^{i-1}(1 - \rho_i^2)^{-1/2} & \rho_i^{i-2} & \rho_i^{i-3} & \dots & 1 \end{bmatrix}_{n \times n} \quad (3)$$

where ρ_i is the coefficient of autocorrelation in the ith equation and |ρ_i| ≤ 1.

In sum, to mechanically take into account the serial correlation of the disturbance terms, the Koyck transformation is applied. This transformation results in the tenability of assumptions (i)-(iv) on the probability distribution of the disturbance terms. Through the use of the Koyck transformation, the disturbance terms are contemporaneously correlated but no longer serially correlated.

Upon this transformation, the likelihood function, under normality, assumes the form:

$$G_{M.L.}(\theta) = -\frac{mn}{2} \log(2\pi) - \frac{n}{2} \log |\Sigma| - \frac{1}{2} \mathbf{u}'(\Sigma^{-1} \otimes \mathbf{I})\mathbf{u} \quad (4)$$

since,

$$P(\mathbf{u}) = \frac{1}{(2\pi)^{mn/2} |\Sigma|^{n/2}} \cdot \exp\left[-\frac{1}{2} \mathbf{u}'(\Sigma^{-1} \otimes \mathbf{I})\mathbf{u}\right] \quad (5)$$

Note that E[**uu'**] = Ω = Σ ⊗ I, where **u** is the vector of all mn disturbances such that u_{ij} = ε_{ij} - ρ_i ε_{i,j-1}.

The last term of the expression in (4) is a nonlinear function of the unknown parameters, and consequently, the set of

likelihood equations, $\frac{\partial G_{M.L.}}{\partial \theta} = \mathbf{0}$, is generally nonlinear as well. If a supremum $\hat{\theta}_{M.L.}$ exists, this estimator satisfies the likelihood equations. Hence, as a practical matter, to maximize G_{M.L.}(θ) it is necessary

to resort to numerical techniques. The asymptotic variance-covariance matrix of the parameter estimates comes about via the inversion of the negative of the matrix of second partial derivatives (the Hessian),

$$-\left[\frac{\partial^2 G_{M.L.}}{\partial \theta \partial \theta'}\right]_{\theta}^{-1} = \hat{\theta}_{M.L.} \quad (6)$$

Asymptotically, ML estimators are optimal according to the usual criteria in econometrics (Barten, 1969). The small sample properties of this method are less satisfactory, and perhaps another method of nonlinear estimation is preferable.² In addition, with the ML technique, likelihood ratio (LR) tests are the natural method of inference. Aside from the small sample properties, the disadvantages of the ML procedure also include the need to specify the distribution for the disturbance terms in the system and the use of numerical techniques to maximize G_{M.L.}(θ).

In practice, practitioners may maximize the likelihood function outright or solve the likelihood equations. Algorithms for such purposes fall into three basic categories: (1) algorithms that employ no derivatives in the process; (2) variable metric methods that employ first partial derivatives and approximations to the second partial derivatives; and (3) methods that employ both first and second partial derivatives. No single best algorithm exists because some algorithms tend to locate inappropriate stationary points, some algorithms converge more slowly than others in terms of the requirement of the number of iterations, and some algorithms place a heavy demand on computer time and effort in programming.

The IZEF Procedure

The IZEF procedure is the iterative version of the basic approach of Zellner

² However, research with Monte Carlo experiments indicates that the ML procedure generally gives rise to parameter estimates with good finite sample properties (Goldfeld and Quandt, 1976).

in handling seemingly unrelated nonlinear regressions. Berndt and Christensen and Christensen and Manser (1972, 1976, 1977) have used the iterative version of Zellner's Aitken (ZEF) estimator in empirical applications. With the specification of the nonlinear system of equations in (1) along with the same assumptions on the probability distribution of the disturbance terms, the first step of the procedure is to obtain least squares estimators $\hat{\theta}$ through the minimization of

$$Q_i(\theta) = \frac{1}{n} \mathbf{u}_i' \mathbf{u}_i \quad i = 1, 2, \dots, m, \quad (7)$$

equation by equation. The second step is to form the residual vectors, $\hat{\mathbf{u}}_i$, and estimate the elements, $\hat{\sigma}_{it}$, of the variance-covariance matrix Σ by

$$\hat{\sigma}_{it} = \frac{1}{n} \hat{\mathbf{u}}_i' \hat{\mathbf{u}}_t \quad i, t = 1, 2, \dots, m \quad (8)$$

to obtain the estimator $\hat{\Sigma}$ of Σ . The third step of the procedure is to obtain the Aitken type estimator $\hat{\theta}$ through the minimization of

$$T(\theta) = \frac{1}{n} \mathbf{u}' (\hat{\Sigma}^{-1} \otimes \mathbf{I}) \mathbf{u}. \quad (9)$$

Note that the maximization of $G_{ML}(\theta)$ is essentially equivalent to the minimization of $T(\theta)$ (compare equations (4) and (9)). The asymptotic variance-covariance matrix of the parameter estimates emerges via the Aitken estimation phase of the procedure.

