A Semiparametric Approach to Estimate Engel curves using the US Micro Data

By

Anil K. Sulgham and Hector O. Zapata

Author Affiliations: Anil K. Sulgham is a Graduate Research Assistant, and Hector O. Zapata is a Professor in the Department of Agricultural Economics and Agribusiness, Louisiana State University, Baton Rouge, Louisiana.

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meetings Long Beach, California, July 23-26, 2006

Contact:
Anil K. Sulgham
Department of Agricultural Economics and Agribusiness
101 Agricultural Administration Building
Louisiana State University
Baton Rouge, LA 70803-5604
Email: asulgh1@lsu.edu

Hector O. Zapata
Department of Agricultural Economics and Agribusiness
101 Agricultural Administration Building
Louisiana State University
Baton Rouge, LA 70803-5604
Email: hzapata@agctr.lsu.edu

A Semiparametric Approach to Estimate Engel curves using the US Micro Data

Introduction

Engel curves explain the change of expenditure for different goods as a function of income (or total expenditure). Ernst Engel (1857) made the first attempt to investigate Engel curves; he studied how household expenditures on food vary with income. He found that food expenditures are an increasing function of income and of family size, but that food budget shares decrease with income. The study adopted a nonparametric approach to construct curves which are currently called regressograms. Since then much of the work on Engel curves involved use of parametric models.

Cross-section Engel functions have been a major subject of research in applied demand analysis. The use of cross-sections from survey data in the estimation of demand systems simplifies demand analysis as it assumes constant prices (all households face same prices). The relationship between consumption, income and prices is transformed into the well known consumption income relationship.

Classical approaches to estimating cross-sectional Engel curves are based on parametric models. Working (1943) proposed the log-linear budget share specification, which is known as the Working-Leser model, since Leser (1963) found that this functional form fit better than some alternatives. The most common demand system specifications such as AIDS, translog, linear expenditure, PIGL, and PIGLOG have been favored because of their exact aggregation or representative agent properties, but model misspecification is a recurrent theme. Recent studies have focused on Engel curves because these are more curvature flexible than the PIGLOG specification, for example Hausman et al., (1991); Hausman et al., (1995); Lewbel, (1991); Blundell and Duncan, (1998); and Blundell et al., (1993) find quadratic terms are needed in the model. In any parametric model the functional shape of the estimate is already given by assumptions. The quality of the resulting estimator depends heavily on the correctness of this specification. If the model is misspecified, then inferences and forecasts from such models are inadequate.

The assumption of a fixed parametric functional form embedded in parametric methods is relaxed in nonparametric models; consequently, there are no parameters to estimate. But parametric properties can be added in a semiparametric framework to arrive to a more economic theory consistent specification. In the context of Engel curve estimation nonparametric smoothing methods have been applied in a few studies, for example, in Bierens and Pott-Butler (1990), Banks et al., (1997), Blundell et al., (1998).

The basic relationship represented by an Engel curve is that of consumption and income. However, the consumption patterns of households also respond to demographic characteristics. For example, it is reasonable to expect a family with two children to spend more on food than a family with one child. Knowledge of the way income effects differ across household types is critical in understanding the impact of tax and welfare programs on expenditure patterns (Blundell et al., 1998). In the context of Engel curves most empirical studies allow demographic and other household characteristics to enter parametrically resulting in semiparametric specification.

Blundell et al. (1998) estimated Engel curves using a partial linear framework developed by Robinson (1988). The study found that under partial linear framework if at least one good has shares that are linear in $\ln x$, for example food shares, then introducing demographics restricts all demands to have shares linear in $\ln x$. To overcome this restriction they propose the extended partial linear model which allows the demographics to enter the model in an additive way and at the same time allowing for shape invariance of Engel curves. Blundell et al. (1998), investigate expenditure shares of couples with one child that are supposed to be related by parametric transformations to the expenditure shares of couples with two children using the UK data by applying the concept of shape invariance.

The objective of this paper is to estimate Engel curves with an emphasis on modeling the demographic characteristics using cross-section data from the 2003 US consumer expenditure survey (CES). The study focuses on finding adequate specification for the US Engel curves estimation using parametric, nonparametric, and semiparametric techniques. Nonparametric kernel regression techniques are used to estimate Engel curves. A semiparametric approach is used to pool across groups mainly households with one child as reference group and households with two children as the non reference group. We also estimate popular parametric specifications to consider as the null against the nonparametric and semiparametric alternatives. Model specification tests are conducted using the recently developed nonparametric specification test (Ellison and Ellison, 2000).

