Efficient Estimation of Semiparametric Equivalence Scales With Evidence From South Africa

Adonis Yatchew, Yiguo Sun and Catherine Deri *

Abstract

A useful method for semiparametric estimation and testing of base-independent equivalence scales is proposed by Pendakur (1998). We embed his specification in a partial linear index model with two resulting advantages. First, existing results on such models may be employed. Second, simultaneous estimation across multiple household types (and multiple goods) may be performed leading to greater efficiency. We also propose a parsimonious variant which incorporates the rule-of-thumb for equivalence scales given by $(A+\beta_2 K)^{\beta_1}$ where A and K are the number of adults and children, β_1 and β_2 are scalars. We employ nearest neighbour smoothers which are computationally fast and simple to implement and we propose a straightforward test of base-independence.

The efficiency gains from the proposed models and estimators are particularly helpful for developing country data where there is often much greater variation in household size and composition. Our analysis of South African household survey data generally supports the hypothesis of base-independent equivalence scales which we estimate to be approximately $(A+.74K)^{.59}$. Finally, we use the South African results to calibrate simulations and Monte Carlo experiments. These indicate very strong (often four-fold) increases in precision when estimating the parsimonious semiparametric specification.

Keywords: base-independent equivalence scales, semiparametric, Engel curves, partial linear index model

JEL Classification: C14, D11, D12, D63

October 16, 2000.

* Department of Economics, University of Toronto, yatchew@chass.utoronto.ca, yiguosun@chass.utoronto.ca, cderi@chass.utoronto.ca The authors are grateful to Gordon Anderson, Dwayne Benjamin, Lorne Brandt, Judith Giles, Joel Horowitz, Esfandiar Maasoumi, Angelo Melino and members of the Canadian Econometrics Study Group for helpful comments.

The estimation of share functions and shift parameters may be equivalently cast into a partial linear index function framework.[2] Suppose there are $q+1$ family types and select the first type as the reference to which the other q types will be compared. Let z be a q-dimensional row vector of dummy variables for the q (non-reference) family types, and ε an i.i.d random variable with mean 0 and variance σ^2. Then under the null hypothesis that there exists a base-independent equivalence scale, our model is given by:

$$y = f\left(\log x - z\delta\right) + z\eta + \varepsilon \tag{1.3}$$

where δ and η are q-dimensional parameter vectors.[3] For purposes of identification, the coefficient of $log\, x$ is set to 1. Within this framework it is straightforward to conduct analyses across multiple household types (and multiple goods). For example, one can test whether all $q+1$ family types can be imbedded in a common model of the form (1.3), whether families with and without children require separate specifications or, whether all family types are fundamentally different. Whereas Pendakur (1998) relies upon simulation methods to obtain standard errors and critical values, the partial linear index model framework — while not precluding simulation — permits one to draw upon a variety of procedures with simple asymptotic properties.

As Blackorby and Donaldson (1989,1993) and Pendakur point out, it is not possible to separately identify the equivalence scale and its elasticity if the share functions are loglinear. This is evident since under such circumstances, f in equation (1.3) is a linear function with say coefficients φ_o, φ_1 and we have $y = \varphi_0 + \varphi_1 \log x + z\left(\eta - \varphi_1\delta\right) + \varepsilon$. Thus one needs to conduct specification tests separately for each family type where the null is loglinear and the alternative is nonparametric.

To estimate the equivalence scale and elasticity as well as the nonparametric function f we adapt

---

[2] For relevant econometrics literature see Ichimura (1993), Powell (1994), Horowitz and Härdle (1996), Carroll, Fan, Gijbels and Wand (1997), Cavanagh and Sherman (1998) and references therein.

[3] Blundell, Duncan and Pendakur (1998) apparently propose such a framework but estimate using only pairs of household types.

1. INTRODUCTION

In a recent paper, Pendakur (1998) proposes a semiparametric estimator of base-independent equivalence scales. Suppose there are two types of households A and B. Let x be household expenditure and let p be a vector of prices corresponding to goods purchased. For the moment we will focus on purchases of a single good, say food. Let $y_A = f_A(p, x_A)$ be the food share of expenditure for households of type A. Under base-independence, the food expenditure share of type B households is related to that of type A households by:[1]

$$y_B = f_B(p, x_B) = f_A\left(p, \frac{x_B}{\Delta_B(p)}\right) + \eta_B(p) \tag{1.1}$$

where $\Delta_B(p)$ is the base-independent equivalence scale and η_B is its elasticity with respect to the price of food. Such scales -- if they are valid -- are extremely convenient because they tell us how much income is required by households of type B in order to achieve the same level of utility as households of type A, i.e., $x_B = \Delta_B(p) \cdot x_A$. Holding prices constant, expenditure shares as a function of the *log* of income are vertical and horizontal translations of each other:

$$y_B = f_B(x_B) = f_A\left(\frac{x_B}{\Delta_B}\right) + \eta_B = f(\log x_B - \delta_B) + \eta_B \tag{1.2}$$

where $\delta_B = \log \Delta_B$ and $f = f_A \circ exp$, a convolution of the type A share function with the exponential function. Pendakur tests this translation 'shape invariance' while estimating the function f nonparametrically. To do so he draws upon earlier results of Härdle and Marron (1990, 1993), Pinkse and Robinson (1995) and others. His analyses are restricted to pair-wise comparisons of household types. He also performs analyses across multiple goods to improve efficiency of estimation of the equivalence scale parameter δ_B.

[1] See Pendakur (1998), p.5, equation (2.5). His results draw upon earlier papers by Lewbel (1989) and Blackorby and Donaldson (1989, 1993).

a symmetric k-nearest neighbor smoother to the partial linear index model setting. The procedure is simple to implement and computationally fast so long as the number of distinct family types is small. However, the computational burden does increase with q. The reason is that estimation of index model parameters (in our case δ) typically requires a grid search.[4]

In Canada, about 75% of the population live in households which consist of singles or couples with 0, 1 or 2 children (a total of 6 family types). On the other hand, in developing countries one tends to observe much larger variation in household size and composition. Couples tend to have more children and it is not uncommon for extended families to occupy a single household. For example, for South Africa, which we study in some detail below, less than 40% of the population live in the types of families listed above. Restricting attention to these six family types would cause us to ignore the majority of the population. In order to build a parsimonious model which can handle a large number of family types while maintaining semiparametric flexibility, consider the following specification for the equivalence scale:

$$\Delta = \exp(\delta) = \left(A + \beta_2 K\right)^{\beta_1} \tag{1.4}$$

where A is the number of adults in the household and K is the number of children. Here β_1 reflects scale economies in the household and β_2 measures the effect on the equivalence scale of children relative to adults. Both parameters are restricted to be between 0 and 1. (See e.g., Citro and Michael (1995, p. 176) who recommend values around .7 for β_1 and β_2.) Then our model becomes:

$$y = f\left(\log x - \beta_1 \log\left(A + \beta_2 K\right)\right) + z\eta + \varepsilon \tag{1.5}$$

While estimation of (1.3) requires a q-dimensional grid search to estimate δ, (1.5) reduces this search to two dimensions. (No similar problem exists for estimation of η.) Furthermore, because

---

[4] See Horowitz and Härdle (1996) for an alternative procedure.

the equivalence scales are functions of only two parameters, they will, in general, be estimated more precisely.[5] Finally, specification (1.5) has the appealing property that it is monotone in A and K and will yield plausible 'spacings' of equivalence scales among families of similar composition — an essential feature from an equity point of view if model estimates are to be used for policy purposes.

After delineating estimators for models (1.3) and (1.5) we produce a simple test of base independence (which is essentially a test of whether the share functions are 'similar in shape') and a test of the parsimonious specification (1.5) against the more general (1.3). The tests compare restricted and unrestricted residual variance estimates.

If reliable data are available on the consumption of various goods, system estimation may be applied since the equivalence scale parameter is common across share equations for different goods. This can be implemented most easily by applying GLS methods to the single equation estimates but we also use system wide estimation procedures. A simple χ^2 test may then be applied to test equality of equivalence scale parameters across goods.

The paper is organized as follows. Section 2 describes single and multi-equation estimators and test procedures. Section 3 applies these techniques to South African data on food and rent expenditures. The estimates in Section 3 are then used to perform efficiency comparisons of various specifications and to calibrate the Monte Carlo experiments in Section 4. Our principal focus is on the equivalence scale parameter δ, since it is usually of greatest interest. We find up to 25% improvement in efficiency when estimating the multi-family model (1.3) relative to pair-wise estimation. On the other hand, the parsimonious model (1.5) yields as much as a *four-fold* increase in precision. Multi-equation estimation (using food and rent data) yields modest gains of less than 5%. An Appendix contains technical details.

---

[5] One might also expect families of similar composition to have similar η's. We did consider specifications of the form $y = f(\log x - \beta_1 \log(A + \beta_2 K)) + \gamma_1 A + \gamma_2 K + \varepsilon$ and found significant deterioration in fit. Obviously more general and nonlinear functions for η as a function of A and K may be specified.

2. STATISTICAL PROCEDURES

Single Equation Estimation

Suppose we are given data $(y_1, x_1, z_1), \ldots, (y_n, x_n, z_n)$ on n households. With mild abuse of notation, y and x will be used to denote both the variable in question and the corresponding column vector of observations on the variable. The context should make it clear which applies. If x is a vector, then $f(x)$ will denote the vector consisting of f evaluated at the components of x. We will use Z to denote the nxq matrix of data on the dummy variables. (Hence, $\log x - Z\delta$ is a vector as is $f(\log x - Z\delta)$.) Equation (1.3) may be written in vector-matrix notation as:

$$\underset{nx1}{y} = f\Big(\underset{nx1}{\log x} - \underset{nxq}{Z}\,\underset{qx1}{\delta}\Big) + \underset{nxq}{Z}\,\underset{qx1}{\eta} + \underset{nx1}{\varepsilon} \tag{2.1}$$

To estimate this model we will rely on the symmetric k-nearest neighbor smoother (also known as the running mean smoother) which averages an equal number of observations on either side of a point to estimate the function at that point. (Assume throughout that k is even.) The estimator is fast and simple to implement. Thus, let S be the smoothing matrix given by:

$$\underset{(n-k)xn}{S} = \begin{bmatrix} \frac{1}{k},\ldots,\frac{1}{k},0,\frac{1}{k},\ldots,\frac{1}{k},0,\ldots\ldots\ldots\ldots,0 \\ 0,\frac{1}{k},\ldots,\frac{1}{k},0,\frac{1}{k},\ldots,\frac{1}{k},0,\ldots\ldots\ldots,0 \\ \vdots \quad \vdots \\ \vdots \quad \vdots \\ 0,\ldots\ldots\ldots,0,\frac{1}{k},\ldots,\frac{1}{k},0,\frac{1}{k},\ldots,\frac{1}{k},0 \\ 0,\ldots\ldots\ldots\ldots,0,\frac{1}{k},\ldots,\frac{1}{k},0,\frac{1}{k},\ldots,\frac{1}{k} \end{bmatrix} \tag{2.2}$$

Analogously to Härdle, Hall and Ichimura (1993), the same asymptotically optimal smoothing parameter k may be used to estimate δ, η and f. Since cross-validation is used to select k while simultaneously estimating the parameters of the model, S excludes the i-th observation when estimating the i-th nonparametric effect.

To estimate (1.3) or (2.1) proceed as follows. Fix k. For any δ let P_δ be the permutation matrix which reorders the vector $\log x - Z\delta$ so that its elements are in increasing order. Let I be the identity matrix and apply $(I-S)P_\delta$ to (2.1) to obtain:

$$(I-S)P_\delta y = (I-S)P_\delta f(\log x - Z\delta) + (I-S)P_\delta Z\eta + (I-S)P_\delta \varepsilon \tag{2.3}$$

which yields the usual estimator of the partial linear model (see Robinson (1988) and Speckman (1988)), that is:

$$\hat{\eta}_\delta = \left[((I-S)P_\delta Z)'(I-S)P_\delta Z\right]^{-1}((I-S)P_\delta Z)'(I-S)P_\delta y \tag{2.4}$$

By a grid search over values of δ and k find:

$$s^2 = \min_k \min_\delta \frac{1}{n}\left((I-S)P_\delta y-(I-S)P_\delta Z\hat{\eta}_\delta\right)'\left((I-S)P_\delta y-(I-S)P_\delta Z\hat{\eta}_\delta\right) \tag{2.5}$$

$\hat{k}$, $\hat{\delta}$ will be the values which achieve this minimum and our estimator of η will be $\hat{\eta} = \hat{\eta}_{\hat{\delta}}$. Large sample standard errors and related asymptotic theory as well as estimation of the parsimonious specification (1.5) are detailed in the Appendix.

Engel's method for determining equivalence scales may be easily applied in this setting. Let y be the share of income spent on food and set $\eta = 0$ in equation (2.1) to obtain $y = f(\log x - Z\delta) + \varepsilon$. For Rothbarth's method let y be income spent on adult goods. (The usual simplification is to use expenditures on all non-food goods -- see Deaton and Meullbauer (1986) and Deaton (1997, p.247-70)). Then set $\eta = 0$ in equation (2.1).