In short, the variances and covariances of the disturbances are estimated from the transformed residuals u_{ij} derived from an equation-by-equation application of least squares. The minimization of $Q_i(\theta)$ to obtain the ordinary least squares estimators ($\hat{\theta}$) and transformed residual vectors $\hat{\mathbf{u}}$, may be carried out using either Hartley's modified Gauss-Newton algorithm or Marquardt's algorithm. The parameters of the system are then estimated simultaneously by applying Aitken's generalized least squares to the whole system of equa-

tions. Obviously, we need not stop there. The resulting parameter estimates can then be used for calculating a new set of residuals leading to a new estimate of $\hat{\Sigma}^{-1}$ which can be used for obtaining new estimates of the parameters of the system and so on. Interestingly, Zellner only mentions this iterative process as a possibility to estimate systems of equations.

Conditions are set forth such that the IZEF estimator $\hat{\theta}$ is weakly consistent for θ , such that, in addition, $\hat{\theta}$ is asymptotically normally distributed (Gallant). Hence, the IZEF estimator $\hat{\theta}$ has similar asymptotic properties as the ML estimator $\hat{\theta}_{ML}$. However, hardly any information is available about the small sample properties of these two estimators. Despite the similarity of asymptotic properties of $\hat{\theta}$ and $\hat{\theta}_{ML}$, practitioners generally abandon the IZEF procedure in favor of the ML procedure.

In sum, the ML procedure and the IZEF procedure employ modern nonlinear algorithms to estimate systems of nonlinear equations. The algorithms typically but not always guarantee convergence of the iterative estimation processes. The computational burden for both techniques may be immense, and in general, problems may arise with respect to the numerical precision of the results. These difficulties are typically a function of the size and complexity of the model. To shed some light on such circumstances, a presentation of the two techniques with respect to the estimation of parameters from a popular nonlinear system of demand equations is in order.

The LES Specification

The linear expenditure system (LES) is a venerable system of demand (or expenditure) equations. For broad aggregate commodities, the LES provides a reasonable model for representation of consumer response to changes in prices and total expenditure. Demand equations (expenditure equations), with prices and total ex-

penditure as explanatory variables, are the traditional tools for analyzing consumer behavior. The selection of this specification for empirical purposes in this paper rests on the fact that the LES is the most widely employed complete demand system (Goldberger; Pollak and Wales; and Brown and Deaton). Additionally, the LES facilitates the estimation of large systems of equations (Braithwait, 1977, 1980).

For statistical purposes, the LES can be written in the following form:³

$$p_{ij}q_{ij} = p_{ij}\gamma_i + \beta_i(y_j - \sum_{i=1}^m p_{ij}\gamma_i) + \epsilon_{ij},$$

$$i = 1, 2, \dots, m; \quad j = 1, 2, \dots, n \quad (10)$$

where q_{ij} corresponds to the j^{th} observation on the quantity of the i^{th} commodity, p_{ij} corresponds to the j^{th} observation on the price of the i^{th} commodity, y_j corresponds to the j^{th} observation on the total expenditure on all commodities, and ϵ_{ij} corresponds to the j^{th} observation on the unobserved random disturbance for the i^{th} commodity. The γ_i and β_i are unknown parameters to be estimated, subject to the

linear constraint that $\sum_{i=1}^m \beta_i = 1$. Hence, the

LES requires the estimation of $2m - 1$ parameters. In addition, other restrictions on the parameters of this system are the following: (1) $\beta_i > 0$ for all i ; (2) $\gamma_i > 0$ for all i ; and (3) $q_{ij} - \gamma_i > 0$ for all i and for all $j = 1, 2, \dots, n$.

In light of such restrictions, a useful interpretation of the parameters of the LES is available. Given y_{ij} and p_{1j}, \dots, p_{mj} , the consumer first purchases minimum or habitual quantities of each good, γ_1 of commodity 1, γ_2 of commodity 2, \dots , and γ_m of commodity m .⁴ At the given prices, this expenditure for observation j is simply

$\sum_{r=1}^m p_{rj}\gamma_r$, the subsistence expenditure for the consumer. The supernumerary expenditure for the consumer for observation j is then $y_j - \sum_{r=1}^m p_{rj}\gamma_r$. The consumer finally distributes the supernumerary expenditure among the m commodities in the proportions β_1, \dots, β_m .

The specification in (10) is based upon the assumption of no autocorrelation of the disturbance terms. Upon the application of the Koyck transformation to take care of the autocorrelation of the disturbance terms, the LES may be written as follows:

$$p_{ij}q_{ij} = \rho_i p_{i,j-1} q_{i,j-1} + \gamma_i (p_{ij} - \rho_i p_{i,j-1})$$

$$+ \beta_i (y_j - \sum_{i=1}^m p_{ij}\gamma_i)$$

$$- \rho_i \beta_i (y_{j-1} - \sum_{i=1}^m p_{i,j-1}\gamma_i) + u_{ij}. \quad (11)$$

From (10), the constraint on the β_i 's and the fact that the total expenditure $y_j = \sum_{i=1}^m p_{ij}q_{ij}$ imply that $\sum_{i=1}^m \epsilon_{ij} = 0$ for all j . This

latter restriction implies that the coefficient of autocorrelation in any equation must be the same ($\rho_i = \rho$ for all i), and the linear combination $\sum_{i=1}^m u_{ij}$ for all j must sum

to zero (Berndt and Savin; Green, Hassan, and Johnson). Interestingly, the linear dependency of the disturbance terms in (10) imposes a restriction on the parameters of the autoregressive processes.