The rest of the paper is organized as follows. Section two provides a description of the empirical model followed by the data and methodology section. Results from the empirical analysis are discussed in section four. Finally summary and conclusions are presented.

Empirical Model

Working (1943) proposed the linear budget share specification, which is known as the Working-Leser model, since Leser (1963) found that this functional form fit better than some alternatives. We use the basic framework of Working-Leser model which is specified as follows:

$$w_j = a_j + b_j \ln x \tag{1}$$

The model underlies the popular Almost Ideal and Translog demand models of Deaton and Muellbauer (1980a) and Jorgenson, Lau, and Stoker (1980) which also have the Piglog specification in which budget shares are linear in log total expenditure. In our empirical analysis the parametric models use the above log linear model and the quadratic model includes extra squared term of log expenditure.

Blundell et al., (1998) extended partial linear Working-Leser framework which they named “extended partial linear model”. The study estimates Engel curves by using shape invariance parameter to pool across demographic groups. The shape invariance parameter which is also referred to as equivalence scales is estimated in the semiparametric framework by minimizing a loss function. The correction for endogeneity of total expenditure is also introduced into the model. In our study we use extended partial linear framework. The formal representation of extended partial linear model is given as

$$w_j = \alpha_j z + g_j(\ln x - \phi z) + v\rho_j + \varepsilon_j \tag{2}$$

where w_j is the expenditure share of the j^{th} good and x the total household expenditure, z represents the demographic variable and v are the residuals obtained from the augmented regression of $\ln x$ on log income of household.

To adjust for endogeneity the popular augmented regression technique suggested by Holly and Sargan (1982) is adopted in the semiparametric framework. Suppose *ln x* is endogenous in model in the sense that $E(\varepsilon \mid \ln x) \neq 0$. To correct for endogeneity we use the two step procedure. The first step involves a regression of log of total expenditure on log of income of the household; the residuals obtained from this step are used in the second stage.

$$\ln x = y\beta + v \text{, with } E(v|y) = 0 \quad (3)$$

The second step partial linear model is specified using the residuals from the first step as the excluded instrumental variable.

$$w_j = g_j(\ln x) + v\rho_j + \varepsilon_j \quad (4)$$

The significance ρ_j parameter confirms the endogeneity while correcting the same.

Data and Methodology

We use cross-sectional data derived from 2003 Consumer Expenditure Survey (CES) conducted by U.S. Department of Labor (Bureau of Labor Statistics). The data selected include observations on households with one or two children under the age of eighteen. In order to preserve demographic homogeneity in all other aspects we include only married couples. This leaves us with a sample size of 682 including 371 couples with one child and 311 couples with two children. Table 1 gives descriptive statistics of the sample used in the study. In our application we consider six broad categories of goods: food, clothing, alcohol, transportation, recreation, and other goods.

The linear and quadratic models are estimated using simple ordinary least squares (OLS) and used as null against alternatives nonparametric and semiparametric specifications. We use the Nadaraya-Watson estimator for estimating the nonparametric regression. The idea of nonparametric regression is to let the data determine the shape of the function to be estimated. Given an underlying data generating function, $Y_i = g(x_i) + \varepsilon_i$ ($\varepsilon \sim iid(0)$), an estimated nonparametric regression curve, $\hat{g}(x_i)$, may be defined over the data points $\{Y_i\, x_i\}$ as follows

$$\hat{g}(x_i) = \frac{\hat{r}(x)}{\hat{f}(x)} \tag{5}$$

where $$\hat{r}(x) = \frac{1}{N}\sum_{i=1}^{n} Y_i K\left(\frac{X - x_i}{h}\right)$$

and $$\hat{f}(x) = \frac{1}{N}\sum_{i=1}^{n} K\left(\frac{X - x_i}{h}\right) \tag{6}$$

K(.) is a weakly positive Kernel function, and *h* is a nonzero bandwidth. The choice of the kernel function, *K(.)*, has a minimal affect on the estimate of $\hat{g}(x_i)$ (see Hardle, 1993), so we choose the triangle kernel for *K(.)*. Cross-validation technique is employed to select the optimal bandwidth for each estimated nonparametric regression. The top and bottom 2.5% of the data are trimmed to avoid the boundary bias in the estimation of nonparametric regression.