We will want to test base-independence as well as a number of other hypotheses such as the validity of the parsimonious specification (1.5) and whether households consisting solely of adults, single parent households and two-parent households can be embedded in a single model. In each case our test statistic will be based on a comparison of a restricted and unrestricted estimated variance, i.e., $s^2_{res} - s^2_{unr}$. For simplicity, we will typically use an 'optimal differencing estimator' [6] to estimate the *unrestricted* variance. Suppose one is given data $(y_1, x_1), \ldots, (y_n, x_n)$ on a smooth generic nonparametric regression model $y = f(x) + \varepsilon$ where y and x are scalars and the data have been reordered so that $x_1 \leq \ldots \leq x_n$. Let m be the order of differencing and $d_0, d_1, \ldots, d_m$ the optimal differencing weights.[7] We will use D to represent a generic square differencing matrix:

$$D = \begin{bmatrix} d_0, d_1, d_2, ..d_m, 0, \ldots\ldots\ldots, 0 \\ 0,\ d_0, d_1, d_2, ..d_m, 0, \ldots\ldots\ldots, 0 \\ \vdots \quad \vdots \\ \vdots \quad \vdots \\ 0, \ldots\ldots\ldots, 0,\ d_0, d_1, d_2, ..d_m, 0 \\ 0, \ldots\ldots\ldots, 0, d_0, d_1, d_2, ..d_m \\ 0, \ldots\ldots\ldots\ldots\ldots\ldots, 0 \\ \vdots \quad \vdots \\ 0, \ldots\ldots\ldots\ldots\ldots\ldots, 0 \end{bmatrix} \tag{2.6}$$

Note that the last m rows are zero. The differencing estimator of the residual variance is defined as:

$$s^2_{diff} = \frac{1}{n} y'D'Dy \tag{2.7}$$

We will use (2.7) to obtain our *unrestricted* residual variance estimators and in each case we will have

---

[6] See Hall, Kay and Titterington (1990), Yatchew (1997,1999a,b).

[7] The weights satisfy the conditions $\Sigma^m_{j=0} d_j = 0$, $\Sigma^m_{j=0} d_j^2 = 1$ and are optimal in the sense that they minimize the large sample variance of s^2_{diff} which will be defined momentarily. See Yatchew (1998, p. 697) for weights up to order 10. Higher order weights are available upon request.

under the null hypothesis:

$$(mn)^{1/2}\frac{\left(s_{res}^2 - s_{unr}^2\right)}{s_{unr}^2} \quad \overset{D}{\rightarrow} \quad N(0,1) \tag{2.8}$$

Consider testing the base-independent specification (1.3) against the alternative that Engel curves for the various family types are not similar in shape. That is, under the alternative we have $q+1$ distinct models:

$$y_j \;=\; f_j(\log x_j) + \varepsilon_j \qquad j = 0,1,\ldots q \tag{2.9}$$

where y_j , $\log x_j$ and ε_j are column vectors of length n_j for the j-th family type. In this case we may use the differencing estimator (2.7) to estimate $s_{diff,j}^2$ the residual variance for each family type j. We then construct s_{unr}^2 as their weighted combination where the weights are the relative sizes of the sub-populations. That is,

$$s_{unr}^2 \;=\; \sum_{j=0}^{q}\frac{n_j}{n}s_{diff,j}^2 \;=\; \frac{1}{n}\sum_{j=0}^{q}y_j'D'Dy_j \tag{2.10}$$

To complete the test, the restricted estimator s_{res}^2 is obtained directly from (2.5) and the test in (2.8) may be applied.[8]

Next, consider testing the base-independent specification (1.3) against the parsimonious version (1.5). In this case, the restricted estimator s_{res}^2 is obtained using (A.4) in the appendix. Obtain $\hat{\delta}$, $\hat{\eta}$ by solving (2.5) and construct the set of ordered pairs: $\left(y_i - z_i\hat{\eta},\ \log x_i - z_i\hat{\delta}\right)$ $i = 1,\ldots,n$ where the $\log x_i - z_i\hat{\delta}$ are in increasing order. Define the unrestricted variance s_{unr}^2 to be the differencing

[8] To test the parsimonious specification (1.5) against the unrestricted model (2.9) calculate s_{res}^2 using (A.4) in the Appendix.

estimator (2.7) applied to these ordered pairs. Finally, calculate the test statistic (2.8).[9]

Finally, to test whether adult households (A), single parent families (S) and couples with children (C) can be embedded in a single model, proceed as follows. Apply (2.5) to the restricted model (1.3) to obtain s^2_{res}. Re-estimate model (1.3) for each family type A, S and C to obtain $\hat{\delta}_A, \hat{\eta}_A$, $\hat{\delta}_S, \hat{\eta}_S$ and $\hat{\delta}_C, \hat{\eta}_C$. Construct the ordered pairs $\left(y_{Ai} - z_{Ai}\hat{\eta}_A,\ \log x_{Ai} - z_{Ai}\hat{\delta}_A\right)$ $i = 1,\ldots,n_A$, $\left(y_{Si} - z_{Si}\hat{\eta}_S,\ \log x_{Si} - z_{Si}\hat{\delta}_S\right)$ $i = 1,\ldots,n_S$ and $\left(y_{Ci} - z_{Ci}\hat{\eta}_C,\ \log x_{Ci} - z_{Ci}\hat{\delta}_C\right)$ $i = 1,\ldots,n_C$ and apply the differencing estimator (2.7) to obtain $s^2_{diff_A}, s^2_{diff_S}$ and $s^2_{diff_C}$. Analogously to (2.10) calculate

$$s^2_{unr} = \frac{n_A}{n} s^2_{diff_A} + \frac{n_S}{n} s^2_{diff_S} + \frac{n_C}{n} s^2_{diff_C} \tag{2.11}$$

and apply (2.8).

In selecting the order of differencing m, the objective is to under-smooth estimation of the alternative relative to the null. This ensures that test statistic (2.8) admits the simple standard normal approximation under the null. In our tests below, we use $m \le 10$. For further details, see the Appendix.

[9] Alternatively, one could add dummy variables to specification (1.5) and test their joint significance.

Multi-Equation Procedures

Consider now the two-good model:

$$
\begin{aligned}
\underset{nx1}{y_1} &= f_1\left(\underset{nx1}{\log x} - \underset{nxq}{Z}\,\underset{qx1}{\delta}\right) + \underset{nxq}{Z}\,\underset{qx1}{\eta_1} + \underset{nx1}{\varepsilon_1} \\
\underset{nx1}{y_2} &= f_2\left(\underset{nx1}{\log x} - \underset{nxq}{Z}\,\underset{qx1}{\delta}\right) + \underset{nxq}{Z}\,\underset{qx1}{\eta_2} + \underset{nx1}{\varepsilon_2}
\end{aligned}
\tag{2.12}
$$

where the covariance matrix for the stacked vector consisting of ε_1 and ε_2 is given by $\underset{(2x2)}{\Sigma} \otimes I_n$ and Σ had distinct elements $\sigma_1^2, \sigma_2^2, \sigma_{12}$. As before, the Z matrix consists of dummy variables to distinguish family types; δ which is common across equations is the vector of *log* equivalence scale parameters and η_1 and η_2 are the price elasticities of the equivalence scale. The parameters of the model may be estimated more efficiently by applying GLS procedures to the single equation estimates and imposing the constraint that δ is identical in both equations. Details are provided in the Appendix.

Alternatively, an asymptotically efficient system estimation procedure is given by the following. Apply the single equation procedure in (2.5) to each equation. Retain the optimized values of the smoothing parameters $\hat{k}_1, \hat{k}_2$ and define corresponding smoothing matrices S_1, S_2 (see equation (2.2)). For any δ let P_δ be the permutation matrix which reorders the vector $\log x - Z\delta$ so that its elements are in increasing order. Next define:

$$
\begin{aligned}
\hat{\eta}_{1_\delta} &= \left[((I-S_1)P_\delta Z)'(I-S_1)P_\delta Z\right]^{-1}((I-S_1)P_\delta Z)'(I-S_1)P_\delta y_1 \\
\hat{\eta}_{2_\delta} &= \left[((I-S_2)P_\delta Z)'(I-S_2)P_\delta Z\right]^{-1}((I-S_2)P_\delta Z)'(I-S_2)P_\delta y_2
\end{aligned}
\tag{2.13}
$$

and let $\hat{\Sigma}$ be a consistent estimator of Σ obtained from single equation estimation.[10] By a grid

[10] Estimates of σ_1^2 and σ_2^2 may be obtained directly from (2.5). To estimate σ_{12}, calculate the sample covariance of the residuals from the single equation estimates.

search over values of δ, find

$$\min_{\delta} \frac{1}{n} \begin{pmatrix} (I-S_1)P_\delta y_1 - (I-S_1)P_\delta Z\hat{\eta}_{1_\delta} \\ (I-S_2)P_\delta y_2 - (I-S_2)P_\delta Z\hat{\eta}_{2_\delta} \end{pmatrix}' \left(\hat{\Sigma}^{-1} \otimes I_n\right) \begin{pmatrix} (I-S_1)P_\delta y_1 - (I-S_1)P_\delta Z\hat{\eta}_{1_\delta} \\ (I-S_2)P_\delta y_2 - (I-S_2)P_\delta Z\hat{\eta}_{2_\delta} \end{pmatrix} \tag{2.14}$$

$\hat{\delta}$ will be the value which achieves the minimum and our estimators of the price elasticities will be $\hat{\eta}_1 = \hat{\eta}_{1_{\hat{\delta}}}$, $\hat{\eta}_2 = \hat{\eta}_{2_{\hat{\delta}}}$. Of course, one could iterate by updating the estimate of Σ; and, one could also search over values of the smoothing parameters (as in (2.5)) but neither will effect asymptotic efficiency. A similar procedure is available if one is estimating the two-equation version of the parsimonious model (1.5). Asymptotic standard errors and related results are given in the Appendix.

3. Analysis of South African Data

We now apply the techniques from the previous sections to data taken from the 1993 South African Living Standards Survey. This World Bank survey, funded by the governments of Denmark, the Netherlands and Norway was undertaken just prior to the country's first democratic elections in April 1994. Covering approximately 9000 households, the principal purpose of the survey was to collect data for policy makers on living standards. In principle the sample design which was employed should have yielded a representative sample. However, a number of factors made perfect implementation impossible, in particular, violence and systematic under-representation of whites in the sample (whites were most likely to refuse the interview).[11]

Initially there were 8794 valid observations. Trimming 5% from each tail of the income distribution reduces the data set to 7914. We focus attention on families with no more than 6 adults and no more than 5 children yielding 7358 observations. Table I summarizes the distribution of family types. Summary statistics for key variables are contained in Appendix Table I.

Table I: Distribution of Family Composition
(Number of Families)

Children / Adults	0	1	2	3	4	5	
1	1109	138	126	85	61	14	1533
2	890	526	524	309	144	65	2458
3	373	314	322	233	138	67	1447
4	222	227	230	160	104	66	1009
5	105	117	144	116	66	43	591
6	50	44	71	78	45	32	320
	2749	1366	1417	981	558	287	7358

Figure 1-A portrays Engel curves for food for three family sub-groupings: families consisting of up to four adults, single parent and two parent families with up to 3 children. As expected these slope

---

[11] Additional information on the methodology, sampling and data from this survey can be obtained at http://www.worldbank.org/lsms/country/za94/docs/za94ovr.txt.