The linear restriction on the disturbance terms means that the variance-covariance matrix of disturbance terms for the full systems is singular. Since the disturbance terms are linearly dependent, the ML procedure and the IZEF procedure may not be employed on all m equations in the system at once. Thus, one of the equations in (11) is completely redundant in the sense that using the information

³ When parameters of demand systems are estimated from sample data, additive disturbances are typically introduced.

⁴ Strictly speaking, each γ_i need not exceed zero. However, this economic interpretation of the γ_i 's is only meaningful for positive parameter values.

contained in any $m - 1$ of the equations, the remaining equation can be obtained by an appropriate linear combination. The singularity of $\Omega = \Sigma \otimes I$ can be handled by discarding one of the equations.

Barten (1969) has shown that the maximum likelihood estimates of the parameters for any system are invariant with respect to the equation deleted provided no autocorrelation exists. Ruble has demonstrated that unlike Zellner's Aitken ZEF parameter estimates, the IZEF estimates do not depend upon the equation omitted. Further, Kmenta and Gilbert have shown that the IZEF procedure produces maximum likelihood estimates for linear equation systems. Christensen and Manser (1976) have claimed, by way of empirical application to estimate indirect and direct translog models, that this result holds for both linear and nonlinear models. Although this claim has been substantiated by Barnett, the theoretical demonstration holds only under certain regularity conditions (Barnett, pp. 355-56).

In addition, the speed of convergence of the ML procedure and the IZEF procedure and the accuracy of the parameter estimates of γ_i , β_i , and ρ depend on starting values of the parameters and the tolerance level for algorithm termination. Further, only in the rarest cases can one ascertain with certainty that a local minimum (maximum) is a global minimum (maximum). A number of shortcomings may occur to prevent the attainment of global minima (maxima): (1) extreme flatness of $G_{ML}(\theta)$ and $T(\theta)$, (2) inaccuracies in the numerical evaluation of partial derivatives, and (3) the incorporation of certain constraints to affect the path of convergence. Hence, an attempt to check on the attainment of global minima (maxima) is in order. Such an attempt typically involves the arduous and inelegant procedure of employing different starting values for the parameters subject to the same tolerance level for termination of the iterative process.

The Data

The data used for the estimation of the LES are the U.S. personal consumption expenditure data for the period 1949-77. Published by the Commerce Department, these data are available for twelve major commodity groups. For empirical purposes, five aggregate commodity groups emerge from the basic twelve commodity groups: (1) food, (2) household and personal items, (3) energy, (4) housing, and (5) miscellaneous. The food category includes food at home and food away from home. The household and personal items category entails clothing, durable goods, nondurable goods, and services. The energy category deals with transportation and utilities. The housing category includes owner and tenant-occupied, non-farm dwellings. The miscellaneous category is a residual category and consists of all remaining items.

The number of observations in the sample is 29, typical of data series available to researchers for the estimation of demand systems. All expenditures used are on a real per capita basis, and total expenditure is the sum of the real per capita expenditures on these five aggregate groups. The real per capita quantities of the five commodities are obtained by dividing their real per capita expenditures by their implicit price deflators (1972 = 100). The implicit price deflators are obtained by dividing nominal per capita expenditures by real per capita expenditures.

Empirical Results

This section presents empirical results of the ML procedure and the IZEF procedure with respect to the estimation of the parameters from the LES using the aforementioned U.S. consumption data. This presentation focuses on the following: (1) evaluation of the parameter estimates and their asymptotic standard errors

TABLE 1. Starting Values for the Parameters.

	Starting Value SET 1	Starting Value SET 2	Starting Value SET 3
β_{FOOD}	.2423	.2908	.1938
β_{HSPER}	.0419	.0503	.0335
β_{ENERGY}	.3673	.4408	.2938
β_{HOUSING}	.5267	.6320	.4214
β_{MISCEL}	.7169	.8603	.5735
γ_{FOOD}	352.237	422.684	281.790
γ_{HSPER}	900.28	1,080.336	720.224
γ_{ENERGY}	200.778	240.934	160.622
γ_{HOUSING}	169.919	203.903	135.935
γ_{MISCEL}	379.843	455.812	303.874
ρ	.7816	.9379	.6252

Source: The author.

and (2) cost in terms of computational effort necessary to employ the techniques. In the application, the ML procedure uses the Quasi-Newton iterative method available from the Shazam package (White), while the IZEF procedure uses the modified Gauss-Newton iterative method available from the SAS/ETS package (SAS Institute, Inc.). For econometric estimation, in general, researchers typically employ either the Shazam routines or the SAS/ETS routines. Potential differences in parameter estimates, standard errors, and computational costs may subsequently be due to differences in the inherent ML and IZEF procedures, to differences in iterative methods from the estimation packages, or to some combination thereof.