The semiparametric approach involves partial linear specification for each of the budget share equations

$$w_{ij} = \alpha_j^{`} z_i + g_j(\ln x_i) + \varepsilon_{ij} \tag{7}$$

in which $\alpha_j^{`}$z represents a linear index in terms of a vector of observable exogenous regressors z_i and unknown parameters α_j. Here we will assume $E(\varepsilon_{ij} \mid z, \ln x) = 0$ and $Var(\varepsilon_{ij} \mid z, \ln x) = \sigma_j^2(z, \ln x)$. Following Robinson (1988), a simple transformation of the model can be used to give an estimator for α_j. Taking expectations of the above conditional on *ln x*, and subtracting from the resulting expression from (7) yields

$$w_{ij} - E(w_{ij} \mid \ln x_i) = \alpha_j^{`}(z_i - E(z_i \mid \ln x_i)) + \varepsilon_{ij} \tag{8}$$

The terms $E(w_{ij} \mid \ln x_i)$ and $E(z_i \mid \ln x_i)$ can be replaced by $\hat{g}_{jh}^{w}(\ln x)$ and $\hat{g}_{h}^{z}(\ln x)$ nonparametric fitted values respectively. A simple OLS estimator is then applied to the residuals obtained to get α_j. To adjust for endogeneity the popular augmented regression technique suggested by Holly and Sargan (1982) is used in the semiparametric framework.

The shape invariance parameter is estimated using a pooling approach. Under the pooling approach nonparametric regressions for the two demographic groups (families with one child and two children) are assumed to be linked by the parameters $(\phi, \{\alpha_j\})$ and follow the shape-invariance specification given by:

$$g_j^1(x) = \alpha_j + g_j^0(x - \phi). \tag{9}$$

Hardle and Marron (1990) suggested a simple loss function equal to the integrated squared distance between the reference function and the transformed function to estimate the parameters. But the approach is valid for fixed design models. The budget data used in the study are suited to random design models because the independent variable, total expenditure, is a random variable. We use a slightly more complex function suggested by Pinkse and Robinson (1995) for the random design case. The definition of this estimator is essentially based on the Nadaraya-Watson estimator.

$$L(\phi, \alpha_j) = \sum_{j=1}^{n} \int_{xlow}^{xhigh} \hat{f}^{1}(x)\hat{f}^{0}(x - \phi)(\hat{g}_j^{1}(x) - \hat{g}_j^{0}(x - \phi) - \alpha_j)^2 w(x)dx \tag{10}$$

Where $\hat{f}^{1}$, and $\hat{f}^{0}$, are the kernel density estimates, $\hat{g}_j^{1}$, and $\hat{g}_j^{0}$ are the fitted values of the nonparametric regressions, and *w(x)* is equation specific non-negative weight function. Estimation of scale parameter ϕ and the shift parameters α_j involves sequential gridsearch methods. Robinson's (1988) method is applied to estimate the initial values for α_j. Conditional on α_j a sequential gridsearch method is applied to the loss function to estimate the scale parameter ϕ. The standard errors obtained should be interpreted with caution as the distribution of loss function is likely non-normal (Pendakur, 2005). Hence we construct bootstrap standard errors for ϕ and α_j through repetition of this gridsearch process for 500 bootstrap samples.

We use the "wild" bootstrap procedure because of the potential heteroscedastic setting. In this case, for each estimated residual ε_i one creates two-point distribution for a random variable, say $\upsilon_i = \varepsilon_i(1-\sqrt{5})/2$, with probability (prob(υ_i) =(5+$\sqrt{5}$)/10) and υ_i = $\varepsilon_i(1+\sqrt{5})/2$, with probability (prob(5-$\sqrt{5}$)/10) (see Yatchew, 2003). The random variable υ_i has properties $E(\upsilon_i) = 0$, $E(\upsilon_i^2) = \hat{\varepsilon}_i^2$, and $E(\upsilon_i^3) = \hat{\varepsilon}_i^3$. Based on this distribution we draw ε_i^B and build the bootstrap data set. The data set is then used for further analysis in the estimation of standard error and the empirical distribution of the loss function.

The empirical distribution of the loss function is developed under the assumption of extended partially linear null. The test for shape invariance is conducted using the bootstrap critical values for the loss function generated from the bootstrap samples (similar to Pendakur, 1999). Finally specification tests are conducted with parametric models as null hypothesis and nonparametric and semiparametric specification as the alternatives. We use the recently developed specification test by Ellison and Ellison (2000) which looks at the orthogonality between nonparametric fitted values and parametric residuals.

$$1/n\sum_{i=1}^{n}\hat{g}(x_i)[y_i-\hat{f}(x_i,b)] \qquad (11)$$

The test simplifies into a weighted function of the sum of squares of parametric residuals, and can be linked to the test proposed by Zheng (1996).