Figure 1A: Engel Curves for Food Share

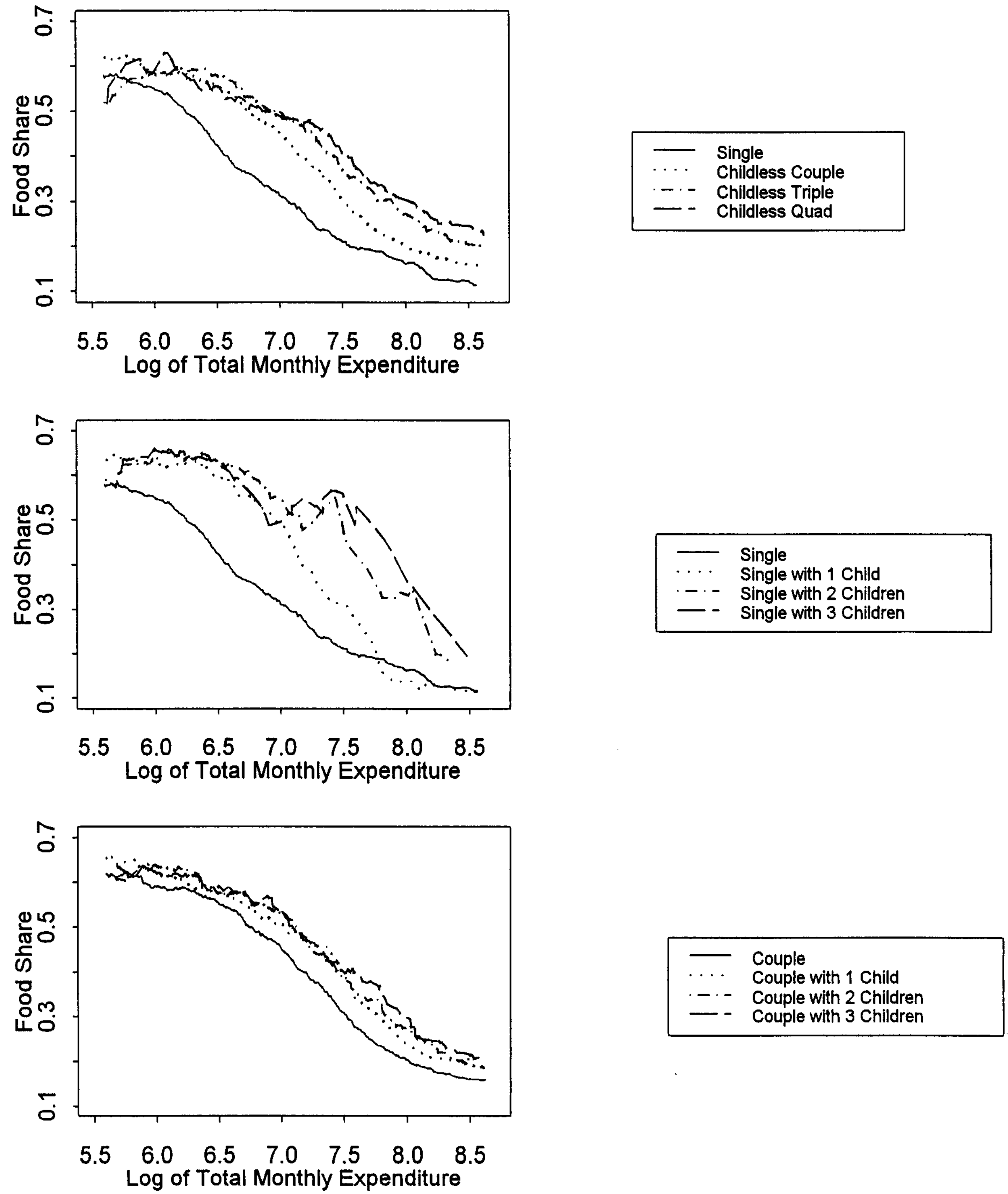
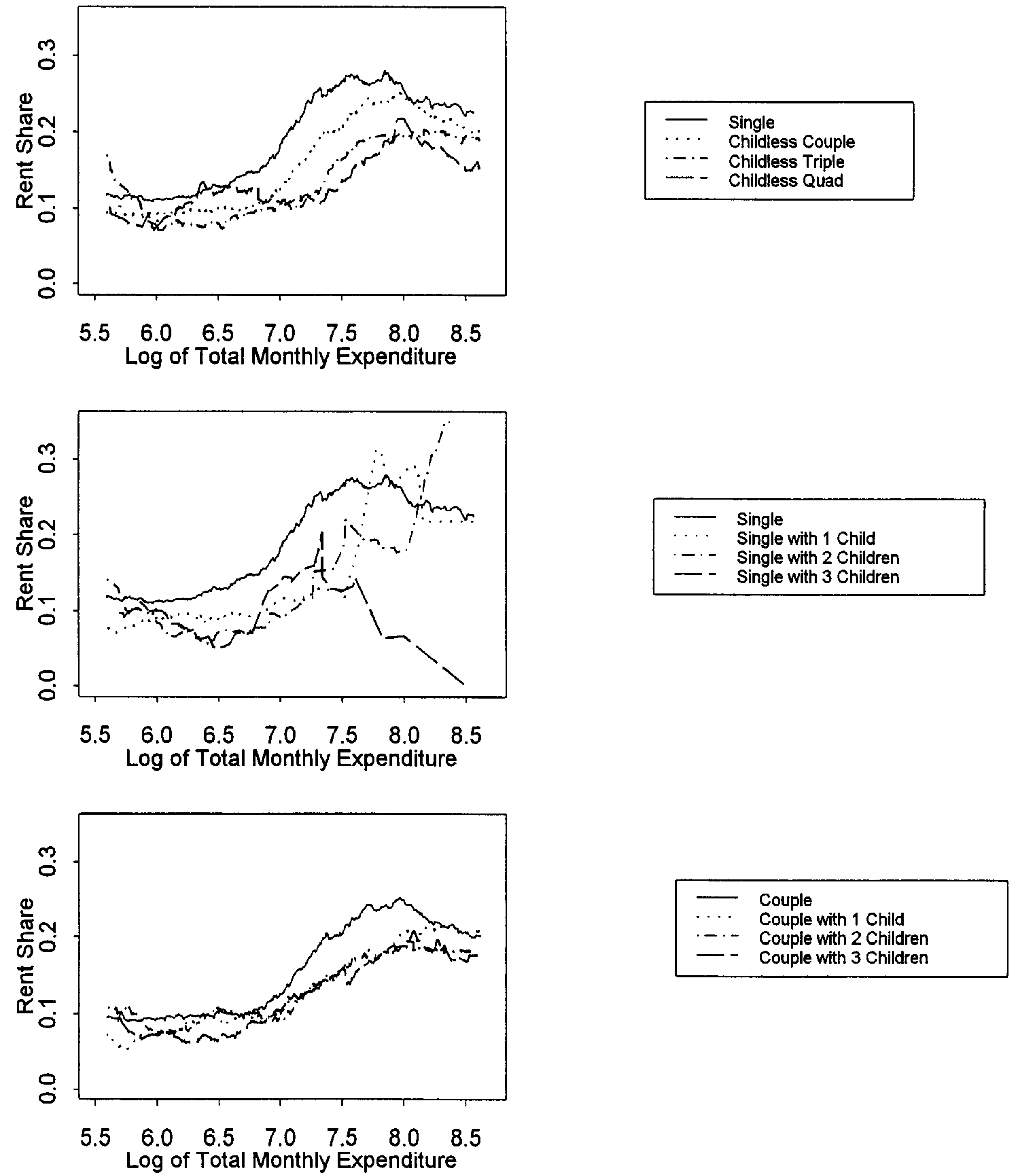


Figure 1B: Engel Curves for Rent Share



downwards and for a given level of expenditure, food share increases with family size. Figure 1-B portrays Engel curves for rent. The rent share is flat at low levels of income, then increases before flattening out yet again. For a given level of expenditure, rent share generally decreases as family size increases.

We begin by using the differencing test of specification against a logarithmic null (Yatchew(1997)). We run separate tests for each of the 36 family types in Table I. Because the null is parametric, under-smoothing the alternative is not necessary. Nevertheless, even for low orders of differencing, the *log* specification for the food and rent Engel curves is rejected for most family types. (At $m=10$, rejection occurs for 30 of 36 family types for the food data and 20 of 36 family types for the rent data.)

Table II-A summarizes estimation results for Pendakur's pair-wise comparison model using the food data. Individually, most estimates of the equivalence scale Δ or the *log* equivalence scale δ are plausible, though some do not satisfy monotonicity. For example, the estimated equivalence scale for 3 adults is 1.95; for 4 adults it is 1.92. The associated standard errors are large, particularly when considered from a policy point of view. Furthermore, the 'spacings' between estimated equivalence scales are not always plausible and as policy parameters, would be difficult to defend from an equity standpoint. The explanatory power is good, R^2 being around the 50% mark. The optimal k generally increases with sample size. Table II-B summarizes analogous results for the rent data. The explanatory power is much lower with the typical R^2 being around 15%.

Table III-A summarizes estimation results for the multi-family model equation (1.3) applied to four families at a time and using food data. Standard errors are generally smaller than the pair-wise estimates in Table II-A. In some cases improvements are quite substantial — the estimated standard error $\hat{\delta}$ for families consisting of 4 adults and no children improves from .29 to .18. Nevertheless, as a practical matter, the 'spacings' are again often implausible and standard errors remain large. For example, the estimated Δ for a couple with one child relative to a childless couple is 1.03 (s.e. .11); with two children, it is 1.52 and with three 1.57. Table III-B summarizes analogous results for the rent data where the estimates are substantially less plausible.

Table IV-A presents results for the parsimonious model (1.5) for food data. The (common) Engel curve is nonparametric but the equivalence scales satisfy a parametric form which provides for plausible spacings and monotonicity in A and K. Standard errors are much smaller -- sometimes by a factor of four -- than those corresponding to pair-wise or even multi-family estimates (Tables II-A, III-A). Furthermore the reduction in explanatory power is modest relative to the completely unconstrained model (2.9) which allows *separate* Engel curves for each family type — R^2 declines from .519 to .514. Applying test (2.8) where s_{res}^2 is the estimate of the residual variance from the parsimonious model, we obtain a value of 2.6 for the (standard normal) test statistic. While -- strictly speaking -- this constitutes a rejection, it is not a strong rejection given the sample size and the severity of the constraints. Nor do the comparisons of explanatory power suggest that a grave mis-specification has been committed. In contrast, the log-linear model which requires a common slope for all Engel curves but allows separate intercepts for each family type yields a somewhat lower R^2 of .490. Table IV-B contains analogous results for the rent data. Again the standard errors are dramatically lower. The R^2 for the parsimonious model is very slightly lower at .142 than for the completely unconstrained model (2.9) where it is .143. Applying (2.8) yields a value of .122 for the test statistic.

Overall, given these results, we consider the parsimonious model to be a good compromise: it has desirable monotonicity and spacing properties; it appears to result in much more precise estimates than pair-wise or multi-family estimation; and, with 36 family types, it is much easier to compute than the base-independent semiparametric model (1.3) which imposes no constraints on equivalence scales across family types.

Joint estimation of the two equation parsimonious model yields only very slightly improved estimates when compared to the Food Equation estimates in Table IV-A. We obtain $\hat{\beta}_1$ = .60 with a standard error of .0538; $\hat{\beta}_2$ = .74 with a standard error of .1788. GLS estimates (see Appendix for methodology) were very similar to the joint estimates.

TABLE II-A: PAIR-WISE ESTIMATES OF EQUIVALENCE SCALES -- FOOD DATA

Family Composition: Adults	Family Composition: Children	Equivalence Scale $\hat{\Delta} = \exp(\hat{\delta})$	$\hat{\delta}$	$\hat{\eta}$	$\hat{k}$	n	s^2	R^2
Reference Family: Single Adult								
1	1	1.23 (.35)	.21 (.28)	.08 (.05)	20	1247	.0197	.51
1	2	1.67 (.53)	.51 (.32)	.09 (.05)	40	1235	.0199	.51
2	0	1.65 (.24)	.50 (.14)	.004 (.02)	118	1999	.0187	.54
2	1	1.79 (.26)	.58 (.15)	.04 (.02)	72	1635	.0190	.52
2	2	2.05 (.31)	.72 (.15)	.02 (.03)	60	1633	.0185	.53
2	3	2.32 (.53)	.84 (.23)	.01 (.04)	66	1418	.0192	.50
3	0	1.95 (.46)	.67 (.24)	.02 (.03)	50	1482	.0185	.50
3	1	1.99 (.41)	.69 (.21)	.03 (.03)	40	1423	.0190	.49
3	2	2.44 (.71)	.89 (.29)	.006 (.05)	46	1431	.0197	.50
4	0	1.92 (.56)	.65 (.29)	.04 (.04)	42	1331	.0195	.47
4	1	2.56 (.57)	.94 (.22)	-.004 (.03)	50	1336	.0195	.47
4	2	3.22 (.84)	1.17 (.26)	-.03 (.04)	46	1339	.0201	.46
Reference Family: Single Parent With One Child								
1	2	1.22 (.56)	.20 (.46)	.02 (.03)	50	264	.0174	.50
Reference Family: Couple With One Child								
2	2	1.05 (.13)	.05 (.12)	.004 (.02)	70	1050	.0159	.59
2	3	1.52 (.24)	.42 (.16)	-.05 (.02)	74	835	.0169	.55

Notes: standard errors in parentheses.

TABLE II-B: PAIR-WISE ESTIMATES OF EQUIVALENCE SCALES -- RENT DATA								
Family Composition		Equivalence Scale $\hat{\Delta} = \exp(\hat{\delta})$	$\hat{\delta}$	$\hat{\eta}$	$\hat{k}$	n	s^2	R^2
Adults	Children							
Reference Family: Single Adult								
1	1	1.19 (.39)	.17 (.33)	-.03 (.02)	80	1247	.0159	.15
1	2	1.20 (.42)	.35 (.18)	-.05 (.02)	70	1235	.0156	.17
2	0	1.40 (.18)	.34 (.13)	-.02 (.01)	92	1999	.0157	.16
2	1	1.86 (.32)	.62 (.17)	-.03 (.01)	84	1635	.0148	.15
2	2	1.88 (.34)	.63 (.18)	-.03 (.01)	88	1633	.0145	.14
2	3	1.95 (.46)	.67 (.23)	-.03 (.01)	84	1418	.0154	.14
3	0	1.36 (.28)	.31 (.20)	-.05 (.01)	88	1482	.0150	.16
3	1	1.65 (.35)	.50 (.21)	-.04 (.01)	88	1423	.0159	.14
3	2	2.39 (.53)	.87 (.22)	-.02 (.01)	68	1431	.0155	.14
4	0	1.51 (.41)	.41 (.27)	-.04 (.02)	70	1331	.0162	.13
4	1	1.88 (.42)	.63 (.22)	-.03 (.02)	70	1336	.0158	.13
4	2	2.27 (.61)	.82 (.27)	-.01 (.02)	70	1339	.0164	.12
Reference Family: Single Parent With One Child								
1	2	1.60 (3.43)	.47 (2.14)	.02 (.02)	56	264	.0088	.33
Reference Family: Couple With One Child								
2	2	1.34 (.35)	.29 (.26)	.01 (.01)	80	1050	.0104	.16
2	3	1.62 (.71)	.48 (.44)	.02 (.02)	80	835	.0110	.15
Notes: standard errors in parentheses.								

TABLE III A : MULTI-FAMILY ESTIMATES OF EQUIVALENCE SCALES - FOOD								
Family Composition		Equivalence Scale $\hat{\Delta} = \exp(\hat{\delta})$	$\hat{\delta}$	$\hat{\eta}$	$\hat{k}$	n	s^2	R^2
Adults	Children							
Reference Family: Single Adult								
2	0	1.75 (.20)	.56 (.12)	-.01 (.02)	70	2594	.0179	.55
3	0	2.41 (.37)	.88 (.15)	-.02 (.03)				
4	0	3.19 (.58)	1.16 (.18)	-.05 (.03)				
Reference Family: Single Adult								
1	1	1.30 (.32)	.26 (.24)	.07 (.04)	58	1458	.0196	.53
1	2	1.82 (.48)	.60 (.26)	.07 (.04)				
1	3	1.97 (.60)	.68 (.30)	.07 (.04)				
Reference Family: Couple With No Children								
2	1	1.03 (.11)	.03 (.11)	.04 (.02)	50	2249	.0169	.60
2	2	1.52 (.17)	.42 (.11)	-.01 (.02)				
2	3	1.57 (.21)	.45 (.13)	-.01 (.02)				
Reference Family: Couple With One Child								
2	2	1.20 (.16)	.18 (.13)	-.02 (.02)	44	1503	.0168	.56
2	3	1.26 (.19)	.23 (.15)	-.02 (.02)				
2	4	1.31 (.26)	.27 (.20)	-.0003 (.03)				
Notes: standard errors in parentheses.								

