JOURNAL OF APPLIED ECONOMETRICS
J. Appl. Econ. **14**: 209–232 (1999)

ESTIMATION IN LARGE AND DISAGGREGATED DEMAND SYSTEMS: AN ESTIMATOR FOR CONDITIONALLY LINEAR SYSTEMS

RICHARD BLUNDELL[a]* AND JEAN MARC ROBIN[b]
[a]*Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE and University College London, UK*
[b]*INRA-LEA, 48 bd. Jourdan, 75014 Paris, CREST-INSEE and EPEE-Universite d'Evry, France*

SUMMARY

Empirical demand systems that do not impose unreasonable restrictions on preferences are typically non-linear. We show, however, that all popular systems possess the property of conditional linearity. A computationally attractive iterated linear least squares estimator (ILLE) is proposed for large non-linear simultaneous equation systems which are conditionally linear in unknown parameters. The estimator is shown to be consistent and its asymptotic efficiency properties are derived. An application is given for a 22-commodity quadratic demand system using household-level data from a time series of repeated cross-sections. Copyright © 1999 John Wiley & Sons, Ltd.

1. INTRODUCTION

The empirical analysis of consumer demand has been given a new impetus by the renewed interest in indirect tax reform and the welfare consequences of such reforms. This is as true of the move toward sales tax reform in the USA as it is toward the reform of VAT in Europe, placing new challenges on the analysis of demand. Estimation is required at the household level and across a large number of disaggregated commodities.

Demand behaviour at the individual household level is non-linear. It is not reasonable to assume linearity of expenditures in terms of total budget and relative prices, even the log linear expenditure share models that form the underlying shape of the popular Translog and Almost Ideal models of Jorgenson, Christensen and Lau (1975a,b) and Deaton and Muellbauer (1980a,b) respectively have been shown to require further non-linear terms. These terms reflect the growing evidence from a series of recent empirical studies that suggest quadratic logarithmic income terms are required for certain expenditure share equations (see, for example, Atkinson *et al.*, 1989; Bierens and Pott-Buter, 1987; Hausman, Newey and Powell, 1995; Lewbel, 1991; Blundell, Pashardes and Weber, 1993).

In addition to this move away from linearity, there is a growing interest in a greater disaggregation of commodities. This partly reflects the detail required for tax reform analysis but also partly reflects the objection to the crude grouping of commodities that used to characterize empirical demand analysis. Demand systems can be expected to cover in excess of 20 commodities and cover several thousand individual observations. In our application the demand for 22 commodities is modelled using household data from 20 annual household expenditure surveys.

* Correspondence to: Prof. Richard Blundell, University College London, Gower Street, London, WC1E 6BT.

CCC 0883-7252/99/030209–24$17·50 *Received 24 March 1997*
 Revised 13 May 1998

For empirical purposes, exact estimation of non-linear equation systems for large data sets with more than a small number of equations has typically been limited by the intrinsic non-linearity of many models. The aim of this paper is to develop and implement a computationally attractive Iterated Linear Least Squares Estimator (ILLE) which is applicable to many popular non-linear demand (and cost share) systems in empirical microeconomics. This class include the Translog and the Almost Ideal models as well as the recent quadratic extensions of these referred to above. All are shown to have the characteristic of *conditional linearity*, that is, they are linear in all the parameters of interest *conditional* on very general functions of explanatory variables and parameters of interest themselves. Many models outside pure empirical consumer demand analysis fit this framework and what we develop here is equally applicable in models of factor demands, etc.

We present theorems that show the consistency and the asymptotic normality of the estimator based on the ILLE procedure. We also prove asymptotic equivalence to the 3SNLS estimator and discuss relative efficiency to the corresponding asymptotically optimal estimator. The estimator is shown to be applicable in the context of a very broad class of non-linear equation systems. The estimator is generalized to allow for endogenous regressors using an extension of the limited information augmented regression technique of Hausman (1978) and Holly and Sargan (1982). Identification conditions are derived for this case and a test for overidentifying restrictions is proposed.

In the application a 22-commodity demand system is estimated using 20 years of repeated cross-sections from the British Family Expenditure Survey covering some five thousand households. A Quadratic Almost Ideal Demand System (QUAIDS), developed in Banks, Blundell and Lewbel (1997), is estimated allowing for the endogeneity of total expenditure. Disposable household income is used as an identifying instrument. The resulting estimates and their implied elasticities are compared to those obtained from an Almost Ideal model. Estimation uses a flexible GAUSS based subroutine which is developed for the QUAIDS model and for demand models in this general class.[1] The subroutine allows for the examination of homogeneity and symmetry as well as the full distribution of elasticities and their precision. Convergence of the ILLE is shown to be rapid and well behaved.

The plan of this paper is as follows. Section 2 characterizes the class of conditionally linear models and reviews the importance of conditional linearity in popular demand models. Section 3 presents the Iterated Linear Least squares Estimator. Section 4 derives the asymptotic properties of the estimator and Section 5 provides the extension to endogenous regressors. Section 6 presents the empirical application to expenditures in the UK Family Expenditure Survey. Section 7 presents a summary and conclusions.

2. CONDITIONALLY LINEAR SYSTEMS

We consider the class of non-linear equations systems that possess a conditional linearity property. That is models which are linear in the parameters *conditional* on the same functions of explanatory variables and parameters themselves. Specifically, we suppose that, when t varies, N endogenous variables $y_{1t}, \ldots, y_{Nt}$ are related to a vector x_t of M conditioning variables by the

[1] Available at cost from the authors.

following system of equations:

$$y_{it} = g(x_t, \theta^0)'\theta_i^0 + u_{it} \qquad i = 1, \ldots, N, \quad t = 1, \ldots, T \tag{1}$$

where $\theta_i^0 \in \mathcal{P} \subset \mathbb{R}^K$ is a parameter with $\theta^0 = (\theta_1^{0'}, \ldots, \theta_N^{0'})'$, and g is a K-vector of functions of x_t and θ^0.

We shall make the following assumptions:

Assumption 1 Let $u_t = (u_{1t}, \ldots, u_{Nt})'$. The sets $\{u_t, x_t\}$, $t = 1, \ldots, T$, are independent sets of identically distributed random vectors.

This assumption is made to allow the use of the Strong Law of Large Numbers and the Central-Limit theorem. It could easily be weakened but at the cost of additional assumptions on, for example, the moments of $\{u_t, x_t\}$. We maintain the independence assumption to simplify the discussion. Note that it does not prevent the inclusion of a particular form of panel data dynamic model.[2] However, the asymptotic results we derive are based on large T and independence across observations $t = 1, \ldots, T$.

Assumption 2 For all $t \in \{1, \ldots, T\}$,

$$\mathrm{E}(u_t \mid x_t) = 0 \qquad \mathrm{E}(u_t u_t' \mid x_t) = \Sigma_0 \tag{2}$$

Estimation will then be based on the following set of identifying conditions:

Assumption 3 There is one and only one solution to the NK identifying restrictions:

$$\mathrm{E}\{g(x_t, \theta)[y_{it} - g(x_t, \theta)'\theta_i]\} = 0 \qquad i = 1, \ldots, N \tag{3}$$

which is θ^0 the true value of the parameter.

2.1. An Example of Conditional Linearity in Demand Analysis

Consider the popular Almost Ideal demand system of Deaton and Muellbauer (1980a) in which budget shares w_{it} for each $i = 1, \ldots, N$ goods are *linear* in the log price N-vector p_t and log total outlay $\ln x_t$. This 'LAIDS' demand system has the form

$$w_{it} = \alpha_i + \gamma_i' p_t + \beta_i(\ln x_t - a(p_t, \theta)) + u_{it}$$

for $i = 1, \ldots, N$ goods, with

$$a(p_t, \theta) = \alpha' p_t + \tfrac{1}{2} p_t' \Gamma p_t$$

where $\alpha = (\alpha_1, \ldots, \alpha_N)$ and

$$\Gamma = \begin{pmatrix} \gamma_1' \\ \vdots \\ \gamma_N' \end{pmatrix}$$

[2] We allow for correlations between u_{it} and u_{jt}, and (y_{it}) could thus refer to a panel where t is the individual index and i the time index.

Note that apart from the function $a(p_t, \theta)$ the system is linear with the same variables appearing on the r.h.s. of each share equation. Each share equation has precisely the form of equation (1).

Many other popular demand systems fit into this formulation (the Translog and Linear Expenditure System, for example), the Almost Ideal model provides a convenient example at this stage. In our application we consider a quadratic extension of the Almost Ideal model. These alternatives and extensions follow naturally. It is also worth pointing out that if the variables in the model are non-stationary, but the error in equation (1) is stationary, then (1) is an example of a non-linear 'cointegrated' model (see Attfield, 1997; Ng, 1995; and Lewbel, 1991, 1996, for example). Although little research has been done on such models, the estimation procedure developed here is computationally efficient, and if the analysis for the stationary case breaks down then standard errors and test critical values can be simulated or bootstrapped.