Empirical Results

The results from kernel regression are presented in Figures 1 to 6 for the six budget shares considered in our study. The figures present unrestricted non-parametric Engel curves for the demographic group (couples with one child) which is also our reference group and the second group (couples with two children). The triangle kernel is used for all the regressions, applying leave-one-out cross-validation methods for bandwidth selection to each non-parametric regression.

The regression curves appear to demonstrate that the Working-Leser linear logarithmic (Piglog) formulation is a reasonable choice for some budget share curves (for example, food, clothing, and alcohol). For other shares, in particular transportation and other goods, a more non-linear relationship between share and log expenditure is evident.

In Figure 1, for example, we see a broadly parallel shift in the food Engel curve, with couples with two children spending more of their budget on food than couples with a single child at the same level of total expenditure. For alcohol, on the other hand, Engel curves for couples with two children shift down relative to the reference group (see Figure 2). There is no strong evidence of demographic variability in clothing, transportation, and other good shares. The similarity in the shapes across demographic characteristic suggests existence of shape invariance.

The parametric and semiparametric estimates are reported in Tables 2-7 for each of the six share equations. The first column in each of the tables presents results for a simple OLS regression of budget share on log expenditure with no semiparametric controls. The second column reports results for a model which adjusts for the number of children in the household using the partially linear framework of Robinson (1988). The estimates of the model that controls for demographics and endogeneity using the partial linear framework are presented in third column. The final two specifications relate to the shape-invariant generalizations proposed by Blundell et al., (1998). The fourth model allows for scale shifts in log expenditure by demographic type using the estimation method of Pinkse and Robinson (1995) and the fifth in addition introduces controls for endogeneity.

In these tables $\hat{\beta}_j^{ols}$ refers to the simple OLS estimate of the slope coefficient. The shape-invariant transformation has two parameters for each share equation; the scaling parameter ϕ in the term $(\ln x - \phi z)$ and an intercept parameter α_j. For the two models in column four and five in each of Tables 2 to Table 7 we estimate the parameters $(\phi, \{\alpha_j\})$ through minimization of loss function presented in equation (10).

The estimate of scale parameter common to all six share equations is 0.258, giving an estimated equivalence scale of (1.294) for couples with two children compared with reference group (couple with one child). This is quite close to the estimates reported by the U.S. House of Representatives (1994), Blundell et al (1998) for UK data and

Pendakur (1999) for Canada data. The parameters α_j specific to each share equation are reported in the tables. Having accounted for the scale parameter ϕ, we find significant shift parameter for food, confirming the initial graphical evidence in Figure 1. We fail to reject the null hypothesis of shape invariance for all share equations based on the critical values generated from the empirical distribution of the loss function. The critical values are obtained from the distribution of loss function generated under the null hypothesis of shape invariance and then compared to the value loss function for each share equation.

The results from tests of the linear and quadratic logarithmic specifications against the nonparametric and semiparametric alternatives are reported in bottom two rows of Tables 2-7. For majority of the expenditure shares (food, alcohol, clothing, recreation, and other shares) in Tables 2-7, in all specifications, we are unable to reject linearity. In contrast, for transportation share, the Piglog of Working-Leser form, and quadratic specifications are strongly rejected. This result is maintained even after controlling for demographic variation and the endogeneity of total expenditure. We also find the correction for endogeneity of log total expenditure to be important in most share equations, most notably food, transportation and other share.

Summary and Conclusions

The objective of this paper was to investigate the shape of Engel curves using theoretically consistent and data coherent methods. By choosing a cross section of US consumers we have focused on the Engel curve relationship. As a baseline specification we have worked with the Working-Leser or Piglog specification in which budget shares are expressed in terms of log total expenditure, this being the Engel curve shape underlying the popular AIDS and Translog demand models.

We also consider parametric models which have more variety of curvature than is permitted by the Piglog. The evidence of quadratic curvature from more recent studies (Banks et al., 1997) warrants the need to investigate quadratic specification. Consequently we used both the Piglog and quadratic logarithmic specifications as null parametric specifications against non-parametric and semiparametric alternatives.