Five models are employed to investigate the differences among parameter estimates and standard errors with different equations deleted. To overcome the singularity of the variance-covariance matrix of disturbance terms, each model consists of four equations with the fifth equation deleted. Model 1 corresponds to the deletion of the miscellaneous category, Model 2 corresponds to the deletion of the food category, Model 3 corresponds to the deletion of the household and personal items category, Model 4 corresponds to the deletion of the energy category, and Model 5

corresponds to the deletion of the housing category.

Three sets of starting values for the parameters are employed to check on the attainment of global minima (maxima). The sets of starting values for the parameters as well as the lower and upper bounds for the parameters are shown in Table 1. The first set of starting values reflects more or less arbitrary choices for the initial approximation. The second set of starting values represents a 20 percent increase over each starting value in the first set, and the third set of starting values represents a 20 percent decrease over each starting value in the first set.⁵ The choices of the starting values for ρ are based on the residuals of the system of equations estimated under the assumption of no autocorrelation of the disturbance terms. The estimation of the parameters in the system is carried out by dropping the first observation of the data set. Different results may be obtained if the first observation is included and the restriction $|\rho| \leq 1$ is imposed (Beach and Mackinnon). A cursory examination of the residuals of the LES from the use of the ML procedure and the IZEF procedure suggests that the assumption of no autocorrelation is not tenable.⁶ The tolerance level for termination of the ML procedure and the IZEF procedure is .0001.

The parameter estimates and their asymptotic standard errors obtained from

⁵ The three sets of starting values are consequently scalar multiples of each other. To more fully explore the issue of sensitivity to starting values, additional distinctive sets of starting values were employed. The empirical results, regardless of the choice of starting values, were essentially the same in all cases.

⁶ Durbin and Malinvaud have suggested that the conventional single-equation Durbin-Watson Statistic be used to check for serial correlation of disturbances in the multivariate equations setting. The appropriate number of degrees of freedom is (K,T) for the IZEF and ML estimates, where K is the number of regressors for each equation and T is the number of annual observations.

TABLE 2. Estimation of the LES with the ML Procedure: Parameter Estimates and Asymptotic Standard Errors.

Model	β_{FOOD}			β_{HSRPER}			β_{ENERGY}			β_{HOUSING}			β_{WISCEL}		
	SVS1	SVS2	SVS3	SVS1	SVS2	SVS3	SVS1	SVS2	SVS3	SVS1	SVS2	SVS3	SVS1	SVS2	SVS3
1	.14847 ^a (.01848) ^b	.14846 (.01917)	.14846 (.01908)	.51444 (.03108)	.51444 (.02934)	.51444 (.03137)	.09636 (.00617)	.09636 (.00587)	.09636 (.00620)	.09117 (.00734)	.09117 (.00630)	.09117 (.00743)	.14956 (N/E) ^c	.14957 (N/E)	.14957 (N/E)
2	.14847 (N/E)	.14847 (N/E)	.14847 (N/E)	.51444 (.03224)	.51444 (.03019)	.51444 (.02984)	.09636 (.00617)	.09636 (.00627)	.09636 (.00554)	.09117 (.00715)	.09117 (.00732)	.09117 (.00723)	.14956 (.01157)	.14956 (.01144)	.14956 (.01106)
3	.14846 (.01978)	.14846 (.01938)	.14847 (.01939)	.51445 (N/E)	.51445 (N/E)	.51601 (N/E)	.09636 (.00634)	.09636 (.00590)	.09495 (.00513)	.09117 (.00724)	.09117 (.00693)	.09167 (.00697)	.14956 (.01214)	.14956 (.01054)	.14890 (.01191)
4	.14847 (.01860)	.14847 (.01932)	.14816 (.01718)	.51444 (.03127)	.51444 (.03181)	.51632 (.02961)	.09636 (N/E)	.09636 (N/E)	.09495 (N/E)	.09117 (.00700)	.09117 (.00706)	.09168 (.00688)	.14956 (.01214)	.14956 (.01113)	.14889 (.01099)
5	.14847 (.01966)	.14846 (.02028)	.14847 (.02002)	.51444 (.03137)	.51444 (.03244)	.51443 (.03240)	.09636 (.00618)	.09636 (.00602)	.09636 (.00592)	.09117 (N/E)	.09118 (N/E)	.09118 (N/E)	.14956 (.01157)	.14956 (.01133)	.14956 (.01130)
1	351.44 (34.87)	351.44 (34.21)	351.44 (34.49)	839.85 (165.74)	839.85 (165.53)	839.85 (165.38)	257.13 (19.78)	257.13 (19.92)	257.13 (19.84)	395.67 (29.52)	395.67 (28.98)	395.67 (30.35)	400.75 (34.96)	400.75 (34.70)	400.75 (36.33)
2	351.44 (36.20)	351.44 (34.78)	351.44 (32.30)	839.85 (161.65)	839.86 (165.05)	839.85 (163.71)	257.13 (19.58)	257.13 (19.80)	257.13 (19.28)	395.67 (29.17)	395.67 (29.09)	395.67 (29.36)	400.75 (34.96)	400.75 (34.86)	400.75 (34.39)
3	351.44 (35.58)	351.44 (33.82)	340.49 (24.32)	839.86 (166.90)	839.86 (170.26)	763.97 (5.63)	257.13 (19.67)	257.13 (19.88)	248.93 (8.07)	395.67 (29.24)	395.67 (30.06)	383.81 (14.86)	400.75 (34.88)	400.75 (34.78)	387.04 (18.78)
4	351.44 (34.79)	351.44 (34.18)	340.35 (23.45)	839.84 (166.84)	839.84 (161.62)	763.21 (5.23)	257.13 (20.08)	257.13 (19.18)	248.85 (7.83)	395.67 (29.67)	395.67 (28.49)	383.69 (14.44)	400.75 (35.59)	400.75 (34.15)	386.89 (17.23)
5	351.44 (35.82)	351.44 (32.46)	351.44 (33.70)	839.85 (166.45)	839.85 (147.20)	839.84 (156.91)	257.13 (20.03)	257.13 (17.58)	257.13 (19.31)	395.67 (29.50)	395.67 (25.90)	395.67 (29.16)	400.75 (35.42)	400.75 (31.57)	400.75 (34.17)