3. THE ITERATED LEAST SQUARES ESTIMATOR

The estimator we are considering exploits the conditional linearity of equation (1), in particular, first, conditioning on $g(x_t, \theta)$ given θ and estimating each θ_i using a linear moment estimator, then using these estimates to update $g(x_t, \theta)$ and continuing the iteration. Although an apparently natural estimator for conditionally linear systems, by ignoring the derivative of $g(x_t, \cdot)$ it does not solve the minimization problem underlying the identifying condition of assumption (3). As a result its asymptotic properties are not obvious implications of the standard GMM arguments. The estimator consists of the following series of iterations: given an initial value $\theta^{(p)}$ for θ, compute $\theta^{(p+1)}$ by regressing y_t on $g(x_t, \theta^{(p)})$; and repeat the iteration until numerical convergence occurs.

This estimator avoids having to estimate the whole system simultaneously since each step is a set of single-equation estimations. For large demand systems, of the sort estimated in Section 6, this is a distinct advantage. If we let $\hat{\theta}$ represent the limit of $\theta^{(p)}$ when p tends to infinity, then, provided it exists, we show that it is a consistent estimator of θ^0.[3]

A formal definition of the Iterated Linear Least Squares Estimator is as follows. Let $\mathbf{y}_i = (y_{i1}, \ldots, y_{iT})'$ and let $G(\theta)$ be the matrix:

$$G(\theta) = \begin{pmatrix} g(x_1, \theta)' \\ \vdots \\ g(x_T, \theta)' \end{pmatrix} \tag{4}$$

The $(p+1)$th iteration of the algorithm yields the following value for the parameter associated with the ith equation:

$$\theta_i^{(p+1)} = [G(\theta^{(p)})' G(\theta^{(p)})]^{-1} G(\theta^{(p)})' \mathbf{y}_i$$

[3] This estimator is reminiscent of, but differs in important respects from, the classical Iterated OLS estimator for bilinear systems of equations which are symmetrically linear with respect to some parameter vector à conditional on another parameter vector β. In conditionally linear systems, all parameters of interest are contained in the θ_i. The bilinear method is known to converge very slowly. In addition, even where non-linear systems of the form (1) above possess a bilinear form, many more parameters appear in the estimation at each step. Moreover, the equation-by-equation characteristic of (1) is lost. Finally, it should be stressed that popular forms are not bilinear.

Hence any limit value of any such recursive sequence will thus be a fixed point of the N-vector $\Psi(\theta)$ of functions in which ith component is

$$\Psi_i(\theta) = [G(\theta)'G(\theta)]^{-1}G(\theta)'\mathbf{y}_i \tag{5}$$

This yields the following precise definition:

Definition 1 An Iterated Linear Least squares Estimator for model (1) is a fixed point of $\Psi(\theta)$.

4. ASYMPTOTIC PROPERTIES

Let $\hat{\theta}$ denote any fixed point of the N-vector $\Psi(\theta)$. Under standard stochastic uniform convergence assumptions that we state in Appendix A (see, for example, White, 1994, Chapter 3), Theorem 1 shows that $\hat{\theta}$ is a consistent estimator of θ^0.

Theorem 1. Consistency Under the regularity assumptions stated in Appendix A, $\hat{\theta}$ converges almost surely to θ^0.

Proof See Appendix A. ■

The asymptotic normality of the estimator is another simple consequence of the asymptotic expansion used in the proof of the convergence theorem.

Theorem 2. Asymptotic normality Under the regularity assumptions stated in Appendix A, $\hat{\theta}$ is asymptotically normal and

$$\sqrt{T}(\hat{\theta} - \theta^0) \rightsquigarrow \mathcal{N}(0, J_0^{-1}(\Sigma_0 \otimes K_0)(J_0')^{-1}) \tag{6}$$

where $\rightsquigarrow$ is weak convergence and

$$K_0 = \mathrm{E}[g(x_t, \theta^0)g(x_t, \theta^0)']$$
$$J_0 = I_N \otimes K_0 + \mathrm{E}\left[\left(\Theta_0' \frac{\partial g(x_t, \theta^0)}{\partial \theta'}\right) \otimes g(x_t, \theta^0)\right]$$

where Θ_0 is the matrix $(\theta_1^0, \ldots, \theta_N^0)$ (i.e. $\theta^0 = \mathrm{vec}(\Theta_0)$).

Proof See Appendix A. ■

An estimator of the asymptotic variance–covariance matrix of $\hat{\theta}$, presented in (6) above, is computed as follows. Σ_0 is estimated using the usual formula:

$$\hat{\Sigma} = \frac{1}{T}(Y - G(\hat{\theta})\hat{\Theta})'(Y - G(\hat{\theta})\hat{\Theta})$$

where

$$Y = (\mathbf{y}_1, \ldots, \mathbf{y}_N) \qquad \hat{\Theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_N)$$

Then, $\hat{\theta}$ solves

$$[I_N \otimes G(\hat{\theta})'][\mathbf{y} - (I_N \otimes G(\hat{\theta}))\hat{\theta}] = 0 \tag{7}$$

Let

$$\nabla(\hat{\theta}) = \frac{\partial([I_N \otimes G(\hat{\theta})]\hat{\theta})}{\partial\theta'}$$

be the Jacobian of $\theta \mapsto [I_N \otimes G(\theta)]\theta$ evaluated at point $\theta = \hat{\theta}$. A first-order Taylor series expansion of equation (7) yields

$$[I_N \otimes G(\theta^0)']\nabla(\theta^0)(\hat{\theta} - \theta^0) = [I_N \otimes G(\theta^0)']\mathbf{u} + o_P(1)$$

where $\mathbf{u} = (u_{11}, \ldots, u_{1T}, u_{21}, \ldots, u_{2T}, \ldots, u_{N1}, \ldots, u_{NT})'$. The term TK_0 can thus be estimated by

$$\hat{K} = G(\hat{\theta})'G(\hat{\theta})$$

To obtain an estimator of the term TJ_0, simply compute

$$\hat{J} = [I_N \otimes G(\hat{\theta})']\nabla(\hat{\theta})$$

An estimator of the asymptotic variance–covariance matrix of $\hat{\theta}$ is therefore

$$\hat{V}_{\mathrm{as}}\hat{\theta} = \hat{J}^{-1}(\hat{\Sigma} \otimes \hat{K})(\hat{J}')^{-1}$$

Note that if Σ^0 is non-singular, then

$$\hat{V}_{\mathrm{as}}\hat{\theta} = [\nabla(\hat{\theta})'(\hat{\Sigma}^{-1} \otimes G(\hat{\theta})\hat{K}^{-1}G(\hat{\theta})')\nabla(\hat{\theta})]^{-1}$$

It is then possible to show that $\hat{\theta}$ is asymptotically equivalent to a non-linear three-stage estimator (NL3S).

Corollary 1 The ILLE $\hat{\theta}$ is asymptotically equivalent to the Non-Linear Three Stage Least Square (NL3S) estimator $\tilde{\theta}$ that is obtained for an identical choice of $G(\theta^0)$ as instruments for each equation; that is,

$$\tilde{\theta} = \arg\min_{\theta}(\mathbf{y} - [I_N \otimes G(\theta)]\theta)'\left(\Sigma_0^{-1} \otimes \mathrm{P}_{G(\theta^0)}\right)(\mathbf{y} - [I_N \otimes G(\theta)]\theta) \tag{8}$$

where $\mathbf{y} = (\mathbf{y}_1', \ldots, \mathbf{y}_N')'$ and $\mathrm{P}_{G(\theta^0)}$ is the matrix of the projection on the space spanned by the columns of $G(\theta^0)$.

Proof See Appendix A. ■

The choice of $\Sigma_0^{-1} \otimes \mathrm{P}_{G(\theta^0)}$ as a metric for the NL3S class of estimators will not be optimal. The variance of $\sqrt{T}$ times the non-linear least squares estimator obtained by minimizing

$$(\mathbf{y} - [I_N \otimes G(\theta)]\theta)'\left(\Sigma_0^{-1} \otimes I_T\right)(\mathbf{y} - [I_N \otimes G(\theta)]\theta)$$

w.r.t. θ, is

$$\text{plim } T\Big[\nabla(\theta^0)'\Big(\Sigma_0^{-1} \otimes I_T\Big)\nabla(\theta^0)\Big]^{-1}$$

whereas that of the ILLE is

$$\text{plim } T\Big[\nabla(\theta^0)'\Big(\Sigma_0^{-1} \otimes \mathrm{P}_{G(\theta^0)}\Big)\nabla(\theta^0)\Big]^{-1} = \text{plim } T\Big[\hat{\nabla}(\theta^0)'\Big(\Sigma_0^{-1} \otimes I_T\Big)\hat{\nabla}(\theta^0)\Big]^{-1}$$

where

$$\hat{\nabla}(\theta^0) = (I_N \otimes \mathrm{P}_{G(\theta^0)})\nabla(\theta^0)$$