The shape invariant semiparametric framework proposed by Blundell et al., (1998) is estimated using the pooling approach suggested by Pinkse and Robinson (1995). To correct for endogeneity Holly and Sargan (1982) augmented regression approach is adopted in the semiparametric framework. Income is used to instrument total expenditure, correcting for endogeneity is found to have an important impact on the curvature of the Engel curve relationship. We fail to reject the assumption of shape invariance based on the critical values generated from the empirical distribution of loss function.

Ellison and Ellison (2000) specification test is used to compare the parametric specification against the nonparametric and semiparametric alternatives. The results from specification tests indicate Working-Leser or Piglog specification was sufficient for most budget shares except for transportation where semiparametric specification had support.

Our study further contributes to the growing evidence on the need for investigating specification of model in the empirical analysis of Engel curves. The parametric models as evidenced above might not be adequate for all goods. The results seem to support the findings from earlier studies using data from other developed nations for example Blundell et al., (1998) for UK data and Pendakur (1999) for Canada data.

The relatively close estimate for the equivalence scales of developed nations indicates similarities in the consumption patterns.

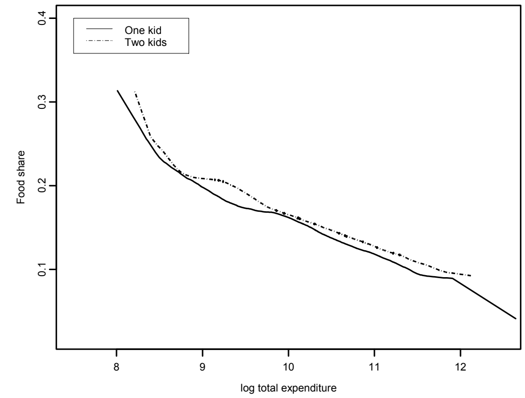

Figure 1. Food Engel curves

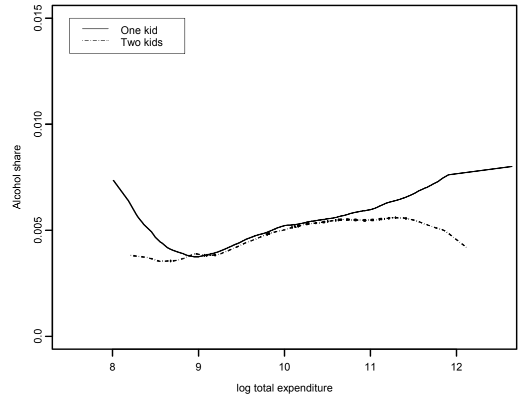

Figure 2. Alcohol Engel curves

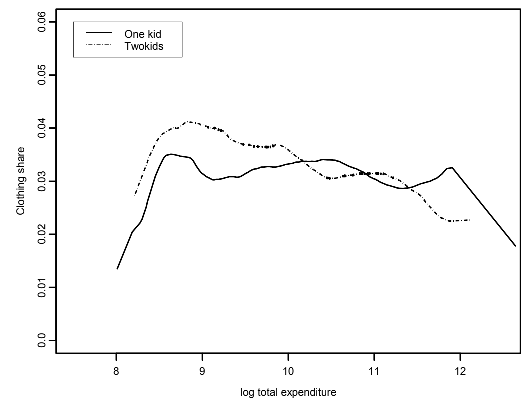

Figure 3. Clothing Engel curves

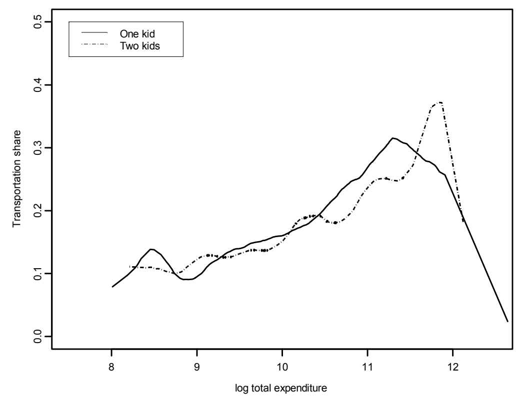

Figure 4. Transportation Engel curves

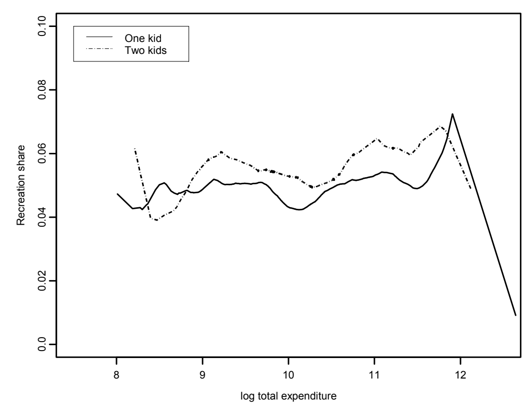

Figure 5. Recreation Engel curves

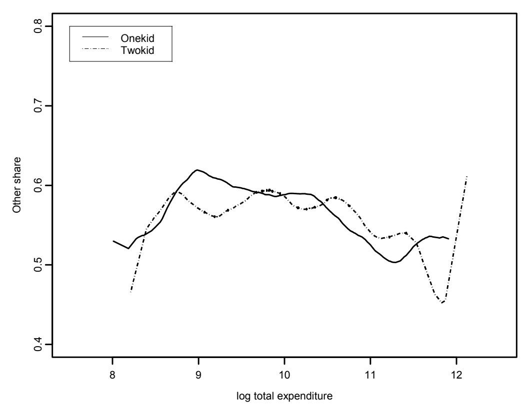

Figure 6. Other Engel curves

Table 1. Descriptive statistics for budget share data

	Couple with one child		Couple with two children	
Variable	Means	Std. deviations	Means	Std. deviations
Food share	0.153	0.073	0.160	0.066