^a Parameter estimate.
^b Asymptotic standard error.
^c Not estimable.

TABLE 3. Estimation of the LES With the IZEF Procedures: Parameter Estimates and Asymptotic Standard Errors.

Model	β_{FOOD}			β_{ENERGY}			$\beta_{HOUSING}$			β_{MISCEL}					
	SVS1	SVS2	SVS3	SVS1	SVS2	SVS3	SVS1	SVS2	SVS3	SVS1	SVS2	SVS3			
1	.1484 ^a (.01822) ^b	.14845 (.01822)	.14844 (.01822)	.51445 (.03128)	.51445 (.03128)	.51446 (.03128)	.09636 (.00629)	.09636 (.00629)	.09636 (.00629)	.09117 (.00747)	.09117 (.00747)	.09117 (.00747)	.14958 (N/E) ^c	.14957 (N/E)	.14957 (N/E)
2	.14850 (N/E)	.14850 (N/E)	.14851 (N/E)	.51441 (.03128)	.51441 (.03128)	.51440 (.03127)	.09636 (.00629)	.09636 (.00629)	.09636 (.00629)	.09117 (.00746)	.09117 (.00746)	.09117 (.00746)	.14956 (.01187)	.14956 (.01187)	.14956 (.01187)
3	.14848 (.01822)	.14847 (.01822)	.14848 (.01822)	.51443 (N/E)	.51444 (N/E)	.51442 (N/E)	.09636 (.00629)	.09636 (.00629)	.09637 (.00629)	.09117 (.00746)	.09117 (.00746)	.09117 (.00746)	.14956 (.01187)	.14956 (.01187)	.14956 (.01187)
4	.14848 (.01822)	.14848 (.01822)	.14849 (.01822)	.51441 (.03128)	.51440 (.03127)	.51439 (.03127)	.09638 (N/E)	.09639 (N/E)	.09639 (N/E)	.09117 (.00746)	.09117 (.00746)	.09117 (.00746)	.14956 (.01187)	.14956 (.01187)	.14956 (.01187)
5	.14844 (.01822)	.14844 (.01822)	.14844 (.08122)	.51446 (.03128)	.51446 (.03128)	.51446 (.03128)	.09636 (.00629)	.09636 (.00629)	.09636 (.00629)	.09119 (N/E)	.09119 (N/E)	.09119 (N/E)	.14955 (.01187)	.14955 (.01187)	.14955 (.01187)
1	351.41 (32.08)	351.42 (32.08)	351.41 (32.08)	839.86 (159.13)	839.86 (159.13)	839.86 (159.14)	257.12 (18.52)	257.12 (18.52)	257.12 (18.52)	395.67 (28.24)	395.67 (28.24)	395.67 (28.24)	400.74 (33.02)	400.74 (33.02)	400.74 (33.02)
2	351.45 (32.08)	351.46 (32.08)	351.46 (32.08)	839.85 (159.05)	839.85 (159.05)	839.85 (159.04)	257.13 (18.52)	257.13 (18.52)	257.13 (18.52)	395.66 (28.23)	395.66 (28.23)	395.66 (28.23)	400.76 (33.02)	400.76 (33.02)	400.76 (33.02)
3	351.46 (32.08)	351.45 (32.08)	351.47 (32.08)	839.85 (159.04)	839.85 (159.05)	839.85 (159.03)	257.13 (18.52)	257.13 (18.52)	257.13 (18.52)	395.66 (28.23)	395.66 (28.23)	395.66 (28.23)	400.76 (33.02)	400.76 (33.02)	400.76 (33.02)
4	351.46 (32.08)	351.46 (32.08)	351.47 (32.08)	839.84 (159.04)	839.84 (159.03)	839.84 (159.02)	257.13 (18.52)	257.13 (18.52)	257.13 (18.52)	395.66 (28.23)	395.66 (28.23)	395.66 (28.23)	400.76 (33.02)	400.76 (33.02)	400.76 (33.02)
5	351.41 (32.08)	351.40 (32.08)	351.41 (32.08)	839.86 (159.14)	839.87 (159.15)	839.86 (159.15)	257.12 (18.52)	257.12 (18.52)	257.12 (18.52)	395.67 (28.24)	395.67 (28.24)	395.67 (28.24)	400.74 (33.02)	400.74 (33.02)	400.74 (33.02)