The extent to which the two variances differ is therefore directly function of the difference between the norm of $\nabla(\theta^0)$ in the metric $\Sigma_0^{-1} \otimes I_T$ and the norm of its projection $\hat{\nabla}(\theta^0)$ in the same metric. If both $\nabla(\theta^0)$ and $I_N \otimes G(\theta^0)$ described the same vector space, then these two variances would be identical. But this would be possible only if $G(\theta^0)$ were not a function of θ^0.[4]

5. ENDOGENOUS REGRESSORS

5.1. An Augmented Regression Approach

Suppose that Assumption 2 is not satisfied because a subset $\tilde{x}_t$ of $\tilde{M} \leqslant M$ explanatory variables cannot be assumed exogenous. Instead, suppose that there exists (z_t) a sequence of random vectors of H instrumental variables, including those explanatory variables which are indeed exogenous, such that

$$\mathrm{E}(u_t \mid z_t) = 0 \qquad \mathrm{E}(u_t u_t' \mid z_t) = \Sigma_0 \tag{9}$$

Initially suppose that the model is linear: $g(x_t, \theta^0) = x_t$. Following the Holly and Sargan (1982) development of the Hausman (1978) test, we may use an augmented regression framework to test and correct for endogeneity. First regress $\tilde{x}_t$ on z_t, compute residuals $\hat{v}_t$, and then regress y_{it} on both x_t and $\hat{v}_t$. The OLS estimator of the parameters of x_t in this augmented regression is identical to the Two-Stage Least Squares (2SLS) estimator. Moreover, testing for the significance of the coefficient of $\hat{v}_t$ is a test of the exogeneity of x_t.

This procedure has proved extremely useful in empirical application and here we adapt it for use when the model is of the form (1). The assumption of orthogonality between residuals u_{it} and instruments z_t used in the linear model is no longer sufficient. Instead, we assume that the errors u_{it} in (1) have the orthogonal decomposition

$$u_{it} = \rho_i^{0'} v_t + \varepsilon_{it}$$

[4] An efficient estimator can nevertheless be obtained in a second step by a one-step Newton–Raphson iteration, that is, by regressing

$$\mathbf{y} - [I_N \otimes G(\hat{\theta})]\hat{\theta} + \nabla(\hat{\theta})\hat{\theta}$$

on the columns of $\nabla(\hat{\theta})$.

where v_t is given by

$$\tilde{x}_t = C_0 z_t + v_t$$

We denote $\rho^0 = (\rho_1^{0'}, \ldots, \rho_{\tilde{M}}^{0'})'$, $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$, and we replace Assumptions (2) and (3) by the following:

Assumption 4 The sets $\{\varepsilon_t, v_t, z_t\}$, $t = 1, \ldots, T$, are independent sets of identically distributed random vectors.

Assumption 5 For all $t \in \{1, \ldots, T\}$,

$$\mathrm{E}(\varepsilon_t \mid z_t, v_t) = 0 \qquad \mathrm{E}(v_t \mid z_t) = 0 \tag{10}$$

Assumption 6 For all $t \in \{1, \ldots, T\}$,

$$\mathrm{E}(\varepsilon_t \varepsilon_t' \mid z_t, v_t) = \Delta_0 \qquad \mathrm{E}(v_t v_t' \mid z_t) = \Omega_0 \tag{11}$$

Finally, estimation is based on the following set of identifying conditions:

Assumption 7 There is one and only one solution to the $N(K + \tilde{M})$ identifying restrictions:

$$\mathrm{E}\left\{ \begin{pmatrix} g(x_t, \theta) \\ v_t \end{pmatrix} [y_{it} - g(x_t, \theta)'\theta_i - \rho_i' v_t] \right\} = 0 \qquad i = 1, \ldots, N \tag{12}$$

which are θ^0 and ρ^0 the true value of the parameters.

If v_t was directly observable then the estimation problem would be exactly similar to the one we have analysed in the preceding section. However, C_0 in $v_t = \tilde{x}_t - C_0 z_t$ is a nuisance parameter which must be first consistently estimated. Let

$$\begin{aligned} \hat{C} &= \left(\sum_{t=1}^{T} \tilde{x}_t z_t' \right) \left(\sum_{t=1}^{T} z_t z_t' \right)^{-1} \\ &= C_0 + \left(\sum_{t=1}^{T} v_t z_t' \right) \left(\sum_{t=1}^{T} z_t z_t' \right)^{-1} \end{aligned}$$

the OLS estimator of C_0, and let $\hat{v}_t = \tilde{x}_t - \hat{C} z_t$.

The ILLE procedure considers the estimators $\hat{\theta}$ and $\hat{\rho}$ obtained from the iterated least squares regression of y_{it} on $g(x_t, \theta)$ and $\hat{v}_t$:

$$\sum_{t=1}^{T} \begin{pmatrix} g(x_t, \hat{\theta}) \\ \hat{v}_t \end{pmatrix} [y_{it} - g(x_t, \hat{\theta})'\hat{\theta}_i - \hat{\rho}_i' \hat{v}_t] = 0$$

Note that

$$y_{it} - g(x_t, \hat{\theta})'\hat{\theta}_i - \hat{\rho}_i'\hat{v}_t = y_{it} - g(x_t, \hat{\theta})'\hat{\theta}_i - \hat{\rho}_i' v_t + \hat{\rho}_i'(v_t - \hat{v}_t) \tag{13}$$

and

$$\sum_{t=1}^{T} \begin{pmatrix} g(x_t, \hat{\theta}) \\ \hat{v}_t \end{pmatrix} \hat{\rho}_i'(v_t - \hat{v}_t) = \sum_{t=1}^{T} \begin{pmatrix} g(x_t, \hat{\theta}) z_t' \\ \mathbf{0}_{\tilde{M}\times H} \end{pmatrix} \left(\sum_{t=1}^{T} z_t z_t' \right)^{-1} \sum_{t=1}^{T} z_t v_t' \hat{\rho}_i$$

where $\mathbf{0}_{\tilde{M}\times H}$ is the $\tilde{M}$ by H matrix of zeroes.[5] It is then straightforward to adapt the sequence of hypotheses and the proofs of the results of the preceding section to show under which regularity conditions this two-stage ILLE yields consistent and asymptotically normal estimates.

We have that

$$\begin{aligned} 0 &= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \begin{pmatrix} g(x_t, \hat{\theta}) \\ \hat{v}_t \end{pmatrix} [y_{it} - g(x_t, \hat{\theta})'\hat{\theta}_i - \hat{\rho}_i'\hat{v}_t] \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \begin{pmatrix} g(x_t, \hat{\theta}) \\ \hat{v}_t \end{pmatrix} [y_{it} - g(x_t, \hat{\theta})'\hat{\theta}_i - \hat{\rho}_i' v_t] \\ &\quad + \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \begin{pmatrix} g(x_t, \hat{\theta}) z_t' \\ \mathbf{0}_{\tilde{M}\times H} \end{pmatrix} \left(\sum_{t=1}^{T} z_t z_t' \right)^{-1} \sum_{t=1}^{T} z_t v_t' \hat{\rho}_i \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \begin{pmatrix} g(x_t, \theta^0) \\ v_t \end{pmatrix} \varepsilon_{it} + \frac{1}{T} \sum_{t=1}^{T} \begin{pmatrix} g(x_t, \theta 0 z_t' \\ \mathbf{0}_{\tilde{M}\times H} \end{pmatrix} \left(\frac{1}{T} \sum_{t=1}^{T} z_t z_t' \right)^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} z_t v_t' \rho_i^0 \\ &\quad - \left[\frac{1}{T} \sum_{t=1}^{T} \begin{pmatrix} g(x_t, \theta^0) \\ v_t \end{pmatrix} \begin{pmatrix} g(x_t, \theta^0) \\ v_t \end{pmatrix}' \right] \sqrt{T} \begin{pmatrix} \hat{\theta}_i - \theta_i^0 \\ \hat{\rho}_i - \rho_i^0 \end{pmatrix} \\ &\quad - \left[\frac{1}{T} \sum_{t=1}^{T} \begin{pmatrix} g(x_t, \theta^0) \\ v_t \end{pmatrix} \theta_i^{0'} \frac{\partial g(x_t, \theta^0)}{\partial \theta'} \right] \sqrt{T}(\hat{\theta} - \theta^0) + o_P(1) \end{aligned}$$

This equality is true for all $i = 1, \ldots, N$. Moreover,

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_t \otimes \begin{pmatrix} g(x_t, \theta^0) \\ v_t \end{pmatrix} + \frac{1}{\sqrt{T}} \sum_{t=1}^{T} R_0' v_t \otimes D_0 L_0 z_t \tag{14}$$