Alcohol share	0.005	0.008	0.004	0.007
Clothing share	0.031	0.025	0.033	0.029
Transportation share	0.187	0.150	0.176	0.150
Recreation share	0.049	0.040	0.055	0.051
Other share	0.573	0.134	0.568	0.137
Log total expenditure	10.150	0.827	10.195	0.757
Log total income	10. 551	1.115	10.687	1.017
Sample size	371		311	

Table 2. Parametric and semiparametric estimates: food Engel curves

	$\phi = 0$			$\phi = 0.258\,(0.047)$	
	1 No corrections	2 Demographics	3 Demographics and endogeneity	4 Demographics	5 Demographics and endogeneity
$\hat{\beta}_j^{ols}$	-0.0464	-0.0466	-0.0464	-0.0465	-0.0461
	(0.0029)	(0.0029)	-(0.0028)	(0.0029)	(0.0028)
$\hat{\alpha}_j$		0.0104	0.0103	0.0222	0.0211
		(0.0045)	(0.0045)	(0.0046)	(0.0045)
$\hat{\rho}_j$			0.0170		0.0444
			(0.0025)		(0.0025)
Loss				0.0002	
				[0.0006]	
H_o: linear parametric form					
$\chi^2_{EE}(1)$	0.082	0.912	0.362	0.174	0.333
Ho: quadratic parametric form					
$\chi^2_{EE}(1)$	0.023	0.939	0.127	0.026	0.049

Notes: $\hat{\beta}_j^{ols}$ is the estimate obtained from OLS regression. Standard errors are enclosed in () parentheses. Bootstrapped 10 % critical values in [ ] parentheses. Bootstrap standard error is reported for the parameter ϕ.

Table 3. Parametric and semiparametric estimates: alcohol Engel curves

	$\phi = 0$			$\phi = 0.258$ (0.047)	
	1 No corrections	2 Demographics	3 Demographics and endogeneity	4 Demographics	5 Demographics and endogeneity
$\hat{\beta}_j^{ols}$	-0.0009	-0.0009	-0.0012	-0.0009	-0.0009
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
$\hat{\alpha}_j$		-0.0005	-0.0006	-0.0003	-0.0005
		(0.0006)	(0.0006)	(0.0006)	(0.0006)
$\hat{\rho}_j$			-0.0018		-0.0020
			(0.0014)		(0.0009)
Loss				8.0E-5	
				[0.0001]	
Ho: linear parametric form					
$\chi^2_{EE}(1)$	0.006	0.005	0.073	0.133	0.203
Ho: quadratic parametric form					
$\chi^2_{EE}(1)$	0.038	0.037	0.090	0.154	0.175

Notes: $\hat{\beta}_j^{ols}$ is the estimate obtained from OLS regression. Standard errors are enclosed in () parentheses. Bootstrapped 10 % critical values in [ ] parentheses. Bootstrap standard error is reported for the parameter ϕ.