^a Parameter estimate.
^b Asymptotic standard error.
^c Not estimable.

TABLE 4. The Estimates and the Asymptotic Standard Errors for the Autoregressive Parameter in the LES.

	ρ ML Procedure			ρ IZEF Procedure		
	SVS1	SVS2	SVS3	SVS1	SVS2	SVS3
1	.95073 (.01548)	.95073 (.01582)	.95073 (.01576)	.95074 (.01257)	.95074 (.01257)	.95075 (.01257)
2	.95073 (.01516)	.95073 (.01333)	.95073 (.01500)	.95071 (.01258)	.95071 (.01258)	.95071 (.01258)
3	.95073 (.01576)	.95073 (.01620)	.94864 (.01495)	.95071 (.01258)	.95071 (.01258)	.95070 (.01258)
4	.95073 (.01497)	.95073 (.01530)	.94862 (.01366)	.95071 (.01258)	.95070 (.01258)	.95070 (.01258)
5	.95073 (.01592)	.95073 (.01552)	.95073 (.01521)	.95075 (.01257)	.95075 (.01257)	.95075 (.01257)

Source: Computations by the author.

the ML procedure and the IZEF procedure are exhibited in Tables 2-4. In general, for the ML method and the IZEF method, the parameter estimates of the LES and their estimated asymptotic standard errors are almost identical for the five models for any given set of starting values. For the two procedures, the parameter estimates are indeed invariant with respect to the equation deleted. This result should not be surprising due to the transformation process used to overcome autocorrelation problems. That is, the situation is consequently the same as that discussed by Barten (1969).

However, the estimated asymptotic standard errors of the ML estimates depend upon the equation omitted, while the estimated asymptotic standard errors of the IZEF estimates are invariant with respect to the equation omitted. The differences among the estimated standard errors of the ML estimates with different equations deleted, though, are for the most part negligible. Also, in most cases, the standard errors of the IZEF estimates for the γ 's and ρ are lower than the corresponding standard errors of the ML estimates. On the other hand, except for β_1 , the standard errors of the IZEF estimates for the β 's are generally higher than the corresponding standard errors of the ML estimates.

In addition, for any given model, the ML procedure and the IZEF procedure generally generate parameter estimates and asymptotic standard errors which are invariant with respect to starting values. However, in two instances, for the ML procedure the parameter estimates and standard errors are noticeably different (particularly for the γ 's) with respect to starting values. The differences in the IZEF parameter estimates and standard errors with respect to starting values are almost nonexistent. Consequently, when using the ML procedure, researchers should perhaps be particularly cautious in selecting starting values.

In every case, all the parameter estimates are statistically significant, with t-values larger than 2, and in agreement with theoretical expectations, the parameter estimates are positive. The statistical significance of ρ is in agreement with findings by Luch and Williams and Green, Hassan, and Johnson.⁷ The magnitude of

⁷ The autocorrelation hypothesis is tested as follows. The LES without the autocorrelation correction, $\rho = 0$, is a restricted version of the LES properly specified to include the autocorrelated error structure. Under the null hypothesis of no autocorrelation of the disturbance terms, it can be shown that $n(\ln|\hat{\Sigma}_R| - \ln|\hat{\Sigma}_{UR}|)$ is distributed asymptotically as a χ^2 statistic with one degree of freedom. n is the

TABLE 5. Own-Price and Expenditure Elasticities for the Commodities Obtained Using the ML and IZEF Procedures.^a

Commodity	Own-Price Elasticity ^b	Expenditure Elasticity ^b
Food	-0.5716	.7103
Household & Personal Items	-0.6900	1.2996
Energy	-0.3828	.8423
Housing	-0.2036	.6818
Miscellaneous Items	-0.4604	.8497

^a These measures are based on the following parameter estimates: $\beta_1 = .14847$; $\beta_2 = .51444$; $\beta_3 = .09636$; $\beta_4 = .09117$; $\beta_5 = .14956$; $\gamma_1 = 351.44$; $\gamma_2 = 839.85$; $\gamma_3 = 257.13$; $\gamma_4 = 395.67$; $\gamma_5 = 400.75$.

^b Evaluated at the sample means.