---

[5] Since

$$\begin{aligned} \sum_{t=1}^{T} \begin{pmatrix} g(x_t, \hat{\theta}) \\ \hat{v}_t \end{pmatrix} \hat{\rho}_i'(v_t - \hat{v}_t) &= \sum_{t=1}^{T} \begin{pmatrix} g(x_t, \hat{\theta}) \\ \hat{v}_t \end{pmatrix} \hat{\rho}_i'(\hat{C} - C_0) z_t \\ &= \sum_{t=1}^{T} \begin{pmatrix} g(x_t, \hat{\theta}) \\ \hat{v}_t \end{pmatrix} z_t'(\hat{C} - C_0)' \hat{\rho}_i \\ &= \sum_{t=1}^{T} \begin{pmatrix} g(x_t, \hat{\theta}) z_t' \\ \mathbf{0}_{\tilde{M}\times H} \end{pmatrix} \left(\sum_{t=1}^{T} z_t z_t' \right)^{-1} \sum_{t=1}^{T} z_t v_t' \hat{\rho}_i \end{aligned}$$

has asymptotically a normal distribution with zero mean and variance

$$\Delta_0 \otimes K_0 + R_0' \Omega_0 R_0 \otimes D_0 L_0 D_0'$$

where $R_0 = (\rho_1, \ldots, \rho_N)$, and with

$$K_0 = \mathrm{E}\left[\begin{pmatrix} g(x_t, \theta^0) \\ v_t \end{pmatrix}\begin{pmatrix} g(x_t, \theta^0) \\ v_t \end{pmatrix}'\right]$$

$$D_0 = \begin{pmatrix} \mathrm{E}[g(x_t, \theta^0) z_t'] \\ \mathbf{0}_{M \times H} \end{pmatrix}$$

$$L_0 = (\mathrm{E}[z_t z_t'])^{-1}$$

Note that $D_0 L_0$ is estimated by regressing with OLS the elements of $g(z_t, \theta^0)$ on the instruments z_t, and that estimates for Δ_0 and Ω_0 are obtained by standard mean-sum-of-squared residual formulas.

It follows that

$$\sqrt{T}\,\mathrm{vec}\begin{pmatrix} \hat{\Theta} - \Theta^0 \\ \hat{R} - R^0 \end{pmatrix} \rightsquigarrow \mathcal{N}(0, J_0^{-1}(\Delta_0 \otimes K_0 + R_0' \Omega_0 R_0 \otimes D_0 L_0 D_0')(J_0')^{-1})$$

with

$$J_0 = I_N \otimes K_0 + Q_0$$

where Θ_0 is the matrix $(\theta_1^0, \ldots, \theta_N^0)$ (i.e. $\theta^0 = \mathrm{vec}(\Theta_0)$) and where Q_0 is the N by N block-matrix whose i by j block is

$$\mathrm{E}\left[\begin{pmatrix} g(x_t, \theta^0) \\ v_t \end{pmatrix} \theta_i^{0\prime} \frac{\partial g(x_t, \theta^0)}{\partial \theta_j'}; \mathbf{0}_{(K+\tilde{M}) \times \tilde{M}}\right]$$

where $\mathbf{0}_{(K+\tilde{M}) \times \tilde{M}}$ is the $K + \tilde{M}$ by $\tilde{M}$ matrix of zeroes.

Notice that the variance formula is identical to that obtained using Theorem 2, except for the term $R_0' \Omega_0 R_0 \otimes D_0 L_0 D_0'$ which is due to the additional term

$$\sum_{t=1}^{T} \begin{pmatrix} g(x_t, \hat{\theta}) \\ \hat{v}_t \end{pmatrix} \hat{\rho}_i'(v_t - \hat{v}_t)$$

in equation (13) arising from the approximation of v_t by $\hat{v}_t$.

5.2. Testing for Overidentifying Restrictions

Exactly as in the linear case, one can test for overidentifying restrictions by regressing the system residuals on instruments. Let the system residuals be given by

$$\hat{\varepsilon}_i = \mathbf{y}_i - G(\hat{\theta})\hat{\theta}_i - \hat{V}\hat{\rho}_i$$

where $\hat{V}$ is the matrix $(\hat{v}_t')$.

Let $\hat{\gamma}_i$ be the coefficient of the OLS regression of $\hat{\varepsilon}_i$ on the columns of $\tilde{Z}$, where $\tilde{Z} = (\tilde{z}_t')$ is a T by $\tilde{H}$ matrix of instruments which are not already counted in the list of explanatory variables x_t:

$$\hat{\gamma} = [I_N \otimes [\tilde{Z}'\tilde{Z}]^{-1}\tilde{Z}'][\mathbf{y} - (I_N \otimes G(\hat{\theta}))\hat{\theta} - (I_N \otimes \hat{V})\hat{\rho}] \tag{15}$$

where $\hat{\theta} = (\hat{\theta}_1', \dots, \hat{\theta}_N')'$ and $\hat{\rho} = (\hat{\rho}_1', \dots, \hat{\rho}_N')'$ and $V = (v_t')$. By construction, the columns of $\hat{V}$ are orthogonal to the columns of $\tilde{Z}$, hence

$$\hat{\gamma} = [I_N \otimes [\tilde{Z}'\tilde{Z}]^{-1}\tilde{Z}'][\mathbf{y} - (I_N \otimes G(\hat{\theta}))\hat{\theta}]$$

Under the identifying Assumptions 4, 5 and 7, $\hat{\gamma}$ converges to 0 a.s. Moreover, a first-order expansion of the left-hand side of equation (15) shows that

$$\hat{\gamma} = [I_N \otimes [\tilde{Z}'\tilde{Z}]^{-1}\tilde{Z}'][\boldsymbol{\varepsilon} + (I_N \otimes V)\rho - \nabla(\theta^0)(\hat{\theta} - \theta^0)] + o_P(\sqrt{T})$$

where $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1', \dots, \boldsymbol{\varepsilon}_N')'$ with $\boldsymbol{\varepsilon}_i = (\varepsilon_{it})$.

Using the expansion of $\sqrt{T}(\hat{\theta} - \theta^0)$ yields (see equation (14))

$$\sqrt{T}(\hat{\theta} - \theta^0) = [J_0^{-1}]_\theta \left\{ \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_t \otimes \begin{pmatrix} g(x_t, \theta^0) \\ v_t \end{pmatrix} + \frac{1}{\sqrt{T}} \sum_{t=1}^{T} R_0' v_t \otimes D_0 L_0 z_t \right\} + o_P(1)$$

where $[J_0^{-1}]_\theta$ is the rectangular matrix formed by using the NK rows of J_0^{-1} which correspond to parameter θ. Hence, denoting $V = (v_t')$:

$$\begin{aligned}
\sqrt{T}\hat{\gamma} = \left[I_N \otimes \left[\frac{1}{T} \tilde{Z}'\tilde{Z} \right]^{-1} \right] \Bigg\{ & \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_t \otimes \tilde{z}_t \\
& - \frac{1}{T}[(I_N \otimes \tilde{Z}')\nabla(\theta^0)][J_0^{-1}]_\theta \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \varepsilon_t \otimes \begin{pmatrix} g(x_t, \theta^0) \\ v_t \end{pmatrix} \\
& + \frac{1}{\sqrt{T}} \sum_{t=1}^{T} R_0' v_t \otimes \tilde{z}_t - \frac{1}{T}[(I_N \otimes \tilde{Z}')\nabla(\theta^0)][J_0^{-1}]_\theta \frac{1}{\sqrt{T}} \sum_{t=1}^{T} R_0' v_t \otimes D_0 L_0 z_t \Bigg\} \\
& + o_P(1)
\end{aligned}$$

Let

$$\tilde{L}_0 = (\mathrm{E}[\tilde{z}_t \tilde{z}_t'])^{-1}$$

$$H_0 = \operatorname*{plim}_{T\to\infty}\left(I_N \otimes \left[\frac{1}{T}\tilde{Z}'\tilde{Z}\right]^{-1}\right)\frac{1}{T}[(I_N \otimes \tilde{Z}')\nabla(\theta^0)][J_0^{-1}]_\theta$$

$$= (I_N \otimes \tilde{L}_0)\left\{I_N \otimes \mathrm{E}(\tilde{z}_t g(x_t, \theta^0)') + \mathrm{E}\left[\Theta_0' \frac{\partial g(x_t, \theta^0)}{\partial \theta'} \otimes \tilde{z}_t\right]\right\}[J_0^{-1}]_\theta$$

The following asymptotic distribution of $\hat{\gamma}$ can then be deduced from the previous expansion:

$$\begin{aligned}\sqrt{T}\hat{\gamma} \rightsquigarrow \mathcal{N}(0, &\Delta_0 \otimes \tilde{L}_0 - (\Delta_0 \otimes \tilde{L}_0 \tilde{D}_0')H_0' - H_0(\Delta_0 \otimes \tilde{D}_0 \tilde{L}_0) \\ &+ R_0'\Omega_0 R_0 \otimes \tilde{L}_0 - (R_0'\Omega_0 R_0 \otimes \tilde{L}_0 Q_0' L_0 D_0')H_0' - H_0(R_0'\Omega_0 R_0 \otimes D_0 L_0 \tilde{\tilde{L}}_0 \tilde{L}_0) \\ &+ H_0(\Delta_0 \otimes K_0 + R_0'\Omega_0 R_0 \otimes D_0 L_0 D_0')H_0')\end{aligned}$$

where

$$\tilde{\tilde{L}}_0 = \mathrm{E}[z_t \tilde{z}_t']$$

$$\tilde{D}_0 = \begin{pmatrix} \mathrm{E}[g(x_t, \theta^0)\tilde{z}_t'] \\ \mathbf{0}_{\tilde{M}\times\tilde{H}} \end{pmatrix}$$

6. APPLICATION

6.1. The Data

The data used in this study are drawn from the detailed expenditure diaries of the UK Family Expenditure Survey (FES) for the period 1974–1993. Prices are measured quarterly. We selected a reasonably homogeneous sample consisting of married couples with two children. This selection is chosen so as to abstract from demographic differences and to focus on the price and income terms. We study the purchases of 22 non-durables and services comprising: wine, spirits, beer, six food categories, household fuel, clothing, household services, personal goods and services, leisure services, fares, tobacco, motoring and petrol. To avoid the problem of zero expenditures in the tobacco and petrol categories we only select car-owning households that have at least one adult who smokes. The definitions of all goods and their mean shares are presented in the first column of Table BI in Appendix B.