Table 4. Parametric and semiparametric estimates: clothing Engel curves

	$\phi = 0$			$\phi = 0.258$ (0.047)	
	1 No corrections	2 Demographics	3 Demographics and endogeneity	4 Demographics	5 Demographics and endogeneity
$\hat{\beta}_j^{ols}$	-0.0019	-0.0019	-0.0018	-0.0019	-0.0019
	(0.0013)	(0.0013)	(0.0015)	(0.0013)	(0.0013)
$\hat{\alpha}_j$		0.0017	0.0017	0.0021	0.0020
		(0.0021)	(0.0021)	(0.0022)	(0.0021)
$\hat{\rho}_j$			0.0006		0.0036
			(0.0044)		(0.0032)
Loss				7.0E-5	
				[0.0002]	
Ho: linear parametric form					
$\chi^2_{EE}(1)$	0.699	0.716	0.647	0.819	0.776
Ho: quadratic parametric form					
$\chi^2_{EE}(1)$	0.833	0.835	0.831	0.866	0.867

Notes: $\hat{\beta}_j^{ols}$ is the estimate obtained from OLS regression. Standard errors are enclosed in () parentheses. Bootstrapped 10 % critical values in [ ] parentheses. Bootstrap standard error is reported for the parameter ϕ.

Table 5. Parametric and semiparametric estimates: transportation Engel curves

	$\phi = 0$			$\phi = 0.258$ (0.047)	
	1 No corrections	2 Demographics	3 Demographics and endogeneity	4 Demographics	5 Demographics and endogeneity
$\hat{\beta}_j^{ols}$	0.0657	0.0659	0.0619	0.0659	0.0662
	(0.0068)	(0.0068)	(0.0079)	(0.0068)	(0.0068)
$\hat{\alpha}_j$		-0.0136	-0.0132	-0.0324	-0.0332
		(0.0108)	(0.0108)	(0.0109)	(0.0109)
$\hat{\rho}_j$			0.0100		0.0273
			(0.0061)		(0.0161)
Loss				0.0014	
				[0.0023]	
Ho: linear parametric form					
$\chi^2_{EE}(1)$	12.814	12.460	21.174	11.356	6.136
Ho: quadratic parametric form					
$\chi^2_{EE}(1)$	9.424	9.492	19.934	6.553	6.150

Notes: $\hat{\beta}_j^{ols}$ is the estimate obtained from OLS regression. Standard errors are enclosed in () parentheses. Bootstrapped 10 % critical values in [ ] parentheses. Bootstrap standard error is reported for the parameter ϕ.

Table 6. Parametric and semiparametric estimates: recreation Engel curves

	$\phi = 0$			$\phi = 0.258$ (0.047)	
	1 No corrections	2 Demographics	3 Demographics and endogeneity	4 Demographics	5 Demographics and endogeneity
$\hat{\beta}_j^{ols}$	0.0026	0.0025	0.0092	0.0025	0.0049
	(0.0022)	(0.0022)	(0.0024)	(0.0022)	(0.0022)
$\hat{\alpha}_j$		0.0064	0.0064	0.0056	0.0056
		(0.0035)	(0.0035)	(0.0035)	(0.0036)
$\hat{\rho}_j$			-0.0136		-0.0035
			(0.0074)		(0.0053)
Loss				0.0014	
				[0.0031]	
Ho: linear parametric form					
$\chi^2_{EE}(1)$	0.070	0.151	0.203	0.042	0.066
Ho: quadratic parametric form					
$\chi^2_{EE}(1)$	0.104	0.066	0.066	0.060	0.061

Notes: $\hat{\beta}_j^{ols}$ is the estimate obtained from OLS regression. Standard errors are enclosed in () parentheses. Bootstrapped 10 % critical values in [ ] parentheses. Bootstrap standard error is reported for the parameter ϕ.

Table 7. Parametric and semiparametric estimates: other Engel curves

	$\phi = 0$			$\phi = 0.258$ (0.047)	
	1 No corrections	2 Demographics	3 Demographics and endogeneity	4 Demographics	5 Demographics and endogeneity
$\hat{\beta}_j^{ols}$	-0.210	-0.0209	-0.0193	0.0209	-0.0216
	(0.0065)	(0.0065)	(0.0072)	(0.0065)	(0.0064)
$\hat{\alpha}_j$		-0.0042	-0.0045	0.0031	0.0049
		(0.0102)	(0.0102)	(0.0103)	(0.0102)
$\hat{\rho}_j$			-0.0014		-0.0686
			(0.0214)		(0.0145)
Loss				0.0016	
				[0.0043]	
Ho: linear parametric form					
$\chi^2_{EE}(1)$	1.359	1.366	0.756	2.468	1.844
Ho: quadratic parametric form					
$\chi^2_{EE}(1)$	0.623	0.630	0.513	0.574	0.529

Notes: $\hat{\beta}_j^{ols}$ is the estimate obtained from OLS regression. Standard errors are enclosed in () parentheses. Bootstrapped 10 % critical values in [ ] parentheses. Bootstrap standard error is reported for the parameter ϕ.