Family Composition		Equivalence Scale $\hat{\Delta} = \exp(\hat{\delta})$	$\hat{\delta}$	$\hat{\eta}$	$\hat{k}$	n	s^2	R^2
Adults	Children							
Reference Family: Single Adult								
2	0	1.94 (.30)	.66 (.15)	.002 (.01)	60	2594	.0150	.16
3	0	2.10 (.42)	.74 (.20)	-.02 (.01)				
4	0	2.51 (.62)	.92 (.25)	-.01 (.02)				
Reference Family: Single Adult								
1	1	1.25 (.39)	.22 (.31)	-03 (.02)	70	1458	.0151	.17
1	2	1.88 (.60)	.63 (.32)	-04 (.02)				
1	3	1.82 (.78)	.60 (.43)	-04 (.02)				
Reference Family: Couple With No Children								
2	1	1.26 (.21)	.23 (.17)	-.01 (.01)	40	2249	.0123	.18
2	2	1.25 (.19)	.22 (.16)	-.02 (.01)				
2	3	1.45 (.30)	.37 (.21)	-.02 (.01)				
Reference Family: Couple With One Child								
2	2	1.27 (.35)	.24 (.28)	.01 (.01)	100	1503	.0103	.16
2	3	1.17 (.38)	.16 (.32)	-.004 (.02)				
2	4	1.38 (.62)	.32 (.45)	-.001 (.02)				

Notes: standard errors in parentheses.

TABLE IV-A: PARSIMONIOUS MODEL ESTIMATES -- FOOD DATA							
$\hat{\beta}_1$ .59 (.054)	$\hat{\beta}_2$ .74 (.184)	$corr(\hat{\beta}_1, \hat{\beta}_2)$ -.458		$\hat{k}$ 188	n 7358	s^2 .01822	R^2 .514
Adults	Children	Equivalence Scale $\hat{\Delta} = \exp(\hat{\delta})$	se $\hat{\Delta}$	Log Equivalence Scale $\hat{\delta}$	se $\hat{\delta}$	$\hat{\eta}$	se $\hat{\eta}$
1	0	1.00	0.00	0.00	0.000	0.000	0.000
1	1	1.39	0.08	0.33	0.055	0.021	0.009
1	2	1.71	0.13	0.54	0.078	0.022	0.013
1	3	1.99	0.18	0.69	0.091	0.009	0.016
1	4	2.25	0.23	0.81	0.100	-0.023	0.020
1	5	2.49	0.27	0.91	0.107	-0.024	0.025
2	0	1.51	0.06	0.41	0.038	0.062	0.015
2	1	1.81	0.09	0.59	0.051	0.036	0.011
2	2	2.09	0.14	0.74	0.068	0.016	0.014
2	3	2.34	0.19	0.85	0.081	-0.004	0.016
2	4	2.57	0.23	0.94	0.091	-0.020	0.020
2	5	2.79	0.28	1.03	0.099	0.000	0.026
3	0	1.91	0.11	0.65	0.060	0.075	0.017
3	1	2.18	0.14	0.78	0.064	0.024	0.013
3	2	2.42	0.18	0.88	0.073	0.021	0.015
3	3	2.65	0.22	0.98	0.083	0.001	0.017
3	4	2.87	0.26	1.05	0.091	-0.019	0.019
3	5	3.07	0.30	1.12	0.098	0.005	0.023
4	0	2.27	0.17	0.82	0.075	0.064	0.020
4	1	2.50	0.19	0.92	0.077	0.016	0.016
4	2	2.73	0.22	1.00	0.082	0.018	0.017
4	3	2.94	0.26	1.08	0.088	-0.001	0.018
4	4	3.14	0.30	1.14	0.095	-0.008	0.020
4	5	3.33	0.33	1.20	0.100	0.001	0.022
5	0	2.58	0.23	0.95	0.087	0.064	0.022
5	1	2.80	0.25	1.03	0.088	0.027	0.019
5	2	3.01	0.27	1.10	0.091	0.030	0.019
5	3	3.21	0.31	1.17	0.096	0.016	0.020
5	4	3.40	0.34	1.22	0.100	-0.015	0.023
5	5	3.58	0.38	1.28	0.105	-0.025	0.026
6	0	2.88	0.28	1.06	0.097	0.157	0.039
6	1	3.08	0.30	1.13	0.097	0.046	0.023
6	2	3.28	0.33	1.19	0.099	0.017	0.023
6	3	3.47	0.36	1.24	0.102	0.038	0.023
6	4	3.65	0.39	1.29	0.106	-0.011	0.027
6	5	3.82	0.42	1.34	0.110	0.063	0.029

TABLE IV-B: PARSIMONIOUS MODEL ESTIMATES -- RENT DATA

$\hat{\beta}_1$ .59 (.094)	$\hat{\beta}_2$ .78 (.352)	$corr(\hat{\beta}_1, \hat{\beta}_2)$ -.521		$\hat{k}$ 178	n 7358	s^2 .01225	R^2 .142
Adults	Children	Equivalence Scale $\hat{\Delta} = \exp(\hat{\delta})$	se $\hat{\Delta}$	Log Equivalence Scale $\hat{\delta}$	se $\hat{\delta}$	$\hat{\eta}$	se $\hat{\eta}$
1	0	1.00	0.00	0.00	0.00	0.00	0.00
1	1	1.41	0.14	0.34	0.10	-0.02	0.01
1	2	1.74	0.24	0.55	0.14	-0.02	0.01
1	3	2.04	0.33	0.71	0.16	-0.02	0.01
1	4	2.31	0.40	0.84	0.17	-0.04	0.02
1	5	2.55	0.47	0.94	0.19	-0.04	0.03
2	0	1.51	0.10	0.41	0.07	-0.01	0.01
2	1	1.83	0.16	0.60	0.09	-0.02	0.01
2	2	2.12	0.24	0.75	0.12	-0.02	0.01
2	3	2.38	0.33	0.87	0.14	-0.02	0.01
2	4	2.62	0.41	0.96	0.15	-0.03	0.01
2	5	2.85	0.48	1.05	0.17	-0.04	0.02
3	0	1.91	0.20	0.65	0.10	-0.02	0.01
3	1	2.19	0.23	0.78	0.11	-0.03	0.01
3	2	2.45	0.30	0.90	0.12	-0.02	0.01
3	3	2.69	0.37	0.99	0.14	-0.02	0.01
3	4	2.91	0.44	1.07	0.15	-0.02	0.01
3	5	3.13	0.51	1.14	0.16	0.00	0.02
4	0	2.27	0.30	0.82	0.13	-0.02	0.01
4	1	2.52	0.33	0.92	0.13	-0.02	0.01
4	2	2.75	0.38	1.01	0.14	-0.01	0.01
4	3	2.97	0.44	1.09	0.15	-0.01	0.01
4	4	3.18	0.51	1.16	0.16	-0.02	0.01
4	5	3.39	0.57	1.22	0.17	-0.02	0.02
5	0	2.58	0.39	0.95	0.15	-0.01	0.01
5	1	2.82	0.42	1.04	0.15	-0.02	0.01
5	2	3.03	0.47	1.11	0.15	0.00	0.01
5	3	3.24	0.52	1.18	0.16	-0.01	0.01
5	4	3.44	0.58	1.24	0.17	0.00	0.02
5	5	3.63	0.64	1.29	0.18	0.00	0.02
6	0	2.88	0.48	1.06	0.17	-0.01	0.02
6	1	3.09	0.51	1.13	0.17	-0.03	0.02
6	2	3.30	0.55	1.19	0.17	-0.02	0.02
6	3	3.50	0.60	1.25	0.17	-0.02	0.02
6	4	3.68	0.66	1.30	0.18	-0.02	0.02
6	5	3.87	0.71	1.35	0.18	-0.04	0.02

4. Efficiency Comparisons and Monte Carlo Results

We now consider two issues: the relative asymptotic efficiency of various specifications, and the accuracy of the asymptotic approximation. Beginning with asymptotic efficiency, consider first the single equation model (1.3). Since the dummy variables in z are mutually orthogonal one might not have expected substantial improvement in precision when estimating δ using multi-family data relative to pair-wise estimation. The benefit, however, arises from the fact that under base independence, f is the same for all households. Roughly speaking, the precision of $\hat{f}$ depends on the total sample size n while the precision of household specific parameters depends directly on the number of observations in that group and indirectly on the precision of $\hat{f}$. In the parsimonious specification (1.5), equivalence scales (1.4) are a two-parameter function of the number of adults and children, thus data on all family types are informative in estimation of a particular equivalence scale. For the two-good model equation (2.12), since δ is common across equations one would expect improved efficiency from joint estimation across goods — whether by combining single equation estimates using GLS or by system estimation. [12]

To perform efficiency comparisons, we assume that the data were actually generated by the parsimonious model with parameters set roughly equal to the estimated values. That is, $\beta_1 = .60$, $\beta_2 = .75$; residuals in the food and rent equations have mean zero and covariance parameters $\sigma_1^2 = .0182$, $\sigma_2^2 = .0123$, $\sigma_{12} = -.0057$; and, the distribution of family types and expenditures are fixed at the values observed in our 7358 observations. We may now *compute* the asymptotic standard errors of various estimators (though some would be difficult to implement in practice). Table V summarizes the results. Consider single equation estimation using food data. As we have seen in Table III-A, the multi-family specification (1.3) can produce significant efficiency improvements relative to pair-wise estimation even if one estimates over relatively small groupings

[12] The cross equation restrictions should improve the efficiency of $\hat{\eta}_1, \hat{\eta}_2$ as well. To see this, refer to (2.12) and consider the extreme case where f_2 is linear. If one tries to apply single equation methods to equation 2, then δ and η_2 will not be separately identified. However, if f_1 is non-linear, then δ may be identified by applying single equation estimation to equation 1, then substituting the resulting $\hat{\delta}$ in equation 2 to produce a consistent estimate of η_2.

consisting of four family types. Gains are greatest for those types with fewest data-points. Column 4 of Table V gives asymptotic standard errors if one were to use *all* 36 family types when estimating model (1.3), (this would require a grid search in a 35 dimensional space). The gain in efficiency relative to pair-wise estimation can be as high as 25% (Table V, columns 3 and 4). However, the parsimonious model (column 5) yields far greater gains — often a three or four-fold reduction in standard errors, (Table V, columns 3 and 5).

Turning now to two-equation estimation, if one estimates equivalence scales separately for the food and rent equations using all families in equation (1.3), then applies GLS, (see Appendix (A.23), (A.24)), there is a modest reduction in standard errors -- on the order of 2-3%, (Table V columns 4 and 6). Joint estimation of both equations using all families (equation (2.14)), yields a further improvement of 1-2% (Table V columns 6 and 7). Joint estimation of the parsimonious model yields about 1-2% improvement over estimation of the food equation alone (Table V columns 5 and 8).

In summary the overwhelming source of efficiency gains arise from parsimonious specification of the equivalence scale as a function of the number of adults and children in the family.

To provide some indication of the accuracy of the asymptotic approximation we perform a small number of Monte Carlo simulations on two specifications: pair-wise estimation as proposed by Pendakur (but using our estimator) and the parsimonious model. We assume the following model for equation (1.3):

$$y \;=\; 1.73 - .2\left(\log x - z\delta\right) + \varphi \sin\left(\log x - z\delta\right) + z\eta + \varepsilon \tag{4.1}$$

We have selected this specification because by varying the coefficient of the *sin* term we can implement simple departures from the loglinear model which itself often provides a good fit to the estimation of food Engel curves. (As has been discussed in the introduction, as φ goes to 0, model (1.3) becomes unidentified.) For the pair-wise specification we assume that there are 1000 observations equally divided between two types of families A and B. The observations for families of type A, have *log x* equally spaced on the interval [5,8]; for type B families *log x* is equally spaced on the interval [5.5,8.5]. (These ranges are roughly consistent with those we observed for singles and couples with no children.) The other parameters are $\delta = .5$, $\eta = .04$, $\varphi = .2$, $\sigma^2 = .002$. The latter was selected so that the R^2 of the regression would be approximately 50%. In estimation we used $k = 40$ symmetric nearest neighbors. When calculating standard errors, the derivative of f needs to be estimated. We used a conventional kernel derivative estimator (see, e.g., Härdle (1990)) with various bandwidths h. Figure 2 summarizes the results based on 5000 replications and h=.3. The means of the sampling distributions of $\hat{\delta}$ and $\hat{\eta}$ are close to the true parameter values. They are roughly symmetric and appear to be approximately normal. The average estimated standard errors differ moderately from the standard deviations of each sampling distribution. We did indeed find that the estimated standard errors were quite sensitive to the selection of the bandwidth h. We performed simulations for various values of φ and found (as expected) that as this parameter approaches zero and the model approaches the loglinearity, the precision of $\hat{\delta}$ and $\hat{\eta}$ deteriorates.

We turn now to simulations of parsimonious model estimation. We set $\beta_1 = .6$, $\beta_2 = .6$, $\eta = 0$ and assume an equal number of families of each of the 36 types in Table I. For a given family type, the observations on $log\, x$ are equally spaced on the interval $\left[5 + .6\log(A + .6K),\ 8 + .6\log(A + .6K)\right]$ which essentially ensures that the values of equivalent income for each family type is the same. Our model is given by:

$$y = 1.73 - .2\left(\log x - .6\log(A + .6K)\right) + .2\sin\left(\log x - .6\log(A + .6K)\right) + z\eta + \varepsilon \tag{4.2}$$

Our simulations are performed for 10, 20 and 40 observations on each family type leading to a total of 360, 720 and 1440 observations respectively. We also simulate for two different values of the residual variance, $\sigma^2 = .002$, $.0005$ with corresponding values of $R^2 \cong .55$, $.85$. Figures 3A,B which summarize the results, display standardized sampling distributions (solid lines) and the standard normal (dotted lines). As expected, precision increases with sample size. However, for sample sizes of 360 and 720, two modes are sometimes observed. For β_2 this is because some estimates are at the boundary of the search-space (grid searches were performed over rectangular subsets of the unit square). For sample sizes of 1440, the standardized sampling distributions appear close to the standard normal.