Source: Computations by the author.

the ML and IZEF parameter estimates are reasonable in light of the fact that the γ 's reflect minimum or habitual quantities of each commodity and that the β 's represent marginal budget shares of each commodity. To illustrate, the consumer allocates at the margin roughly 15 percent of total expenditure to food (at home and away from home); 51 percent to clothing, durable goods, nondurable goods, and services; 10 percent to transportation and utilities; 9 percent to housing; and 15 percent to miscellaneous items. The minimum annual per capita expenditure, in 1972 dollars, is approximately \$350 for food, \$840 for household and personal items, \$260 for energy, \$395 for housing, and \$400 for miscellaneous items.

Additional magnitudes of interest in the context of complete demand systems are typically own-price elasticities and total expenditure elasticities. Such concepts play a large role in the comprehension of consumer behavior and the formulation of economic policy. With reference to the LES, the expenditure elasticity for the i^{th} commodity is $\beta_i y_i / p_i q_{ij}$, and similarly, the own-price elasticity for the i^{th} commodity is $(\gamma_i (1 - \beta_i) - q_{ij}) / q_{ij}$. Thus, the ML and

IZEF estimates of the parameters of the LES determine estimates of economically meaningful magnitudes.

The estimated own-price and expenditure elasticities for the commodities, evaluated at the sample means, are exhibited in Table 5. The own-price and expenditure elasticities for the various commodities are very plausible. The demands for the five goods are inelastic, with household and personal items relatively least inelastic and housing relatively most inelastic. Except for household and personal items, the commodities are expenditure inelastic. Food, energy, housing, and miscellaneous items are necessities while household and personal items are luxuries.

The goodness-of-fit criterion (R^2) offers information complementary to the theoretical and statistical support of parameter estimates. In all cases, the LES specification with the autocorrelation correction accounts for approximately 98 percent of the variation in real per capita expenditure on food and more than 99 percent of the variation in real per capita expenditure on household personal items, energy, housing, and miscellaneous items.⁸

The estimation of the LES, at least with reference to this empirical problem, was rather easy and inexpensive. In all cases, the algorithms required less than 17 sec-

number of sample observations, Σ_R is the restricted estimator of the variance-covariance matrix of disturbance terms, and $\hat{\Sigma}_{UR}$ is the unrestricted estimator of the variance-covariance matrix of disturbance terms (Theil, 1971). The autoregressive parameter is significantly different from zero at any reasonable level of significance in all cases.

⁸ The R^2 statistic is computed as one minus the ratio of the residual sum of squares to the total sum of squares.

onds of computation (CPU) time. All computations occurred on an IBM 370 Model 158 Dual Processor. However, in terms of CPU time, the IZEF procedure was two and a half to five times as fast, depending upon the equation deleted and the set of starting values. Consequently, the ML procedure was considerably more time-consuming than the IZEF procedure.

Concluding Comments

This paper provides a theoretical and empirical discussion of the ML method and the IZEF method in the estimation of a nonlinear system of demand equations (the LES) when the disturbance terms are both contemporaneously and serially correlated. The IZEF estimator $\hat{\theta}$ has similar asymptotic properties as the ML estimator $\hat{\theta}_{ML}$. These estimators converge in probability to the true θ , and these estimators follow asymptotic normal distributions.

In agreement with Christensen and Manser (1976), sample evidence exists to indicate that the IZEF procedure generates parameter estimates and estimated asymptotic standard errors which are essentially equivalent to those from the ML procedure. In this study, both procedures produced invariant and reasonable parameter estimates with respect to various equations deleted. Also, the ML procedure (except for two instances) and the IZEF procedure produced parameter estimates which were invariant with respect to starting values. However, the ML procedure generates standard errors which depend upon the equation omitted and the starting values, while the IZEF procedure generates standard errors invariant with respect to the equation omitted and the starting values. Additionally, the ML procedure is considerably more time-consuming than the IZEF procedure. In sum, despite their striking similarity, the ML and IZEF procedures reveal subtle, yet rather important, differences. Researchers

consequently should exercise caution in the choice of techniques to be employed in the estimation of nonlinear systems of demand equations. On balance, from the sample evidence in this empirical problem, the IZEF technique is preferred over the ML technique.

There remains some need for further research. An empirical presentation of the ML procedure and the IZEF procedure in the estimation of other nonlinear systems of demand equations besides the venerable LES is in order. Additionally, for generalization purposes, comparisons of the ML method and the IZEF method via Monte Carlo experiments using the same iteration techniques for nonlinear demand systems may be worthwhile. From an econometric viewpoint, since the estimation of nonlinear systems of equations is essentially in the early stages of development, additional research is very likely to pay huge dividends.

References

- Barnett, W. A. "Maximum Likelihood and Iterated Aitken Estimation of Nonlinear Systems of Equations," *Journal of the American Statistical Association*, 71(1976): 354-60.
- Barten, A. P. "The System of Consumer Demand Functions Approach: A Review," *Econometrica*, 45(1977): 23-51.
- Barten, A. P. "Maximum Likelihood Estimation of a Complete System of Demand Equations," *European Economic Review*, 1(1969): 7-13.
- Beach, C. M. and J. G. Mackinnon. "Maximum Likelihood Estimation of Singular Equation Systems with Autoregressive Disturbances," Discussion Paper No. 276, Queen's University, 1977.
- Berndt, E. R. and L. R. Christensen. "The Translog Function and the Substitution of Equipment, Structures, and Labor in the U.S. Manufacturing 1929-1968," *Journal of Econometrics*, (1973): 81-114.