The demand model we estimate is the QUAIDS model introduced by Banks, Blundell and Lewbel (1997). We turn first to a brief description of that model.

6.2. The QUAIDS Model

The Quadratic Almost Ideal Demand System (QUAIDS) provides a useful framework for demand analysis. It processes many of the attractive features of the Almost Ideal Model and the Translog Model while allowing for more non-linear Engel curve behaviour. It is also *conditionally linear* in price aggregators.

As before, we write w_{it} as the expenditure share on commodity i for observation t with total budget x_t and the log price N-vector p_t. The QUAIDS model expenditure shares have the form

$$w_{it} = \alpha_i + \gamma_i' p_t + \beta_i(\ln x_t - a(p_t, \theta)) + \lambda_i \frac{(\ln x_t - a(p_t, \theta))^2}{b(p_t, \theta)} + u_{it} \tag{16}$$

with the following non-linear price aggregators:

$$a(p_t, \theta) = \alpha' p_t + \tfrac{1}{2} p_t' \Gamma p_t$$
$$b(p_t, \theta) = \exp \beta' p_t$$

where $\alpha = (\alpha_1, \ldots, \alpha_N)$ and

$$\Gamma = \begin{pmatrix} \gamma_1' \\ \vdots \\ \gamma_N' \end{pmatrix}$$

The standard *L*inear Almost Ideal Demand System (LAIDS), described in Section 6.1, simply sets $\lambda_i = 0$ across all commodities. The denominator $b(p_t, \theta)$ in the quadratic term of the share equation (16) is required to maintain the integrability of the quadratic expenditure share system (see Banks, Blundell and Lewbel, 1997).

Given the likely correlation between u_{it} and the log total budget variable $\ln x_t$, we follow the approach for estimation with endogenous regressors developed in Section 5. We augment the QUAIDS specification (16) by writing

$$u_{it} = \rho_i v_t + \varepsilon_{it} \text{ for goods } i = 1, \ldots, N$$

and assuming $E(\varepsilon_{it} \mid x_t, p_t) = 0$. To construct v_t for use in the augmented system we first estimate a reduced-form equation for $\ln x$. This is presented in Table BII of Appendix B. In addition to a linear trend, seasonal dummies and relative prices, income and income squared are used as additional instruments. They are clearly significant.

6.3. Estimation Results

For both QUAIDS and LAIDS models convergence of the ILLE was achieved at a high level of tolerance within six iterations. Table I presents the estimated income coefficients for the QUAIDS model.[6] The corresponding LAIDS coefficient estimates are in Table BIII of Appendix B. The importance of the quadratic expenditure terms for many commodities is

[6] Full estimates are available on request from the authors as is the complete Gauss language subroutine of the ILLE for the QUAIDS model.

clearly established and suggests that the LAIDS model is misspecified. This is also reflected in the estimated elasticities presented below. For some commodities, mainly food items, shares linear in log expenditure is not strongly rejected. For these commodities the Working-Leser model underlying the LAIDS specification does seem to provide a sufficient description of Engel curve behaviour confirming the findings in the Banks, Blundell and Lewbel (1997) study which simply considered the aggregate of food items. For other goods, such as entertainment, the quadratic term is very important. The attraction of our proposed estimator is that it allows a system wide assessment of the properties of demand at a very fine level of disaggregation.

The exogeneity of log expenditure is also clearly rejected through the $|t|$-values for the significance of the $\hat{v}$ residual. The reduced form from which $\hat{v}$ is derived is presented in Table BII of Appendix B. The adjustment for endogeneity is therefore included in all results and all standard errors are adjusted for the non-linear price terms as well as the addition of the residual $\hat{v}_{it}$. The overidentification tests do not indicate any serious difficulties with the instrument set and the correction for endogeneity of $\ln x$ is included throughout the following results. The incorrect omission of the $(\ln x)^2$ term induces rejection of the overidentification test for many goods in the case of the LAIDS model as can be seen from Table BIII. All standard errors are adjusted for the non-linear price terms as well as the addition of the residual $\hat{v}_{it}$.

Homogeneity in prices and total expenditure is imposed throughout. There is no evidence of the rejection of homogeneity for any of the 22-commodity share equations as the final column of Table I shows. However, symmetry was less easily acceptable with a χ^2 statistic of 353 and degrees of freedom 210. The results in Table I do not assume symmetry, but the elasticity results presented in Table II do maintain symmetry.

Table I. Estimated income effects, QUAIDS

Commodity	$\ln x$	$(\ln x)^2$	Exog. $\hat{v}\|t\|$	Ov. id. χ^2_2	Hom. $\|t\|$
Beer	– 0·0434 (0·013)	– 0·0066 (0·002)	0·618	0·659	0·013
Wine	– 0·0046 (0·011)	– 0·0029 (0·002)	3·890	4·452	0·065
Spirits	– 0·0060 (0·009)	– 0·0034 (0·001)	0·921	4·901	0·043
Bread	– 0·0117 (0·006)	0·0018 (0·001)	0·625	4·926	0·098
Meat	– 0·0559 (0·013)	– 0·0047 (0·002)	2·201	0·525	0·010
Dairy	0·0016 (0·009)	0·0060 (0·001)	0·603	1·492	0·186
Vegetables	– 0·0143 (0·014)	0·0009 (0·002)	0·618	0·225	0·010
Other food	– 0·0206 (0·011)	0·0007 (0·002)	1·616	1·009	0·208
Food out	– 0·0506 (0·013)	– 0·0128 (0·002)	4·693	2·365	0·136
Electricity	0·0108 (0·012)	0·0071 (0·002)	1·576	3·407	0·059
Gas	– 0·0187 (0·013)	– 0·0014 (0·002)	2·359	0·198	0·031
Adult clothing	– 0·0226 (0·021)	– 0·0087 (0·003)	2·246	1·001	0·057
Children's clothing	– 0·0844 (0·016)	– 0·0167 (0·003)	3·585	1·588	0·038
Household services	0·1518 (0·013)	0·0189 (0·002)	5·422	4·482	0·194
Personal goods	0·0260 (0·011)	0·0007 (0·002)	1·273	4·406	0·428
Leisure goods	– 0·0330 (0·012)	– 0·0066 (0·002)	0·096	0·610	0·198
Entertainment	0·2779 (0·018)	0·0321 (0·003)	7·558	3·838	0·059
Leisure services	– 0·0064 (0·016)	0·0008 (0·002)	0·390	0·985	0·016
Fares	0·0019 (0·012)	– 0·0026 (0·002)	5·423	1·884	0·022
Motoring	0·0092 (0·017)	0·0030 (0·003)	5·504	3·929	
Petrol	– 0·0840 (0·045)	– 0·0104 (0·007)	2·525	0·749	0·010
Tobacco	– 0·0229 (0·015)	0·0047 (0·002)	5·289	0·403	0·054

Note: Estimated standard errors in parentheses.