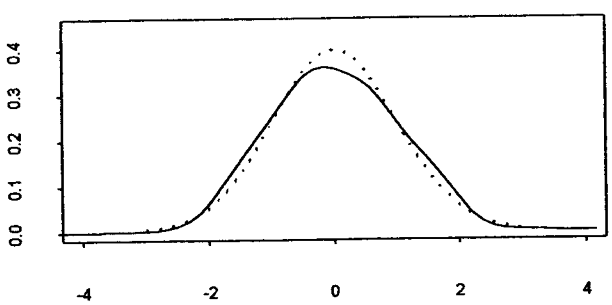
TABLE V — RELATIVE EFFICIENCY OF ESTIMATED EQUIVALENCE SCALES

Family Configuration		Single Equation Model— Food Data			Two Equation Models — Food and Rent Data		
n adults	n kids	Pair-Wise Models Eqn (1.2)	All Family Types Eqn (1.3)	Parsimonious Model Eqn (1.5)	GLS All Family Types Eqns (1.3), (A.24)	Joint Estimation All Family Types Eqn (2.14)	Joint Estimation Parsimonious Model Eqn (A.16)
1	0	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.200	0.183	0.047	0.179	0.177	0.046
	2	0.215	0.191	0.067	0.188	0.185	0.064
	3	0.251	0.205	0.078	0.202	0.200	0.075
	4	0.315	0.255	0.086	0.253	0.251	0.082
	5	0.688	0.581	0.092	0.580	0.578	0.088
2	0	0.104	0.099	0.032	0.093	0.091	0.031
	1	0.124	0.111	0.043	0.105	0.104	0.041
	2	0.124	0.110	0.058	0.105	0.103	0.055
	3	0.150	0.123	0.069	0.119	0.118	0.066
	4	0.217	0.160	0.078	0.157	0.155	0.074
	5	0.302	0.233	0.084	0.230	0.228	0.081
3	0	0.140	0.126	0.051	0.120	0.118	0.049
	1	0.143	0.125	0.054	0.120	0.119	0.052
	2	0.149	0.120	0.063	0.116	0.115	0.060
	3	0.164	0.132	0.071	0.129	0.128	0.067
	4	0.221	0.170	0.077	0.168	0.166	0.074
	5	0.338	0.248	0.084	0.246	0.244	0.080
4	0	0.177	0.147	0.064	0.140	0.138	0.062
	1	0.168	0.136	0.066	0.132	0.131	0.063
	2	0.162	0.135	0.070	0.132	0.130	0.067
	3	0.201	0.155	0.076	0.151	0.150	0.072
	4	0.267	0.205	0.081	0.203	0.201	0.077
	5	0.290	0.215	0.086	0.212	0.211	0.082
5	0	0.233	0.191	0.075	0.186	0.184	0.072
	1	0.224	0.178	0.075	0.173	0.171	0.072
	2	0.201	0.163	0.078	0.161	0.159	0.074
	3	0.238	0.189	0.082	0.186	0.184	0.078
	4	0.347	0.271	0.086	0.269	0.266	0.082
	5	0.402	0.272	0.089	0.270	0.267	0.085
6	0	0.315	0.265	0.083	0.259	0.257	0.080
	1	0.337	0.255	0.083	0.251	0.250	0.079
	2	0.269	0.211	0.085	0.207	0.206	0.081
	3	0.260	0.210	0.088	0.207	0.205	0.084
	4	0.370	0.298	0.091	0.296	0.294	0.086
	5	0.391	0.332	0.094	0.329	0.327	0.089

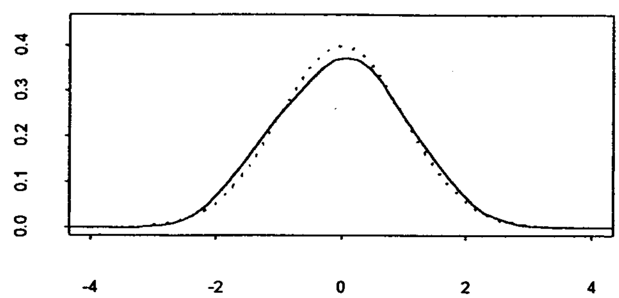
FIGURE 2: PAIR-WISE ESTIMATION — MONTE CARLO

	$\hat{\delta}$ $(\delta = .5)$		$\hat{\eta}$ $(\eta = .04)$
mean	.4971	mean	.0402
s.e.	.1292	s.e.	.0089
average estimated s.e.	.1395	average estimated s.e.	.0076

Sampling Distribution of $\hat{\delta}$



Sampling Distribution of $\hat{\eta}$



Sample size **n=1000**. Data generating mechanism given by equation (4.1). Replications = 5000; bandwidth used to estimate f is h =.3. Dotted line represents N(0,1).

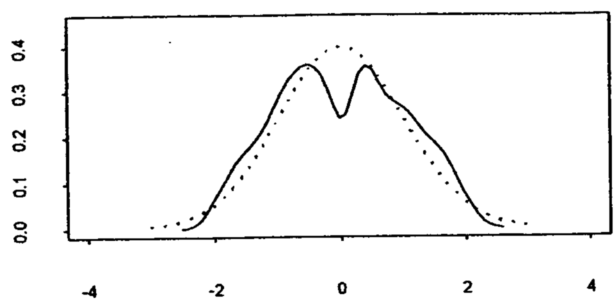
FIGURE 3A: PARSIMONIOUS MODEL — MONTE CARLO
Model $R^2 \cong .55$, $\sigma^2 \cong .002$

$\hat{\beta}_1$ $(\beta_1 = .6)$		$\hat{\beta}_2$ $(\beta_2 = .6)$	

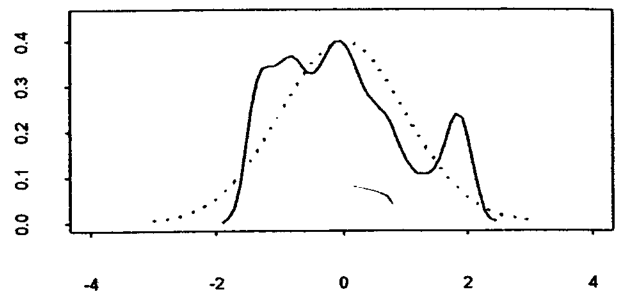
Sample Size n = 360

mean	.5952	mean	.6110
s.e.	.1765	s.e.	.2503
average estimated s.e.	.1287	average estimated s.e.	.2702

Sampling Distribution of $\hat{\beta}_1$



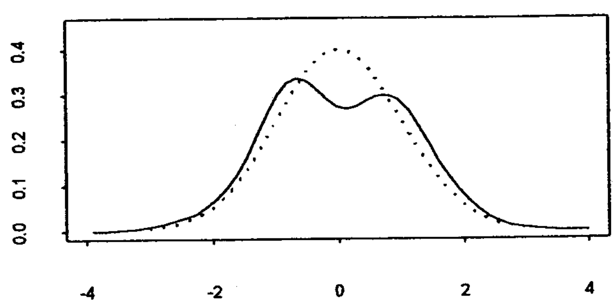
Sampling Distribution of $\hat{\beta}_2$



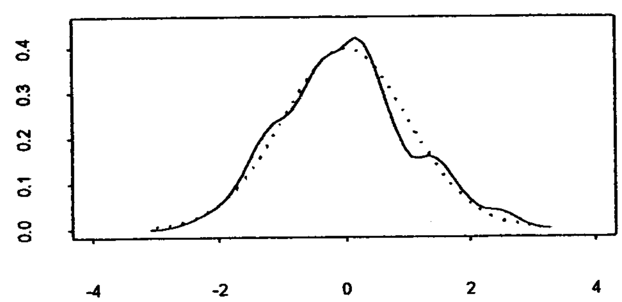
Sample Size n = 720

mean	.5964	mean	.5980
s.e.	.1013	s.e.	.2171
average estimated s.e.	.0923	average estimated s.e.	.2208

Sampling Distribution of $\hat{\beta}_1$



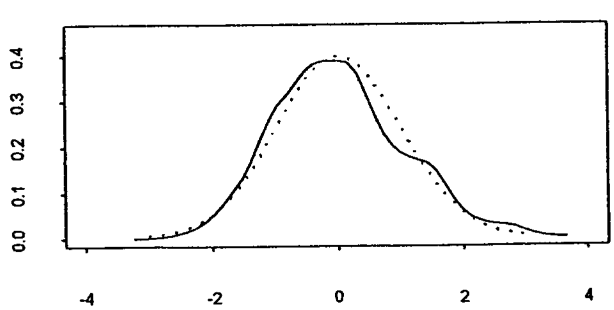
Sampling Distribution of $\hat{\beta}_2$



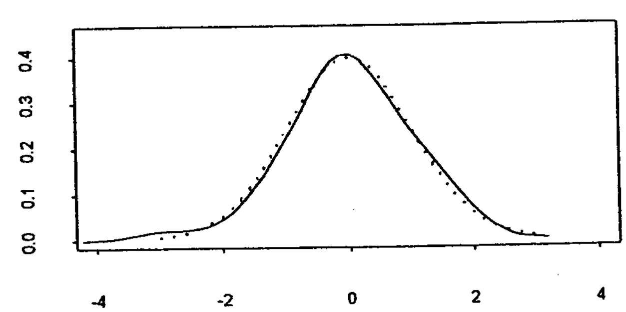
Sample Size n = 1440

mean	.5960	mean	.6001
s.e.	.0866	s.e.	.1993
average estimated s.e.	.0761	average estimated s.e.	.1624

Sampling Distribution of $\hat{\beta}_1$



Sampling Distribution of $\hat{\beta}_2$

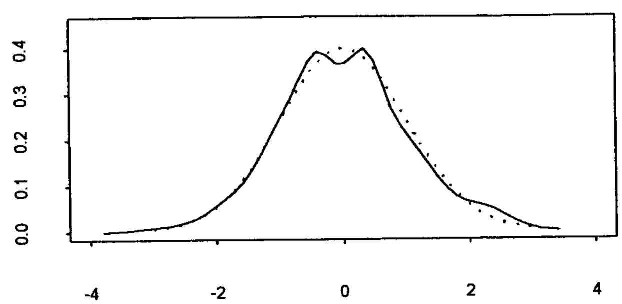


Data generating mechanism given by equation (4.2). Replications = 1000 . Dotted line represents N(0,1).

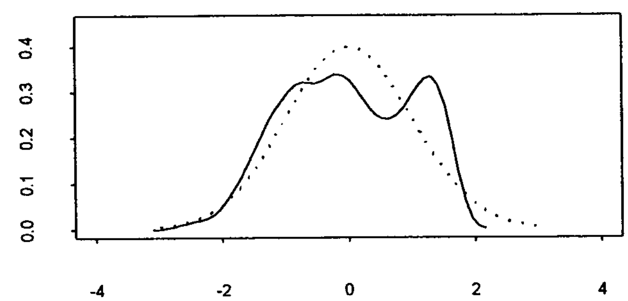
FIGURE 3B: PARSIMONIOUS MODEL — MONTE CARLO
Model $R^2 \cong .85$, $\sigma^2 \cong .0005$

$\hat{\beta}_1$ $(\beta_1 = .6)$		$\hat{\beta}_2$ $(\beta_2 = .6)$	
Sample Size n = 360			
mean	.5977	mean	.5658
s.e.	.0777	s.e.	.1196
average estimated s.e.	.0572	average estimated s.e.	.1343

Sampling Distribution of $\hat{\beta}_1$

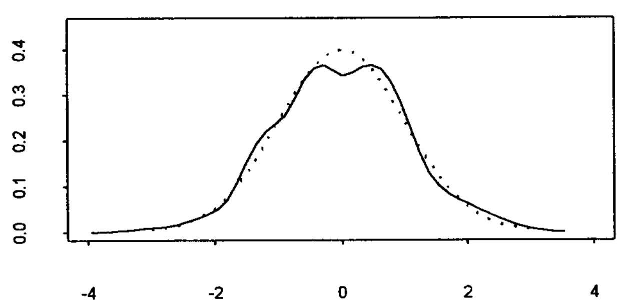


Sampling Distribution of $\hat{\beta}_2$

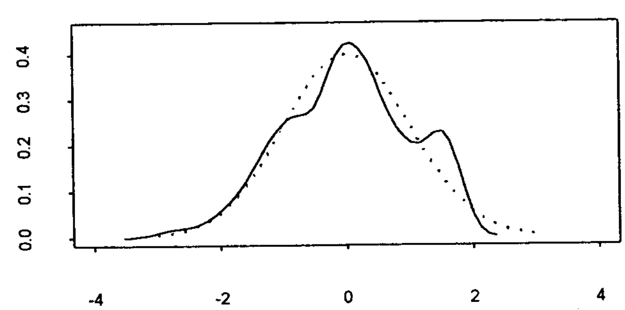


$\hat{\beta}_1$		$\hat{\beta}_2$	
Sample Size n = 720			
mean	.5938	mean	.5984
s.e.	.0600	s.e.	.1120
average estimated s.e.	.0489	average estimated s.e.	.1152

Sampling Distribution of $\hat{\beta}_1$

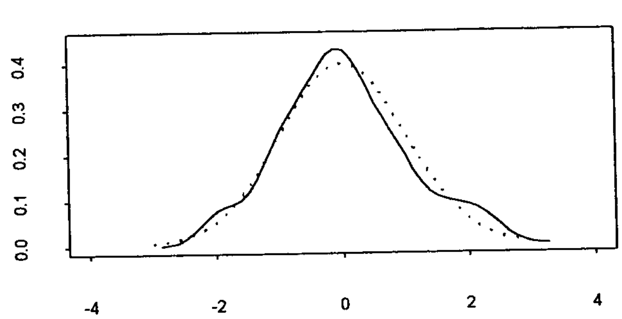


Sampling Distribution of $\hat{\beta}_2$

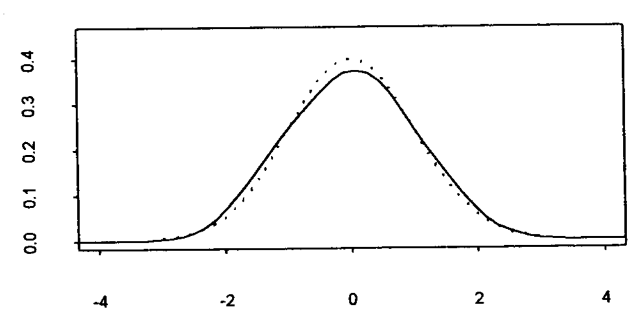


$\hat{\beta}_1$		$\hat{\beta}_2$	
Sample Size n = 1440			
mean	.5991	mean	.6020
s.e.	.0413	s.e.	.1002
average estimated s.e.	.0362	average estimated s.e.	.0963

Sampling Distribution of $\hat{\beta}_1$



Sampling Distribution of $\hat{\beta}_2$



Data generating mechanism given by equation (4.2). Replications = 1000 . Dotted line represents N(0,1).