- Berndt, E. K., B. H. Hall, R. E. Hall, and J. A. Hausman. "Estimation and Inference in Nonlinear Structural Models," *Annals of Economic and Social Measurements*, 3/4(1974): 653-65.
- Berndt, E. K. and N. E. Savin. "Estimation and Hypothesis Testing in Singular Equation Systems with Autoregressive Disturbances," *Econometrica*, 43(1975): 937-57.
- Braithwait, S. D. "An Empirical Comparison of Alternative Multi-Level Demand Systems for the U.S.," Working paper No. 74, Bureau of Labor Statistics, 1977.
- Braithwait, S. D. "The Substitution Bias of the Laspeyres Price Index: An Analysis Using Estimated Cost-of-Living Indexes," *American Economic Review*, 70(1980): 64-77.
- Brown, J. A. C. and A. Deaton. "Surveys in Applied Economics: Models of Consumer Behavior," *The Economic Journal*, 82(1972): 1145-1236.
- Christensen, L. R. and M. E. Manser. "The Translog Utility Function and the Substitution of Meats in U.S. Consumption, 1946-1968," BLS Discussion Paper No. 6, 1972.
- Christensen, L. R. and M. E. Manser. "Cost-of-Living Indices and Price Indexes for U.S. Meat and Produce, 1947-1971." In *Household Production and Consumption* (N. E. Terleckyj, ed.), 1976.
- Christensen, L. R. and M. E. Manser. "Estimating U.S. Consumer Preferences for Meat with a Flexible Utility Function," *Journal of Econometrics*, 5(1977): 37-53.
- Durbin, J. "Testing for Serial Correlation in Systems of Simultaneous Regression Equations," *Biometrika*, December 1957, 370-77.
- Gallant, A. R. "Seemingly Unrelated Nonlinear Regressions," *Journal of Econometrics*, 3(1975): 35-50.
- Goldberger, A. S. "Functional Forms of Utility: A Review of Consumer Demand Theory," Systems Foundation, Methodology and Policy Workshop Paper 6703, University of Wisconsin, 1967.
- Goldfeld, S. M. and R. E. Quandt. *Nonlinear Methods in Econometrics*, Amsterdam: North-Holland, 1972.
- Goldfeld, S. M. and R. E. Quandt. *Studies in Nonlinear Estimation*, Ballinger Publishing Company, 1976.
- Green, R. D., Z. A. Hassan, and S. R. Johnson. "Maximum Likelihood Estimation of Linear Expenditure Systems with Serially Correlated Errors," *European Economic Review*, 11(1978): 207-19.
- Hartley, H. "The Modified Gauss-Newton Method of the Fitting of Nonlinear Regression Functions by Least Squares," *Technometrics*, 3(1961): 269-80.
- Kmenta, J. and R. F. Gilbert. "Small Sample Properties of Alternative Estimators of Seemingly Unrelated Regressions," *Journal of the American Statistical Association*, 63(1968): 1180-1200.
- Lluch, C. and R. Williams. "Consumer Demand Systems and Aggregate Consumption in the USA: An Application of the Extended Linear Expenditure System," *Canadian Journal of Economics*, 8(1975): 49-66.
- Malinvaud, E. *Statistical Methods of Econometrics*, Chicago Rand McNally and Company, 1970.
- Marquardt, D. W. "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *Journal of the Society of Industrial Applied Mathematics*, 11(1963): 431-41.
- Parks, R. W. "Efficient Estimation of a System of Regression Equations when Disturbances are Both Serially and Contemporaneously Correlated," *Journal of the American Statistical Association*, 62(1967): 500-09.
- Parks, R. W. "Maximum Likelihood Estimation of the Linear Expenditure System," *Journal of the American Statistical Association*, 66(1971): 900-03.
- Pollak, R. A. and T. J. Wales. "Estimation of the Linear Expenditure System," *Econometrica*, 37(1969): 611-28.
- Powell, A. A. *Empirical Analytics of Demand Systems*, Lexington: Heath, 1974.
- Ruble, W. L. "Improving the Computation of Simultaneous Stochastic Linear Equation Estimates," unpublished Ph.D. dissertation, Michigan State University, 1968.
- SAS/ETS User's Guide, 1980 Edition, SAS Institute, Inc., 1980.
- Theil, H. *Principles of Econometrics*, New York: Wiley, 1971.

Theil, H. *Theory and Measurement of Consumer Demand, Volume I*, Amsterdam: North-Holland, 1975.

White, K. J., "A General Computer Program for Econometric Methods—SHAZAM," *Econometrica*, January 1978.

Zellner, A. "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregate Bias," *Journal of the American Statistical Association*, 57(1962): 348-67.