Table II. Estimated budget and own-price uncompensated elasticities. Quartiles, QUAIDS

Commodity	η^{50}	η^{25}	η^{75}	ε^{50}	ε^{25}	ε^{75}
Beer	0·992 (0·078)	1·056 (0·096)	0·889 (0·080)	−2·181 (1·300)	−2·353 (1·459)	−2·320 (1·433)
Wine	2·398 (0·267)	2·643 (0·272)	2·179 (0·278)	−0·741 (0·737)	−0·717 (0·824)	−0·750 (0·707)
Spirits	1·804 (0·136)	1·788 (0·117)	2·230 (0·251)	−2·081 (0·963)	−1·999 (0·901)	−2·821 (1·707)
Bread	0·412 (0·039)	0·455 (0·034)	0·418 (0·041)	−0·317 (0·609)	−0·379 (0·556)	−0·297 (0·630)
Meat	0·701 (0·036)	0·725 (0·039)	0·638 (0·040)	−0·316 (0·385)	−0·328 (0·387)	−0·246 (0·434)
Dairy	0·405 (0·029)	0·430 (0·029)	0·297 (0·042)	0·050 (0·192)	−0·039 (0·178)	0·343 (0·268)
Vegetables	0·568 (0·035)	0·600 (0·036)	0·624 (0·032)	−0·308 (0·107)	−0·367 (0·091)	−0·382 (0·095)
Other food	0·565 (0·034)	0·607 (0·033)	0·572 (0·036)	−0·606 (0·224)	−0·643 (0·196)	−0·605 (0·223)
Food out	1·627 (0·102)	1·629 (0·075)	1·365 (0·058)	−0·364 (1·125)	−0·443 (0·985)	−0·535 (0·803)
Electricity	0·234 (0·043)	0·113 (0·054)	0·226 (0·050)	−0·803 (0·243)	−0·790 (0·266)	−0·775 (0·275)
Gas	0·640 (0·069)	0·584 (0·097)	0·744 (0·048)	−0·168 (0·289)	0·004 (0·336)	−0·442 (0·187)
Adult clothing	1·894 (0·155)	1·867 (0·118)	1·878 (0·185)	−1·191 (1·179)	−1·183 (1·058)	−1·189 (1·323)
Children's clothing	1·392 (0·093)	1·398 (0·063)	1·261 (0·093)	−1·213 (1·455)	−1·188 (1·226)	−1·188 (1·438)
Household services	1·732 (0·280)	1·491 (0·121)	1·612 (0·188)	−2·634 (0·985)	−2·344 (0·672)	−2·038 (0·573)
Personal goods	1·400 (0·068)	1·434 (0·075)	1·448 (0·071)	−2·751 (1·691)	−2·913 (1·880)	−2·925 (1·896)
Leisure goods	1·146 (0·049)	1·187 (0·054)	1·137 (0·065)	−1·806 (1·979)	−1·862 (1·048)	−2·093 (1·357)
Entertainment	5·797 (0·661)	7·840 (0·740)	4·011 (0·277)	−3·423 (0·500)	−4·025 (0·519)	−2·296 (0·886)
Leisure services	0·301 (0·107)	0·387 (0·106)	0·370 (0·076)	−1·210 (0·534)	−1·178 (0·454)	−1·194 (0·488)
Fares	2·410 (0·285)	2·772 (0·362)	2·173 (0·202)	−1·598 (1·239)	−1·719 (1·553)	−1·528 (1·153)
Motoring	0·802 (0·088)	0·758 (0·107)	0·848 (0·078)	−1·331 (1·265)	−1·371 (1·414)	−1·296 (1·136)
Petrol	0·805 (0·080)	0·804 (0·119)	0·696 (0·087)	−0·361 (0·211)	−0·182 (0·276)	−0·253 (0·265)
Tobacco	0·244 (0·047)	0·185 (0·054)	0·201 (0·064)	−0·469 (0·270)	−0·444 (0·282)	−0·412 (0·293)

Table III. Estimated budget and own-price uncompensated elasticities. Quartiles, LAIDS

Commodity	η^{50}	η^{25}	η^{75}	ε^{50}	ε^{25}	ε^{75}
Beer	0·856 (0·073)	0·836 (0·084)	0·826 (0·091)	−2·207 (1·363)	−2·373 (1·519)	−2·462 (1·617)
Wine	2·313 (0·242)	2·399 (0·252)	2·258 (0·183)	−0·771 (0·784)	−0·754 (0·846)	−0·778 (0·766)
Spirits	1·682 (0·126)	1·618 (0·105)	2·278 (0·246)	−2·158 (0·989)	−2·050 (0·915)	−3·155 (1·934)
Bread	0·430 (0·032)	0·480 (0·029)	0·411 (0·036)	−0·317 (0·602)	−0·374 (0·554)	−0·292 (0·626)
Meat	0·658 (0·034)	0·666 (0·034)	0·611 (0·042)	−0·298 (0·388)	−0·313 (0·392)	−0·203 (0·450)
Dairy	0·463 (0·026)	0·504 (0·025)	0·322 (0·039)	0·001 (0·188)	−0·072 (0·177)	0·259 (0·257)
Vegetables	0·574 (0·033)	0·609 (0·030)	0·617 (0·032)	−0·315 (0·106)	−0·368 (0·091)	−0·381 (0·094)
Other food	0·562 (0·032)	0·606 (0·029)	0·558 (0·035)	−0·616 (0·222)	−0·652 (0·196)	−0·611 (0·224)
Food out	1·500 (0·076)	1·432 (0·061)	1·367 (0·047)	−0·301 (1·204)	−0·399 (1·019)	−0·491 (0·878)
Electricity	0·345 (0·039)	0·279 (0·044)	0·285 (0·050)	−0·822 (0·235)	−0·807 (0·260)	−0·805 (0·259)
Gas	0·618 (0·067)	0·532 (0·085)	0·742 (0·049)	−0·152 (0·288)	0·034 (0·340)	−0·424 (0·189)
Adult clothing	1·800 (0·133)	1·683 (0·106)	1·923 (0·128)	−1·186 (1·270)	−1·162 (1·083)	−1·205 (1·463)
Children's clothing	1·249 (0·069)	1·199 (0·054)	1·259 (0·067)	−1·168 (1·580)	−1·138 (1·266)	−1·174 (1·646)
Household services	1·884 (0·113)	1·808 (0·096)	1·541 (0·053)	−2·242 (0·700)	−2·138 (0·590)	−1·772 (0·415)
Personal goods	1·392 (0·064)	1·425 (0·068)	1·420 (0·062)	−2·854 (1·690)	−3·006 (1·860)	−2·984 (1·844)
Leisure goods	1·087 (0·045)	1·092 (0·047)	1·124 (0·063)	−1·846 (1·001)	−1·898 (1·066)	−2·207 (1·451)
Entertainment	4·610 (0·549)	6·742 (1·236)	2·807 (0·151)	−2·750 (0·467)	−3·715 (0·359)	−1·893 (0·714)
Leisure services	0·325 (0·089)	0·416 (0·075)	0·364 (0·087)	−1·201 (0·526)	−1·173 (0·451)	−1·188 (0·486)
Fares	2·270 (0·277)	2·515 (0·334)	2·185 (0·199)	−1·671 (1·277)	−1·796 (1·593)	−1·623 (1·249)
Motoring	0·814 (0·080)	0·789 (0·093)	0·831 (0·076)	−1·396 (1·230)	−1·452 (1·397)	−1·358 (1·120)
Petrol	0·866 (0·111)	0·831 (0·141)	0·843 (0·135)	−0·274 (0·117)	−0·085 (0·176)	−0·149 (0·147)
Tobacco	0·287 (0·042)	0·242 (0·046)	0·205 (0·056)	−0·480 (0·265)	−0·449 (0·280)	−0·419 (0·290)

In Table II we present the homogeneity and symmetry-constrained estimates of the quartile points of the distribution of uncompensated own price elasticities $\varepsilon^{\%}$ and income elasticities $\eta^{\%}$. Overall these appear to be reasonable.[7] Own-price elasticities are negative. Indeed, the eigenvalues of the Slutsky substitution matrix are negative or very small over the whole distribution of households. Table II shows some important variability in the distribution of elasticities. The food items are largely price inelastic and, with the distinct exception of food purchased outside the home, are all clear necessities. At the other extreme, entertainment is highly income and price elastic. Clothing items and household goods and services lie somewhere between.

For comparison, Table III presents the elasticity estimates for the LAIDS model. Here there are some noticeable differences with the QUAIDS estimates especially for the income elasticities η. In general, there is less dispersion of elasticities for the LAIDS model. As an example, consider the entertainments group for which the quadratic income term was clearly important.

7. SUMMARY AND CONCLUSIONS

There is an increasing requirement for the empirical analysis of consumer behaviour at the household level and at a finely disaggregated level of commodity groupings. To provide a realistic guide to consumer behaviour the relationship between expenditures, prices and total outlay is going to be non-linear. For empirical purposes, exact estimation of non-linear equation systems for large data sets with more than a small number of equations has typically been limited because of the intrinsic non-linearity of many models. However, we have noted that almost all popular non-linear models of consumer behaviour have the property of being conditionally linear. That is, they are linear in parameters conditional on complicated functions of prices, income and the preference parameters themselves.

In this paper we have proposed an Iterated Linear Least squares Estimator (ILLE) for such conditionally linear systems which fully identifies the parameters of the non-linear consumer demand models analysed. This ILLE is shown to be a computationally attractive consistent estimator. Its asymptotic distribution is derived and extended to the case of endogenous regressors.

The estimator is applied to the analysis of a 22 disaggregated demand system based on household data drawn from 20 years of British Family Expenditure Survey data. A Quadratic Almost Ideal Demand System (QUAIDS) is estimated. Homogeneity and symmetry restrictions are imposed and total expenditure is allowed to be endogenous. Estimation allowed the log total expenditure term, that enters as a quadratic in the QUAIDS model, to be endogenous. There is strong evidence in favour of including the quadratic terms. This is reflected in a comparison of the distribution of income and price elasticities between the QUAIDS model and the more standard LAIDS specification.

[7] Although we consider a fine disaggregation of commodities over a demographically homogeneous selection of households, there remains some possible misspecification since we do not account for the age of children (see Blundell, 1980) and we do allow for possible non-separabilities between goods and leisure (see Blundell and Walker, 1982; Browning and Meghir, 1991).