5. CONCLUSIONS

A range of methodologies are proposed in the vast literature on equivalence scales.[13] Among these are subjective approaches, techniques based on assessments of minimum nutritional standards and utility based analyses. The latter include Engel and Rothbarth methods as well as more general demand system approaches. Strong parametric assumptions are usually part of the specification. Pendakur's generalization allows the calculation of equivalence scales by comparing pairs of nonparametrically estimated Engel curves. We embed Pendakur's semiparametric specification in a partial linear index model framework which permits us to simultaneously analyze data on multiple family types. We also propose a parsimonious variant which has further benefits. First, since this specification reduces to two the number of parameters within the index function portion of the model, estimates of equivalence scales are considerably more efficient and much easier to compute. This is particularly useful when studying developing country data where the number of prevalent family types is usually substantially larger than in data on developed countries. Second, the specification has the appealing property that it is monotone in the number of adults and children and yields plausible 'spacings' of equivalence scales among families of similar composition - an essential feature from an equity point of view if model estimates are to be used for policy purposes.

These benefits are highlighted in our analysis of South African household survey data where our estimated equivalence scales are approximately $(A+.74K)^{.59}$. There is up to 25 percent improvement in efficiency when estimating the multi-family model as compared with the pair-wise model. The parsimonious model, on the other hand, yields as much as a *four-fold* increase in precision. Multiple equation estimation (using food and rent data), yields modest gains of less than 5 percent. In Monte Carlo simulations we find the asymptotically normal approximation to be reasonable at sample sizes over 1000. Finally, we note that estimated standard errors are quite sensitive to selection of the smoothing parameter used to estimate the derivative of the Engel curve. Bootstrap standard errors, though costly to compute, could provide a useful check on those calculated by asymptotic methods.

[13] See Blundell, Preston and Walker (1994), Citro and Michael (1995), Van Praag and Warnaar (1997) and the numerous references within.

REFERENCES

Blackorby, C. and D. Donaldson (1989): "Adult Equivalence Scales, Interpersonal Comparisons of Well-Being and Applied Welfare Eocnomics", University of British Columbia, Department of Economics Discussion paper 89-24.

Blackorby, C. and D. Donaldson (1993): "Adult Equivalence Scales and the Economic Implementation of Interpersonal Comparisons of Well-Being" *Social Choice and Applied Welfare*, 10, 335-361.

Blundell, R., A. Duncan and K. Pendakur (1998): "Semiparametric Estimation and Consumer Demand". *Journal of Applied Econometrics*, 13, 435-461.

Blundell, R., I. Preston and I. Walker (1994) *The Measurement of Household Welfare*, Cambridge University Press.

Carroll R.J., J. Fan, I. Gijbels and M.P. Wand (1997): "Generalized Partially Linear Single Index Models", *Journal of the American Statistical Association*, 92, 477-89.

Cavanagh, C., R. Sherman (1998): "Rank Estimator for Monotonic Index Models", *Journal of Econometrics*, 84, 351-81.

Citro, C. and R. T. Michael, editors (1995) *Measuring Poverty -- A New Approach*, National Academy Press, Washington D.C.

Deaton, A. (1997): *The Analysis of Household Surveys*, John Hopkins University Press.

Deaton, A. and J. Muellbauer" (1986): "On Measuring Child Costs: With Applications to Poor Countries", *Journal of Political Economy*, 94, 720-744.

Fan, Y. and Q. Li (1996): "Consistent Model Specification Tests: Omitted Variables and Semiparametric Functional Forms", *Econometrica*, 64, 865-890.

Hall, P., J.W. Kay and D.M. Titterington (1990): "Asymptotically Optimal Difference-Based Estimation of Variance in Nonparametric Regression", *Biometrika*, 77, 521-8.

Härdle W. (1990): *Applied Nonparametric Regression*, Cambridge University Press.

Härdle W., P. Hall and H. Ichimura (1993): "Optimal Smoothing in Single-Index Models", *Annals of Statistics*, 21, 157-78.

Härdle, W. and J.S. Marron (1990): "Semiparametric Comparison of Regression Curves", *Annals of Statistics*, 18, 63-89.

Härdle, W. and J.S. Marron (1993): "Comparing Nonparametric versus Parametric Regression Fits", *Annals of Statistics*, 21,1926-47.

Horowitz, J. and W. Härdle (1996): "Direct Semiparametric Estimation of Single-Index Models With Discrete Covariates", *Journal of the American Statistical Association*, 91, 1632-1640.

Ichimura, H. (1993): "Semiparametric Least Squares (SLS) and Weighted SLS Estimation of Single-Index Models", *Journal of Econometrics*, 58, 71-120.

Klein, R.W., and R.H. Spady (1993): "An Efficient Semiparametric Estimator for Binary Response Models" *Econometrica*, 61: 387-421.

Lewbel, A. (1989): "Household Equivalence Scales and Welfare Comparisons", *Journal of Public Economics*, 39, 377-391

Pendakur, K. (1998); "Semiparametric Estimates and Tests of Base-Independent Equivalence Scales", *Journal of Econometrics*, 88, 1-40.

Pinkse, C. and P. M. Robinson (1995): "Pooling Nonparametric Estimates of Regression Functions with Similar Shape", in *Advances in Econometrics and Quantitative Economics*, G.S. Maddala, P.C.B Phillips and T.N. Srinivasan, eds., Blackwell, Oxford, 172-195.

Powell, J. L.(1994), "Estimation of Semiparametric Models", in the *Handbook of Econometrics, Vol. IV*, R. Engle and D. McFadden, eds., North Holland, 2443-2521.

Robinson, P.M. (1988): "Root-*N*-Consistent Semiparametric Regression", *Econometrica*, 56, 931-954.

Speckman, P. (1988): "Kernel Smoothing in Partial Linear Models", *Journal of the Royal Statistical Society, B*, 50, 413-36.

Van Praag, B.M.S. and M. F. Warnaar (1997): "The Cost of Children and the Use of Demographic Variables in Consumer Demand", in *Handbook of Population and Family Economics,* Vol. 1A, M.R. Rozenzweig and O. Stark, eds., North Holland, 241-273.

Yatchew, A.J. (1997): "An Elementary Estimator of the Partial Linear Model", *Economics Letters,* 57, 135-143, additional examples contained in *Economics Letters* 1998, 59, 403-405.

Yatchew, A.J. (1998): "Nonparametric Regression Techniques in Economics", *Journal of Economic Literature,* vol. XXXVI, 669-721.

Yatchew, A.J. (1999a): "An Elementary Nonparametric Differencing Test of Equality of Regression Functions ", *Economics Letters,* 62, 271-278.

Yatchew, A.J. (1999b):"Differencing Methods in Nonparametric Regression: Simple Techniques for the Applied Econometrician", manuscript.

Yatchew, A.J. (2000): "Scale Economies in Electricity Distribution: A Semiparametric Analysis", *Journal of Applied Econometrics*, 15, 187-210.

APPENDIX

SINGLE EQUATION ESTIMATION

PROPOSITION 1: Consider the model $y_i = f\left(r\left(w_i, \beta\right)\right) + z_i \eta + \varepsilon_i$ where w_i and z_i are finite-dimensional vectors of exogenous variables; f is a nonparametric function; r is a known function; β and η are finite-dimensional parameter vectors; β_o, η_o and f_o are the true parameter values; $r_o = r\left(w, \beta_o\right)$; and $\varepsilon_i \mid w_i, z_i$ are i.i.d with mean 0, variance σ^2 . Set up the model in matrix notation where the i-th rows of W and Z are w_i and z_i respectively:

$$y = f\left(r\left(W, \beta\right)\right) + Z\eta + \varepsilon \qquad \text{(A.1)}$$

To estimate this model, proceed analogously to equations (2.3)-(2.5). Fix k. For any β let P_β be the permutation matrix which reorders the vector $r\left(W, \beta\right)$ so that its elements are in increasing order. Apply $\left(I - S\right)P_\beta$ to (A.1) to obtain:

$$\left(I - S\right)P_\beta y = \left(I - S\right)P_\beta f\left(r\left(W, \beta\right)\right) + \left(I - S\right)P_\beta Z\eta + \left(I - S\right)P_\beta \varepsilon \qquad \text{(A.2)}$$

define

$$\hat{\eta}_\beta = \left[\left(\left(I - S\right)P_\beta Z\right)'\left(I - S\right)P_\beta Z\right]^{-1}\left(\left(I - S\right)P_\beta Z\right)'\left(I - S\right)P_\beta y \qquad \text{(A.3)}$$

and by a grid search over values of β and k find:

$$s^2 = \min_k \min_\beta \frac{1}{n}\left(\left(I - S\right)P_\beta y - \left(I - S\right)P_\beta Z\hat{\eta}_\beta\right)'\left(\left(I - S\right)P_\beta y - \left(I - S\right)P_\beta Z\hat{\eta}_\beta\right) \qquad \text{(A.4)}$$

$\hat{k}, \hat{\beta}$ will be the values which achieve this minimum and the estimator of η will be $\hat{\eta} = \hat{\eta}_{\hat{\beta}}$.[14]

---

[14] We have employed matrix notation in defining our estimators in equations (2.3)-(2.5) and (A.2)-(A.4) for expositional purposes only. The calculation of objects like $\left(I - S\right)P_\beta y$ is performed without multiplication of large matrices. For fixed β, reorder $r\left(W, \beta\right)$ so that its elements are in increasing order. Sort y into the same order and call the reordered vector y_β. Then use (2.2) to calculate $\left(I - S\right)P_\beta y$ as $y_\beta - \left(y_{\beta_{-k/2}} + \ldots + y_{\beta_{-1}} + y_{\beta_{+1}} + \ldots + y_{\beta_{+k/2}}\right)/k$ where the subscripts denote lags and leads. The matrix $\left(I - S\right)P_\beta Z$ is calculated in a similar fashion.

To obtain large sample standard errors define the following conditional covariance matrices:

$$\begin{aligned}
\Sigma_z &= E\left[\left(z-E(z\,|\,r_o)\right)'\left(z-E(z\,|\,r_o)\right) \,|\, r_o\right] \\
\Sigma_{zf'} &= E\left[f'(r_o)\left(\frac{\partial r}{\partial\beta}-E\left(\frac{\partial r}{\partial\beta}\Big|r_o\right)\right)'\left(z-E(z\,|\,r_o)\right) \,|\, r_o\right] \\
\Sigma_{f'} &= E\left[f'(r_o)^2\left(\frac{\partial r}{\partial\beta}-E\left(\frac{\partial r}{\partial\beta}\Big|r_o\right)\right)'\left(\frac{\partial r}{\partial\beta}-E\left(\frac{\partial r}{\partial\beta}\Big|r_o\right)\right) \,|\, r_o\right]
\end{aligned} \tag{A.5}$$

Let

$$V = \begin{bmatrix} \Sigma_z & \Sigma_{zf'} \\ \Sigma_{zf'}' & \Sigma_{f'} \end{bmatrix} \tag{A.6}$$

then

$$n^{1/2}\begin{pmatrix} \hat\eta - \eta_o \\ \hat\beta - \beta_o \end{pmatrix} \xrightarrow{D} N\left(0,\, \sigma^2 V^{-1}\right) \tag{A.7}$$

and $\hat{k} = O_P\left(n^{4/5}\right)$ which is the optimal rate. Let $\hat{f}'$ be any consistent estimator of the first derivative of f and define $diag\left(\hat{f}'(\cdot)\right)$ to be the $\left(n-\hat{k}\right)xn$ diagonal matrix with diagonal elements $\hat{f}'\left(r\left(w,\hat\beta\right)\right)$, $i = \hat{k}/2+1,\ldots,n-\hat{k}/2$, ($\hat{k}$ is even). Define R to be the matrix whose *i-th* row is the vector partial derivative $\partial r\,(w_i,\hat\beta)\,/\partial\beta$. Then

$$\begin{aligned}
\frac{1}{n}\left((I-S)P_{\hat\beta}Z\right)'\left((I-S)P_{\hat\beta}Z\right) &\xrightarrow{P} \Sigma_z \\
\frac{1}{n}\left((I-S)P_{\hat\beta}R\right)' diag\left(\hat{f}'(\cdot)\right)\left((I-S)P_{\hat\beta}Z\right) &\xrightarrow{P} \Sigma_{zf'} \\
\frac{1}{n}\left((I-S)P_{\hat\beta}R\right)' diag\left(\hat{f}'(\cdot)\right) diag\left(\hat{f}'(\cdot)\right)\left((I-S)P_{\hat\beta}R\right) &\xrightarrow{P} \Sigma_{f'} \quad \blacksquare
\end{aligned} \tag{A.8}$$

COMMENTS: Proof of the above result which we sketch below, is a straightforward variation on existing proofs in the literature — particularly Ichimura (1993), Härdle, Hall and Ichimura (1993) and Carroll et al (1997).