References

Banks, J., R. Blundell, and A. Lewbel. "Quadratic Engel Curves and Consumer Demand." Review of Economics and Statistics 79(1997):527-539.

Bierens, H. and H. A. Pott-Buter. "Specification of Household Expenditure Functions and Equivalence Scales by Nonparametric Regression." Econometric Reviews 9(1990):123-210.

Blackorby, C. and D. Donaldson. "Adult equivalence scales and the economic implementation of interpersonal comparisons of well-being." Social Choice and Welfare 10(1993) 333-361.

Blundell R, Duncan A, Pendakur K. "Semiparametric estimation of consumer demand." Journal of Applied Econometrics 13(1998): 435-461.

Blundell R, Pashardes P, Weber G. "What do we learn about consumer demand patterns from micro data?" American Economic Review 83(1993): 570-597.

Deaton, A. S., and Muellbauer, J. "An Almost Ideal Demand System." American Economic Review 70(1980): 312-336.

Ellison, G., and Ellison, S.F., "A simple framework for nonparametric specification testing." Journal of Econometrics 96(2000):1-23.

Engel, E. "Die Productions- und Consumptions verhaeltnisse des Koenigsreichs Sachsen." reprinted with Engel Anlage I(1895):1-54.

Gozalo, P. "Nonparametric bootstrap analysis with applications to demographic effects in demand functions." Journal of Econometrics 81(1997):357-393.

Hardle W. "Applied Nonparametric Regression. Cambridge University Press: New York(1993).

Hausman, J., W. Newey, H. Ichimura and J. Powell. "Identification and estimation of polynomial errors in variables models." Journal of Econometrics 50(1991):273-296.

Hausman, J.A., W. K. Newey, and J. L. Powell. "Nonlinear Errors in Variables: Estimation of Some Engel Curves." Journal of Econometrics 65(1995):205- 253.

Hardle, W. and J. Marron. "Semiparametric comparison of regression curves." Annals of Statistics 18 (1990):63-89.

Holly, A., and J. Sargan. "Testing for Exogeneity in a Limited Information Framework." Cahiers de Recherches Economiques, No, 8204. Universite dc Lausanne 1982.

Jorgenson. D,, L. Christensen, and L. Lau. "Transcendental Logarithmic Utility Functions." American Economic Review 65(1975):367-383.

Leser, C. "Forms of Engel functions." Econometrica 31(1963):694-703.

Lewbel, A. "The Rank of Demand Systems: Theory and Nonparametric Estimation." Econometrica 59(1991):711-730.

Nadaraya, E. "On Estimating Regression." Theory of Probability and its Applications 9(1964):141-142.

Pendakur, K. "Estimates and Tests of Base-Independent Equivalence Scales." Journal of Econometrics 88(1999):1-40.

Pinkse, C. and P. Robinson. "Pooling nonparametric estimates of regression functions with a similar shape." in G. Maddala, P. Phillips and T. N. Srinivasan (eds), Advances in Econometrics and Quantitative Economics (1995):172-195.

Pollak, R. A., and Wales, T. J. "Estimation of Complete Demand Systems from Household Budget Data." American Economic Review 68(1978):348-359.

Robinson, P. "Root-N-consistent semiparametric regression." Econometrica 56(1988): 931-954.

U.S. House of Representatives, Committee on Ways and Means. Where Your Money Goes. The 1994-1995 Green Book, Washington D.C, U.S. Government Printers (1994).

Watson, G. "Smooth Regression Analysis." Sankhya 26(1964):359-372.

Working, H. "Statistical Laws of Family Expenditures." Journal of the American Statistical Association 38(1943):43-56.

Yatchew, A. "Semiparametric Regression for the Applied Econometrician." Themes in Modern Econometrics, ed. P.C.B. Phillips, Cambridge University Press (2003).

Zheng, J. "A consistent test of functional form via nonparametric estimation techniques." Journal of Econometrics 75(1996):263-289.