APPENDIX A: PROOFS OF THEOREMS

Proof to Theorem 1

We make the following regularity assumptions:

1. The parameter set $\mathcal{P}$ is a convex open set.
2. For all t, $g(x_t, \theta)$ is continuous with respect to θ.
3. $\partial g(x_t, \theta)/\partial\theta'$ exists and is a continuous function of θ.
4. The matrix

$$\frac{1}{T}\sum_{t=1}^{T} g(x_t, \theta)g(x_t, \theta)'$$

   almost certainly converges to $\mathrm{E}[g(x_t, \theta)g(x_t, \theta)']$ uniformly in θ.
5. The matrix $\mathrm{E}[g(x_t, \theta)g(x_t, \theta)']$ is non-singular.
6. The matrix

$$\frac{1}{T}\sum_{t=1}^{T} \frac{\partial}{\partial\theta'}\{g(x_t, \theta)[y_{it} - g(x_t, \theta)'\theta_i]\}$$

   almost certainly converges to

$$\mathrm{E}\left[\frac{\partial}{\partial\theta'}\{g(x_t, \theta)[y_{it} - g(x_t, \theta)'\theta_i]\}\right]$$

   uniformly in θ.
7. For all θ, the matrix whose ith row is

$$\mathrm{E}\left[\frac{\partial}{\partial\theta'}\{g(x_t, \theta)[y_{it} - g(x_t, \theta)'\theta_i]\}\right]$$

   is non-singular.

The estimator $\hat{\theta} = (\hat{\theta}'_1, \ldots, \hat{\theta}'_N)'$ is the solution to the system:

$$\hat{\theta}_i = [G(\hat{\theta})'G(\hat{\theta})]^{-1}G(\hat{\theta})'\mathbf{y}_i \qquad i = 1, \ldots, N$$

or, equivalently,

$$[G(\hat{\theta})'G(\hat{\theta})]^{-1}G(\hat{\theta})'[\mathbf{y}_i - G(\hat{\theta})\hat{\theta}_i] = 0 \qquad i = 1, \ldots, N$$

Since

$$G(\hat{\theta})'G(\hat{\theta}) = \frac{1}{T}\sum_{t=1}^{T} g(x_t, \hat{\theta})g(x_t, \hat{\theta})'$$

converge, by assumption, to a non-singular matrix uniformly in $\hat{\theta}$, then for large enough T, the estimator $\hat{\theta}$ is also such that

$$G(\hat{\theta})'[\mathbf{y}_i - G(\hat{\theta})\hat{\theta}_i] = 0 \qquad i = 1, \ldots, N$$

or, equivalently,

$$\frac{1}{T}\sum_{t=1}^{T} g(x_t, \hat{\theta})[y_{it} - g(x_t, \hat{\theta})'\hat{\theta}_i] = 0 \qquad i = 1, \ldots, N$$

Since all partial derivatives are continuous and since the parameter set is an open convex set, the theorem of intermediate values then implies that

$$\begin{aligned} 0 &= \frac{1}{T}\sum_{t=1}^{T} g(x_t, \hat{\theta})[y_{it} - g(x_t, \hat{\theta})'\hat{\theta}_i] \\ &= \frac{1}{T}\sum_{t=1}^{T} g(x_t, \theta^0)[y_{it} - g(x_t, \theta^0)'\theta_i^0] \\ &\quad + \frac{1}{T}\sum_{t=1}^{T} \frac{\partial}{\partial\theta'}\{g(x_t, \theta^*)[y_{it} - g(x_t, \theta^*)'\theta_i^*]\}(\hat{\theta} - \theta^0) \qquad i = 1, \ldots, N \end{aligned}$$

where θ^* is some intermediate value between $\hat{\theta}$ and θ^0. Clearly, the strong law of large numbers, together with

$$\mathrm{E}(u_t \mid x_t) = 0$$

implies that

$$\frac{1}{T}\sum_{t=1}^{T} g(x_t, \theta^0)[y_{it} - g(x_t, \theta^0)'\theta_i^0] \to 0 \quad \text{a.s.}$$

And since the matrix $B_T((x_t, x_t), \theta^*)$ whose ith row is

$$\frac{1}{T}\sum_{t=1}^{T} \frac{\partial}{\partial\theta'}\{g(x_t, \theta^*)[y_{it} - g(x_t, \theta^*)'\theta_i^*]\} \tag{A1}$$

converge a.s. to a non-singular matrix uniformly in θ^*, it thus follows that there is a value T^*, such that for all $T > T^*$, $B_T(x_t, \theta^*)$ is a.s. non-singular whatever the value of θ^*. Hence $\hat{\theta} = \theta^0$ converges to 0 for almost all sets $\{x_t, \mathbf{u}_t\}$. ■

Proof of Theorem 2

Under the conditions of Theorem 1 and the following additional assumption:

8. The following function of variables and parameters:

$$B_i((y_{it}, x_t), \theta) = \frac{\theta}{\partial\theta'}\{g(x_t, \theta)[y_{it} - g(x_t, \theta)'\theta_i]\} \tag{A2}$$

is continuous in θ uniformly in (y_{it}, x_t).

Consider again the asymptotic expansion used in Theorem 1:

$$\begin{aligned}0 &= \frac{1}{T}\sum_{t=1}^{T} g(x_t, \theta^0)u_{it} \\ &+ \frac{1}{T}\sum_{t=1}^{T} \frac{\partial}{\partial \theta'}\{g(x_t, \theta^*)[y_{it} - g(x_t, \theta^*)'\theta_i^*]\}(\hat{\theta} - \theta^0) \qquad i = 1, \ldots, N\end{aligned}$$

Let $B_i((y_{it}, x_t), \theta^*)$ be the matrix

$$\frac{\partial}{\partial \theta'}\{g(x_t, \theta^*)[y_{it} - g(x_t, \theta^*)'\theta_i^*]\}$$

Then

$$\begin{aligned}&\left|\frac{1}{T}\sum_{t=1}^{T} B_i((y_{it}, x_t), \theta^*) - \frac{1}{T}\sum_{t=1}^{T} B_i((y_{it}, x_t), \theta^0)\right| \\ &< \frac{1}{T}\sum_{t=1}^{T} |B_i((y_{it}, x_t), \theta^*) - B_i((y_{it}, x_t), \theta^0)| \\ &\leqslant \sup_{y_{it}, x_t} |B_i((y_{it}, x_t), \theta^*) - B_i((y_{it}, x_t), \theta^0)|\end{aligned}$$

When T tends to infinity, θ^* tends to θ^0 a.s., since θ^* is between $\hat{\theta}$ and θ^0. Hence, the uniform continuity of $B_i((y_{it}, x_t), \theta^*)$ with respect to θ^* implies that the r.h.s. of the inequality tends to zero.

Moreover,

$$\begin{aligned}&\frac{1}{T}\sum_{t=1}^{T} \frac{\partial}{\partial \theta'}\{g(x_t, \theta^0)[y_{it} - g(x_t, \theta^0)'\theta_i^0]\} \\ &= \frac{1}{T}\sum_{t=1}^{T} \frac{\partial g(x_t, \theta^0)}{\partial \theta'}[y_{it} - g(x_t, \theta^0)'\theta_i^0] \\ &\quad - \frac{1}{T}\sum_{t=1}^{T} g(x_t, \theta^0)\theta_i^{0\prime}\frac{\partial g(x_t, \theta^0)}{\partial \theta'} - \frac{1}{T}\sum_{t=1}^{T} g(x_t, \theta^0)(e_i' \otimes g(x_t, \theta^0)')\end{aligned}$$

where e_i is the ith column of the identity matrix I_N. The strong law of large numbers and the fact that $\mathrm{E}(u_t \mid x_t) = 0$ then implies that

$$\frac{1}{T}\sum_{t=1}^{T} \frac{\partial g(x_t, \theta^0)}{\partial \theta'}[y_{it} - g(x_t, \theta^0)'\theta_i^0] \to 0 \text{ a.s.}$$

and

$$\frac{1}{T}\sum_{t=1}^{T} g(x_t, \theta^0)\theta_i^{0'} \frac{\partial g(x_t, \theta^0)}{\partial \theta'} \to \mathrm{E}\left[g(x_t, \theta^0)\theta_i^{0'} \frac{\partial g(x_t, \theta^0)}{\partial \theta'}\right]$$

and

$$\frac{1}{T}\sum_{t=1}^{T} g(x_t, \theta^0)(e_i' \otimes g(x_t, \theta^0)') \to \mathrm{E}[g(x_t, \theta^0)(e_i' \otimes g(x_t, \theta^0)')]$$

Moreover, the central-limit theorem implies that

$$\begin{pmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^{T} g(x_t, \theta^0)u_{1t} \\ \vdots \\ \frac{1}{\sqrt{T}}\sum_{t=1}^{T} g(x_t, \theta^0)u_{Nt} \end{pmatrix} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} \mathbf{u}_t \otimes g(x_t, \theta^0)$$
$$\rightsquigarrow \mathcal{N}(0, \Sigma_0 \otimes \mathrm{E}[g(x_t, \theta^0)g(x_t, \theta^0)'])$$