In model (1.3) where separate dummies are permitted for each non-reference family, $\beta = \delta$, $w = (x,z)$, $r(w,\delta) = \log x - z\delta$ and $R = -Z$.

In the parsimonious model (1.5) with separate dummies for each family type entering only in the elasticity term, $\beta = (\beta_1, \beta_2)$, $w = (x, A, K)$, $r(w,\beta) = \log x - \beta_1 \log(A + \beta_2 K)$ and the matrix R has i-th row:

$$\left(\frac{\partial r(w_i,\beta)}{\partial \beta_1}, \frac{\partial r(w_i,\beta)}{\partial \beta_2} \right)_{\hat\beta_1,\hat\beta_2} = \left(-\log\left(A_i + \hat\beta_2 K_i \right), \quad -\frac{\hat\beta_1 K_i}{A_i + \hat\beta_2 K_i} \right) \tag{A.9}$$

PROOF OF PROPOSITION 1: Let $r_i = r(w_i, \beta)$ and $r_{oi} = r(w_i, \beta_o)$. To derive the asymptotic distribution, consider first the optimization problem:

$$\begin{aligned} &\min_{\beta\eta} \frac{1}{n} \sum \left(y_i - E(y_i | r_i) - (z_i - E(z_i | r_i))\eta\right)^2 \\ = &\min_{\beta\eta} \frac{1}{n} \sum \left(y_i - H_i - (z_i - G_i)\eta\right)^2 \end{aligned}$$

where $F_i \equiv F(r_i, \beta) \equiv E\left(f(r_{oi}) | r_i\right)$, $G_i \equiv G(r_i, \beta) \equiv E(z_i | r_i)$, $H_i \equiv H(r_i, \beta) \equiv E(y_i | r_i) = F_i + G_i \eta_o$. Differentiating and setting the first order conditions equal to 0, one obtains:

$$\begin{aligned} \eta: &\quad \frac{1}{n} \sum \left(z_i - G_i\right)' \left(y_i - H_i - (z_i - G_i)\eta\right) &= 0 \\ \beta: &\quad \frac{1}{n} \sum \left(\frac{\partial H_i}{\partial \beta} - \frac{\partial G_i}{\partial \beta} \eta \right)' \left(y_i - H_i - (z_i - G_i)\eta\right) &= 0 \end{aligned}$$

Expanding in a first order Taylor series, multiplying by $n^{1/2}$ and setting to 0 those terms which converge to 0, yields:

$$n^{1/2}\frac{1}{n}\sum\left(z_i - G_{oi}\right)'\varepsilon_i - \frac{1}{n}\sum\left(z_i - G_{oi}\right)'\left(z_i - G_{oi}\right) n^{1/2}\left(\hat{\eta} - \eta_o\right) - \frac{1}{n}\sum\left(z_i - G_{oi}\right)'\left[\frac{\partial H_i}{\partial \beta} - \frac{\partial G_i}{\partial \beta}\eta\right]_{\eta_o \beta_o} n^{1/2}\left(\hat{\beta} - \beta_o\right) \cong 0$$

$$n^{1/2}\frac{1}{n}\sum\left[\frac{\partial H_i}{\partial \beta} - \frac{\partial G_i}{\partial \beta}\eta\right]'_{\eta_o \beta_o}\varepsilon_i - \frac{1}{n}\sum\left[\frac{\partial H_i}{\partial \beta} - \frac{\partial G_i}{\partial \beta}\eta\right]'_{\eta_o \beta_o}\left(z_i - G_{oi}\right) n^{1/2}\left(\hat{\eta} - \eta_o\right) - \frac{1}{n}\sum\left[\frac{\partial H_i}{\partial \beta} - \frac{\partial G_i}{\partial \beta}\eta\right]'_{\eta_o \beta_o}\left[\frac{\partial H_i}{\partial \beta} - \frac{\partial G_i}{\partial \beta}\eta\right]_{\eta_o \beta_o} n^{1/2}\left(\hat{\beta} - \beta_o\right) \cong 0$$

where $G_{oi} \equiv G\left(r_{oi}, \beta_o\right) = E\left(z_i | r_{oi}\right)$. Next note that

$$\left[\frac{\partial H_i}{\partial \beta} - \frac{\partial G_i}{\partial \beta}\eta\right]_{\eta_o \beta_o} = \left[\frac{\partial F_i}{\partial \beta}\right]_{\beta_o} = -f'\left(r_{oi}\right)\left(\frac{\partial r_{oi}}{\partial \beta} - E\left(\frac{\partial r_{oi}}{\partial \beta} | r_{oi}\right)\right)$$

where the second equality is obtained following Klein and Spady (1993, p. 401-2).

Define U to be the matrix whose i-th row is given by $\left(z_i - E\left(z_i | r_{oi}\right)\right) / \sqrt{n}$ and W to be the matrix whose i-th row is $-f'\left(r_{oi}\right)\left(\frac{\partial r_{oi}}{\partial \beta} - E\left(\frac{\partial r_{oi}}{\partial \beta} | r_{oi}\right)\right) / \sqrt{n}$. Define $\hat{V} = \begin{bmatrix} U'U & U'W \\ W'U & W'W \end{bmatrix}$ then the system of equations may be rewritten as :

$$\hat{V}\begin{bmatrix} n^{1/2}\left(\hat{\eta} - \eta_o\right) \\ n^{1/2}\left(\hat{\beta} - \beta_o\right) \end{bmatrix} \cong \begin{bmatrix} U'\varepsilon \\ W'\varepsilon \end{bmatrix} \tag{A.10}$$

Furthermore $\hat{V} \xrightarrow{P} V$ and $\begin{bmatrix} U'\varepsilon \\ W'\varepsilon \end{bmatrix} \xrightarrow{D} N\left(0, \sigma^2 V\right)$ in which case $\begin{bmatrix} n^{1/2}\left(\hat{\eta} - \eta_o\right) \\ n^{1/2}\left(\hat{\beta} - \beta_o\right) \end{bmatrix} \xrightarrow{D} N\left(0, \sigma^2 V^{-1}\right)$. ■

COMMENT: Applying partitioned matrix inversion to (A.10) obtain the following which will be used below:

$$n^{1/2}\left(\hat{\beta} - \beta_o\right) \cong \left[W'\left(I - U\left[U'U\right]^{-1}U'\right)W\right]^{-1} W'\left[I - U\left[U'U\right]^{-1}U'\right]\varepsilon \quad ■ \tag{A.11}$$

TESTING PROCEDURES: In order to implement the tests outlined in Section 2, we calculate restricted and unrestricted estimates of the residual variance. In each case the null hypothesis which yields s^2_{res} is a base-independent specification (1.3) or (1.5). Using analysis similar to Hall, Härdle and Ichimura (1993), it can be shown that:

$$s^2_{res} = \frac{1}{n}\varepsilon'\varepsilon + O_P\left(\frac{1}{n^{4/5}}\right) \tag{A.12}$$

Let D be an optimal differencing matrix D. Then $D'D \cong I - \frac{1}{2m}\left(L_1 + L_1' + \ldots + L_m + L_m'\right)$, where m is the order of differencing and the L_i are lag matrices. That is, for $i>0$, L_i has 0's everywhere except on the i-th diagonal below the main diagonal where it has 1's. Consider a generic nonparametric regression $y = f(x) + \varepsilon$, where y and x are scalars, x has bounded domain and the data have been reordered so that $x_1 \le .... \le x_n$. Then using arguments as in Yatchew (1997, Appendix A) or Yatchew (2000, Appendix 1), s^2_{diff} defined in (2.7) satisfies:

$$\begin{aligned} s^2_{diff} &= \frac{1}{n}\varepsilon'\varepsilon - \frac{1}{2mn}\varepsilon'\left(L_1 + L_1' + \ldots + L_m + L_m'\right)\varepsilon + O_P\left(\frac{1}{n^{3/2}}\right) \\ &= \frac{1}{n}\varepsilon'\varepsilon - \frac{1}{mn}\varepsilon'\left(L_1 + L_2 + \ldots + L_m\right)\varepsilon + O_P\left(\frac{1}{n^{3/2}}\right) \end{aligned} \tag{A.13}$$

from which it follows that the unrestricted residual variance estimator s^2_{unr} in (2.10), which is a linear combination of differencing estimators, also satisfies (A.13). Combining these results we have:

$$(mn)^{1/2}\left(s^2_{res} - s^2_{unr}\right) = \frac{(mn)^{1/2}}{mn}\varepsilon'\left(L_1 + L_2 + \ldots + L_m\right)\varepsilon + O_P\left(\frac{(mn)^{1/2}}{n^{4/5}}\right) \tag{A.14}$$

Using a finitely dependent central limit theorem, the first term on the right hand side is $N\left(0, \sigma^4\right)$ from which (2.8) follows immediately.

Next, suppose we are testing the parsimonious model (1.5) against the base-independent specification (1.3). To obtain the unrestricted estimate s^2_{unr} we apply differencing to the ordered pairs

$\left(y_i - z_i\hat{\eta},\ \log x_i - z_i\hat{\delta}\right)$ $i = 1,\ldots,n$. In this case

$$s^2_{unr} = \frac{1}{n}\varepsilon'\varepsilon - \frac{1}{2mn}\varepsilon'\left(L_1 + L_1' + \ldots + L_m + L_m'\right)\varepsilon + O_P\left(\frac{1}{n}\right) \qquad \text{(A.15)}$$

The remainder term has a slower rate of convergence than in (A.13) because $n^{1/2}$consistent estimates are used for η and δ. Once again (2.8) follows directly. It can also be shown that the unrestricted estimator in (2.11) also satisfies (A.15).

The simplicity of (2.8) derives from the simple structure of $D'D$ and from the fact that the asymptotic distribution of the test statistic depends only on the second term on the right hand side of (A.13). By fixing *m* we are effectively under-smoothing the alternative hypothesis.[15] The order of differencing *m* may be permitted to increase with sample size so long as remainder terms still vanish. For example, in (A.14) we require that $m^{1/2}/n^{3/10} \rightarrow 0$ that is, $m/n^{3/5} \rightarrow 0$. ■

[15] This technique is used in nonparametric regression to simplify distribution theory. See e.g, Fan and Li (1996, p. 871) and Härdle (1990, p. 101).

PROPOSITION 2: As before, let $r_i = r(w_i, \beta)$ and $r_{oi} = r(w_i, \beta_o)$ and consider the two-good analogue to (A.1):

$$\begin{aligned} y_1 &= f_1(r(W,\beta)) + Z\eta_1 + \varepsilon_1 \\ y_2 &= f_2(r(W,\beta)) + Z\eta_2 + \varepsilon_2 \end{aligned} \tag{A.16}$$

where

$$\Sigma = cov\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \qquad \Sigma^{-1} = \begin{bmatrix} \sigma^{11} & \sigma^{12} \\ \sigma^{12} & \sigma^{22} \end{bmatrix} \tag{A.17}$$

Define the following conditional covariance matrices:

$$\begin{aligned}
\Sigma_z &= E\left[\left(z - E(z|r_o)\right)'\left(z - E(z|r_o)\right) \mid r_o\right] \\
\Sigma_{zf_1'} &= E\left[f_1'(r_o)\left(\frac{\partial r}{\partial \beta} - E\left(\frac{\partial r}{\partial \beta}|r_o\right)\right)'\left(z - E(z|r_o)\right) \mid r_o\right] \\
\Sigma_{zf_2'} &= E\left[f_2'(r_o)\left(\frac{\partial r}{\partial \beta} - E\left(\frac{\partial r}{\partial \beta}|r_o\right)\right)'\left(z - E(z|r_o)\right) \mid r_o\right] \\
\Sigma_{f_1'f_1'} &= E\left[f_1'(r_o)^2\left(\frac{\partial r}{\partial \beta} - E\left(\frac{\partial r}{\partial \beta}|r_o\right)\right)'\left(\frac{\partial r}{\partial \beta} - E\left(\frac{\partial r}{\partial \beta}|r_o\right)\right) \mid r_o\right] \\
\Sigma_{f_2'f_2'} &= E\left[f_2'(r_o)^2\left(\frac{\partial r}{\partial \beta} - E\left(\frac{\partial r}{\partial \beta}|r_o\right)\right)'\left(\frac{\partial r}{\partial \beta} - E\left(\frac{\partial r}{\partial \beta}|r_o\right)\right) \mid r_o\right] \\
\Sigma_{f_1'f_2'} &= E\left[f_1'(r_o)f_2'(r_o)\left(\frac{\partial r}{\partial \beta} - E\left(\frac{\partial r}{\partial \beta}|r_o\right)\right)'\left(\frac{\partial r}{\partial \beta} - E\left(\frac{\partial r}{\partial \beta}|r_o\right)\right) \mid r_o\right]
\end{aligned} \tag{A.18}$$

Let

$$V = \begin{bmatrix} \Sigma^{-1} \otimes \Sigma_z & \begin{pmatrix} \sigma^{11} \\ \sigma^{12} \end{pmatrix} \otimes \Sigma'_{zf_1'} + \begin{pmatrix} \sigma^{12} \\ \sigma^{22} \end{pmatrix} \otimes \Sigma'_{zf_2'} \\ \left(\sigma^{11} \;\; \sigma^{12}\right) \otimes \Sigma_{zf_1'} + \left(\sigma^{12} \;\; \sigma^{22}\right) \otimes \Sigma_{zf_2'} & \sigma^{11}\Sigma_{f_1'f_1'} + 2\sigma^{12}\Sigma_{f_1'f_2'} + \sigma^{22}\Sigma_{f_2'f_2'} \end{bmatrix} \tag{A.19}$$

Define $\hat{\beta}$, $\hat{\eta}_1 = \hat{\eta}_{1_{\hat{\beta}}}$, $\hat{\eta}_2 = \hat{\eta}_{2_{\hat{\beta}}}$ analogously to the procedure outlined in equations (2.13)-(2.14), then

$$n^{1/2} \begin{pmatrix} \hat{\eta}_1 - \eta_1 \\ \hat{\eta}_2 - \eta_2 \\ \hat{\beta} - \beta \end{pmatrix} \xrightarrow{D} N\left(0, V^{-1}\right) \tag{A.20}$$

Recall that in estimating the multi-equation model, we use the optimized values of the smoothing parameters $\hat{k}_1$, $\hat{k}_2$ from the single equation procedures. Let $\hat{k}_{max} = \max\left\{\hat{k}_1, \hat{k}_2\right\}$. Let $\hat{f}_1^{\prime}$ be any consistent estimator of the first derivative of $f_1^{\prime}$, define $diag\left(\hat{f}_1^{\prime}(\cdot)\right)$ to be the $\left(n - \hat{k}_{max}\right) xn$ diagonal matrix with diagonal elements $\hat{f}_1^{\prime}\left(r\left(w_i, \hat{\beta}\right)\right)$, $i = \frac{1}{2}\hat{k}_{max}+1, \ldots, n - \frac{1}{2}\hat{k}_{max}$. Define $\hat{f}_2^{\prime}$ and $diag\left(\hat{f}_2^{\prime}(\cdot)\right)$ in a similar fashion. Let S be the smoothing matrix with smoothing parameter $\hat{k}_{max}$, then

$$\begin{aligned}
\frac{1}{n}\left((I-S)P_{\hat{\beta}}Z\right)^{\prime}\left((I-S)P_{\hat{\beta}}Z\right) &\xrightarrow{P} \Sigma_z \\
\frac{1}{n}\left((I-S)P_{\hat{\beta}}R\right)^{\prime} diag\left(\hat{f}_1^{\prime}(\cdot)\right)\left((I-S)P_{\hat{\beta}}Z\right) &\xrightarrow{P} \Sigma_{zf_1^{\prime}} \\
\frac{1}{n}\left((I-S)P_{\hat{\beta}}R\right)^{\prime} diag\left(\hat{f}_2^{\prime}(\cdot)\right)\left((I-S)P_{\hat{\beta}}Z\right) &\xrightarrow{P} \Sigma_{zf_2^{\prime}} \\
\frac{1}{n}\left((I-S)P_{\hat{\beta}}R\right)^{\prime} diag\left(\hat{f}_1^{\prime}(\cdot)\right)^2\left((I-S)P_{\hat{\beta}}R\right) &\xrightarrow{P} \Sigma_{f_1^{\prime}f_1^{\prime}} \\
\frac{1}{n}\left((I-S)P_{\hat{\beta}}R\right)^{\prime} diag\left(\hat{f}_2^{\prime}(\cdot)\right)^2\left((I-S)P_{\hat{\beta}}R\right) &\xrightarrow{P} \Sigma_{f_2^{\prime}f_2^{\prime}} \\
\frac{1}{n}\left((I-S)P_{\hat{\beta}}R\right)^{\prime} diag\left(\hat{f}_1^{\prime}(\cdot)\right) diag\left(\hat{f}_2^{\prime}(\cdot)\right)\left((I-S)P_{\hat{\beta}}R\right) &\xrightarrow{P} \Sigma_{f_1^{\prime}f_2^{\prime}} \quad \blacksquare
\end{aligned} \tag{A.21}$$

COMMENTS: Since only consistent estimation of V is required, much broader choice of the smoothing parameter is available. We chose $\hat{k}_{max}$ for simplicity. ■

PROOF OF PROPOSITION 2: To derive the asymptotic distribution, consider the optimization problem:

$$\min_{\eta_1\eta_2\beta} \frac{1}{n}\begin{pmatrix} y_1 - H_1 - (Z-G)\eta_1 \\ y_2 - H_2 - (Z-G)\eta_2 \end{pmatrix}' \left(\Sigma^{-1}\otimes I_n\right) \begin{pmatrix} y_1 - H_1 - (Z-G)\eta_1 \\ y_2 - H_2 - (Z-G)\eta_2 \end{pmatrix}$$

$$= \min_{\eta_1\eta_2\beta} \left\{ \begin{array}{l} \sigma^{11}(y_1 - H_1 - (Z-G)\eta_1)'(y_1 - H_1 - (Z-G)\eta_1) \\ + 2\sigma^{12}(y_1 - H_1 - (Z-G)\eta_1)'(y_2 - H_2 - (Z-G)\eta_2) \\ + \sigma^{22}(y_2 - H_2 - (Z-G)\eta_2)'(y_2 - H_2 - (Z-G)\eta_2) \end{array} \right\}$$

where F_1 has i-th entry $E(f_1(r_{oi})|\, r(w_i,\beta))$, F_2 has i-th entry $E(f_2(r_{oi})|\, r(w_i,\beta))$, G has i-th row $G_i \equiv E(z_i|\, r(w_i,\beta))$, $H_1 \equiv F_1 + G\eta_{1_o}$ and $H_2 \equiv F_2 + G\eta_{2_o}$. Differentiating and setting the first order conditions equal to 0, one obtains:

$$\begin{array}{ll}
\eta_1: & \sigma^{11}(Z-G)'(y_1 - H_1 - (Z-G)\eta_1) + \sigma^{12}(Z-G)'(y_2 - H_2 - (Z-G)\eta_2) = 0 \\
\eta_2: & \sigma^{12}(Z-G)'(y_1 - H_1 - (Z-G)\eta_1) + \sigma^{22}(Z-G)'(y_2 - H_2 - (Z-G)\eta_2) = 0 \\
\beta: & \sigma^{11}\left(\frac{\partial H_1}{\partial\beta} - \frac{\partial G}{\partial\beta}\eta_1\right)'(y_1 - H_1 - (Z-G)\eta_1) + \sigma^{12}\left(\frac{\partial H_1}{\partial\beta} - \frac{\partial G}{\partial\beta}\eta_1\right)'(y_2 - H_2 - (Z-G)\eta_2) \\
& + \sigma^{12}\left(\frac{\partial H_2}{\partial\beta} - \frac{\partial G}{\partial\beta}\eta_2\right)'(y_1 - H_1 - (Z-G)\eta_1) + \sigma^{22}\left(\frac{\partial H_2}{\partial\beta} - \frac{\partial G}{\partial\beta}\eta_2\right)'(y_2 - H_2 - (Z-G)\eta_2) = 0
\end{array}$$

Expanding in a first order Taylor series, multiplying by $n^{1/2}$ and setting to 0 those terms which converge to 0, yields:

$$\hat{V}\begin{bmatrix} n^{1/2}(\hat{\eta}_1 - \eta_{1_o}) \\ n^{1/2}(\hat{\eta}_2 - \eta_{2_o}) \\ n^{1/2}(\hat{\beta} - \beta_o) \end{bmatrix} \cong \begin{bmatrix} \Sigma^{-1}\otimes U' \\ \left[(\sigma^{11}\ \sigma^{12})\otimes W_1'\right] + \left[(\sigma^{12}\ \sigma^{22})\otimes W_2'\right] \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \qquad \text{(A.22)}$$

where

$$\hat{V} = \begin{bmatrix} \Sigma^{-1} \otimes U'U & \begin{pmatrix} \sigma^{11} \\ \sigma^{12} \end{pmatrix} \otimes U'W_1 + \begin{pmatrix} \sigma^{12} \\ \sigma^{22} \end{pmatrix} \otimes U'W_2 \\ \left(\sigma^{11} \;\; \sigma^{12}\right) \otimes W_1'U + \left(\sigma^{12} \;\; \sigma^{22}\right) \otimes W_2'U & \sigma^{11} W_1'W_1 + \sigma^{12}\left(W_1'W_2 + W_2'W_1\right) + \sigma^{22} W_2'W_2 \end{bmatrix}$$

As before U is the matrix whose i-th row is given by $\left(z_i - E\left(z_i | r_{oi}\right)\right) / \sqrt{n}$; W_1 and W_2 are matrices with i-th rows $-f_1'(r_{oi})\left(\frac{\partial r_{oi}}{\partial \beta} - E\left(\frac{\partial r_{oi}}{\partial \beta} | r_{oi}\right)\right) / \sqrt{n}$ and $-f_2'(r_{oi})\left(\frac{\partial r_{oi}}{\partial \beta} - E\left(\frac{\partial r_{oi}}{\partial \beta} | r_{oi}\right)\right) / \sqrt{n}$ respectively. It is straightforward to show that $\hat{V} \overset{P}{\to} V$ and (A.20) follows immediately. ■

COMBINING SINGLE EQUATION ESTIMATES: suppose we estimate model (1.3) for each of the food and rent equations. The resulting estimates of δ, say $\hat{\delta}_1$ and $\hat{\delta}_2$. Define U to have i-th row $\left(z_i - E\left(z_i | \log x_i - z_i \delta_o\right)\right) / \sqrt{n}$; W_1 and W_2 to have i-th rows $f_1'\left(\log x_i - z_i \delta_o\right)\left(z_i - E\left(z_i | \log x_i - z_i \delta_o\right)\right) / \sqrt{n}$ and $f_2'\left(\log x_i - z_i \delta_o\right)\left(z_i - E\left(z_i | \log x_i - z_i \delta_o\right)\right) / \sqrt{n}$ respectively. Define $\bar{W}_1 = \left(I - U[U'U]^{-1} U'\right) W_1$ and $\bar{W}_2 = \left(I - U[U'U]^{-1} U'\right) W_2$. Then

$$\begin{aligned} n^{1/2}\left(\hat{\delta}_1 - \delta_o\right) &\cong \left[\bar{W}_1' \bar{W}_1\right]^{-1} \bar{W}_1' \varepsilon_1 \\ n^{1/2}\left(\hat{\delta}_2 - \delta_o\right) &\cong \left[\bar{W}_2' \bar{W}_2\right]^{-1} \bar{W}_2' \varepsilon_2 \end{aligned} \qquad \text{(A.23)}$$

(see (A.11)), from which we have

$$Cov \begin{bmatrix} n^{1/2}\left(\hat{\delta}_1 - \delta_o\right) \\ n^{1/2}\left(\hat{\delta}_2 - \delta_o\right) \end{bmatrix} \cong plim \begin{bmatrix} \sigma_1^2 \left[\bar{W}_1' \bar{W}_1\right]^{-1} & \sigma_{12} \left[\bar{W}_1' \bar{W}_1\right]^{-1} \bar{W}_1' \bar{W}_2 \left[\bar{W}_2' \bar{W}_2\right]^{-1} \\ \sigma_{12} \left[\bar{W}_2' \bar{W}_2\right]^{-1} \bar{W}_2' \bar{W}_1 \left[\bar{W}_1' \bar{W}_1\right]^{-1} & \sigma_2^2 \left[\bar{W}_2' \bar{W}_2\right]^{-1} \end{bmatrix} \qquad \text{(A.24)}$$

Using the residuals from the single equation procedures we may obtain consistent estimates of σ_1^2, σ_2^2, σ_{12} (see footnote 10). We then may use (A.24) in a GLS procedure to obtain a single more efficient estimate of δ. A simple χ^2 test may be applied to test whether the δ's being estimated in each equation are equal. A similar procedure may be implemented if one is interested in combining single equation estimates of the parsimonious model (1.5). ■

APPENDIX TABLE I: Summary Statistics

Family Composition Adults	Children	n	Log of Total Monthly Expenditure Mean	Min	Max	Food Share Mean	Rent Share Mean
1	0	1109	6.644	5.591	8.562	0.401	0.157
2	0	890	7.098	5.594	8.622	0.405	0.155
3	0	373	7.147	5.596	7.848	0.437	0.131
4	0	222	7.287	5.615	8.622	0.425	0.137
5	0	105	7.272	5.672	8.517	0.418	0.135
6	0	50	7.297	5.759	8.509	0.441	0.127
1	1	138	6.485	5.604	8.562	0.547	0.106
2	1	526	6.992	5.595	8.606	0.470	0.123
3	1	314	7.078	5.603	8.616	0.466	0.118
4	1	227	7.048	5.614	8.620	0.472	0.118
5	1	117	7.196	5.712	8.601	0.455	0.115
6	1	44	7.035	5.706	8.259	0.517	0.088
1	2	126	6.574	5.599	8.371	0.573	0.104
2	2	524	7.162	5.604	8.619	0.450	0.129
3	2	322	7.026	5.611	8.587	0.500	0.111
4	2	230	7.048	5.661	8.604	0.496	0.121
5	2	144	6.993	5.699	8.615	0.501	0.113
6	2	71	7.170	5.819	8.499	0.508	0.103
1	3	85	6.540	5.697	8.480	0.588	0.094
2	3	309	6.983	5.682	8.595	0.493	0.113
3	3	233	6.928	5.650	8.526	0.524	0.109
4	3	160	6.808	5.602	8.562	0.521	0.108
5	3	116	6.965	5.633	8.550	0.522	0.104
6	3	78	7.140	5.994	8.581	0.521	0.104
1	4	61	6.478	5.627	7.625	0.616	0.071
2	4	144	6.823	5.613	8.594	0.543	0.099
3	4	138	6.790	5.600	8.458	0.561	0.093
4	4	104	6.887	5.686	8.268	0.558	0.090
5	4	66	6.815	5.628	8.030	0.547	0.104
6	4	45	6.969	5.741	8.271	0.535	0.084
1	5	14	6.540	5.880	7.390	0.711	0.073
2	5	65	6.700	5.610	8.470	0.584	0.075
3	5	67	6.707	5.607	8.378	0.564	0.105
4	5	66	6.959	5.723	8.422	0.568	0.093
5	5	43	6.807	5.592	7.993	0.554	0.105
6	5	32	7.042	5.929	8.197	0.606	0.075
ALL FAMILIES:		7358	6.954	5.591	8.622	0.470	0.125