Thus

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} g(x_t, \theta^0) \otimes \mathbf{u}_t \rightsquigarrow \mathcal{N}(0, \Sigma_0 \otimes K_0),$$

where

$$K_0 = \mathrm{E}[g(x_t, \theta^0)g(x_t, \theta^0)']$$

It has been thus shown that

$$\sqrt{T}(\hat{\theta} - \theta^0) \rightsquigarrow \mathcal{N}(0, J_0^{-1}(\Sigma_0 \otimes K_0)(J_0')^{-1})$$

where

$$\begin{aligned} J_0 &= \left\{ I_N \otimes \mathrm{E}[g(x_t, \theta^0)g(x_t, \theta^0)'] + \mathrm{E}\left[\Theta^{0'} \frac{\partial g(x_t, \theta^0)}{\partial \theta'} \otimes g(x_t, \theta^0)\right]\right\} \\ &= I_N \otimes K_0 + \mathrm{E}\left[\left(\Theta_0' \frac{\partial g(x_t, \theta^0)}{\partial \theta'}\right) \otimes g(x_t, \theta^0)\right] \end{aligned}$$

where Θ_0 is the matrix $(\theta_1^0, \ldots, \theta_N^0)$. ■

Proof of Corollary 1

Let us denote as θ the solution of the first-order conditions of problem (8):

$$\nabla(\tilde{\theta})'\left(\Sigma_0^{-1} \otimes \mathrm{P}_{G(\theta^0)}\right)(\mathbf{y} - [I_N \otimes G(\tilde{\theta})]\tilde{\theta}) = 0$$

A first-order asymptotic expansion yields

$$\sqrt{T}(\tilde{\theta} - \theta^0) \rightsquigarrow \mathcal{N}\left(0, \left[\nabla(\theta^0)'\left(\Sigma_0^{-1} \otimes \mathrm{P}_{G(\theta^0)}\right)\nabla(\theta^0)\right]^{-1}\right)$$

and $\tilde{\theta}$ and $\hat{\theta}$ have hence the same asymptotic distribution. ■

APPENDIX B: DATA AND ADDITIONAL RESULTS

Table BI. Data description — the 22 commodities

Commodity group	Definition	$\bar{w}_i$	SD (w)
Beer	Beer, on and off licence sales	0·0303	0·0409
Wine	Wine, on and off licence sales	0·0094	0·0202
Spirits	Spirits, on and off licence sales	0·0153	0·0310
Bread	Bread, flour, rice and cereals	0·0407	0·0215
Meat	All meat and fish	0·0774	0·0480
Dairy	All dairy products	0·0586	0·0311
Vegetables	Fresh, tinned and dried vegetables and fruit	0·0521	0·0260
Other food	Tea, coffee, drinks, sugar, jams and sweets	0·0591	0·0310
Food consumed outside the home	Restaurants and canteen meals	0·0645	0·0502
Electricity	Account and slot meter payments	0·0429	0·0299
Gas	Account and slot meter payments	0·0320	0·0301
Adult clothing	Adult clothing and footwear	0·0356	0·0615
Children's clothing and footwear	Children's clothing and footwear	0·0629	0·0637
Household services	Post, phone, domestic services and fees	0·0578	0·0585
Personal goods and services	Personal and chemist's goods and services	0·0482	0·0490
Leisure goods	Records, CDs, toys, books and gardening	0·0527	0·0514
Entertainment	Entertainment	0·0385	0·0676
Leisure services	TV licences and rentals	0·0177	0·0167
Fares	Rail, bus and other fares	0·0147	0·0324
Motoring	Maintenance, tax and insurance	0·0570	0·0714
Petrol	Petrol and oil	0·0678	0·0577
Tobacco	Cigarettes, pipe tobacco and cigars	0·0649	0·0498

Table BII. The estimated reduced form for ln x

Variable	Coeff. (s.e.)
Trend	−0·0145 (0·0216)
S_1	−0·1395 (0·0243)
S_2	−0·0866 (0·0209)
S_3	−0·0725 (0·0169)
Income	−0·5475 (0·0461)
Income^2	0·0988 (0·0048)
lnprices	Yes
R^2	0·725
F (pval)	466·3 (0·000)
T	4951

Table BIII. Estimated income effects, LAIDS

Commodity	ln x	Exog. $\|t\|$	Ov. id. χ^2_2
Beer	−0·0049 (0·003)	1·327	0·015
Wine	0·0124 (0·001)	3·473	6·987
Spirits	0·0137 (0·002)	0·207	7·909
Bread	−0·0231 (0·001)	0·790	4·955
Meat	−0·0280 (0·003)	2·571	1·017
Dairy	−0·0347 (0·002)	0·347	1·825
Vegetables	−0·0203 (0·002)	0·534	0·257
Other food	−0·0258 (0·002)	1·867	1·262
Food out	0·0248 (0·003)	3·717	7·592
Electricity	−0·0316 (0·002)	0·949	9·148
Gas	−0·0103 (0·002)	2·410	0·563
Adult clothing	0·0286 (0·004)	2·828	1·895
Child clothing	0·0150 (0·004)	4·501	5·564
Household services	0·0384 (0·004)	6·268	2·636
Personal goods	0·0209 (0·003)	1·067	4·173
Leisure goods	0·0062 (0·003)	0·424	1·931
Entertainment	0·0853 (0·004)	8·788	36·361
Leisure services	−0·0114 (0·001)	0·319	1·681
Fares	0·0166 (0·002)	4·878	3·386
Motoring	−0·0100 (0·004)	5·680	3·949
Petrol	−0·0099 (0·008)	0·702	0·386
Tobacco	−0·0520 (0·003)	5·265	0·970

Note: Standard errors in parentheses.

ACKNOWLEDGEMENTS

We have greatly benefited from helpful comments from James Banks, Ian Crawford, Eric Renault and two anonymous referees. This study is part of the research program of the ESRC Centre for the Microeconomic Analysis of Fiscal Policy at IFS. Material from the FES made available by the CSO through the ESRC data archive has been used by permission of the controller of HMSO. The usual disclaimer applies.

REFERENCES

Atkinson, A. B., J. Gomulka and N. H. Stern (1989), 'Spending on alcohol: evidence from the Family Expenditure Survey 1970–1983', *Economic Journal*, **100**, 808–827.

Attfield, C. L. F. (1997), 'Estimating a cointegrating demand system', *European Economic Review*, **41**, 61–73.

Banks, J., R. W. Blundell and A. Lewbel (1997), 'Quadratic Engel curves, welfare measurement and consumer demand', *Review of Economics and Statistics*, **LXXIX(4)**, 527–539.

Bierens, H. J. and H. A. Pott-Buter (1987), 'Specification of household Engel curves by nonparametric regression', *Econometric Reviews*, **9**, 123–184.

Blundell, R. W. (1980), 'Estimating continuous consumer equivalence scales in an expenditure model with labour supply', *European Economic Review*, **14**, 145–157.

Blundell, R. W., P. Pashardes and G. Weber (1993), 'What do we learn about consumer demand patterns from micro data?', *American Economic Review*, **83**, 570–597.

Blundell, R. W. and I. Walker (1982), 'Modeling the joint determination of household labour supplies and commodity demands', *Economic Journal*, **92**, 351–364.

Browning, M. and C. Meghir (1991), 'The effects of male and female labour supply on commodity demands', *Econometrica*, **59**, 925–951.

Deaton, A. S. and J. Muellbauer (1980a), 'An almost ideal demand system', *American Economic Review*, **70**, 312–326.

Deaton, A. S. and J. Muellbauer (1980b), *Economics and Consumer Behaviour*, Cambridge University Press, Cambridge.

Hausman, J. A. (1978), 'Specification tests in econometrics', *Econometrica*, **46**, 1251–1271.

Hausman, J. A., W. K. Newey and J. L. Powell (1995), 'Nonlinear errors in variables: estimation of some Engel curves', *Journal of Econometrics*, **65**, 205–234.

Holly, A. and D. Sargan (1982), 'Testing for exogeneity in a limited information framework', *Cahiers de Recherches Economiques*, No. 8204, University of Lausanne.

Jorgenson, D. W. and L. J. Lau (1975b), 'The structure of consumer preferences', *Annals of Social and Economic Measurement*, **4**, 49–101.

Jorgenson, D. W., L. J. Lau and T. M. Stoker (1980), 'Welfare comparison and exact aggregation', *American Economic Review*, **70**, 268–272.

Lewbel, A. (1991), 'The rank of demand systems: theory and nonparametric evidence', *Econometrica*, **59**, 711–730.

Lewbel, A. (1996), 'Aggregation without separability: a generalized composite commodity theorem', *American Economic Review*, **86**, 524–543.

Ng, S. (1995), 'Testing for homogeneity in demand systems when the regressors are non-stationary', *Journal of Applied Econometrics*, **10**, 147–163.

White, H. (1994), *Asymptotic Theory for Econometricians*, Academic Press, New York.

