

# Tricks With Hicks: The EASI Demand System

Arthur Lewbel and Krishna Pendakur  
Boston College and Simon Fraser University

Revised May 2009

## Abstract

We invent Implicit Marshallian Demands, a new type of demand function that combines desirable features of Hicksian and Marshallian demand functions. We propose and estimate the Exact Affine Stone Index (EASI) Implicit Marshallian Demand system. Like the Almost Ideal Demand (AID) system, EASI budget shares are linear in parameters given real expenditures. However, unlike the AID, EASI demands can have any rank and its Engel curves can be polynomials or splines of any order in real expenditures. EASI error terms equal random utility parameters to account for unobserved preference heterogeneity. EASI demand functions can be estimated using GMM or three stage least squares, and, like AID, an approximate EASI model can be estimated by linear regression.

JEL Codes: D11, D12, C31, C33, C51

Keywords: Consumer demand, Demand systems, Hicks, Marshallian, Cost functions, Expenditure functions, Utility, Engel curves.

The authors wish to thank Richard Blundell, Angus Deaton, David Donaldson, Stefan Sperlich, and many anonymous referees for helpful comments. Any errors are our own.

Forthcoming, American Economic Review

# 1 Introduction

Recent empirical work with large consumer expenditure data sets finds Engel curves (income expansion paths) with significant curvature and variation across goods. For example, some goods have Engel curves that are close to linear or quadratic, while others are more S-shaped (see, Blundell, Chen and Kristensen (2007)). Typical parametric demand models cannot encompass this variety of shapes, and are constrained by Gorman (1981) type rank restrictions.

Other current research shows the importance of allowing for unobserved preference heterogeneity in demand systems, and the difficulty of doing so in a coherent fashion. In most empirical models of consumer demand, model error terms cannot be interpreted as random utility parameters representing unobserved heterogeneity. See, for example, Brown and Walker (1989), McFadden and Richter (1990), Brown and Matzkin (1998), Lewbel (2001), and Beckert and Blundell (2004).

Despite these empirical issues, Deaton and Muellbauer's (1980) Almost Ideal Demand (AID) model, which has linear Engel curves and does not incorporate unobserved heterogeneity, remains very popular. This popularity is at least partly because alternative models involve nonlinear functions of many prices and parameters, which are often numerically difficult or intractable to implement. In addition, the AID model has a very convenient approximate form which may be estimated by linear methods.

In this paper, we develop an approach to the specification and estimation of consumer demands that addresses the above issues while maintaining the simplicity of the AID model. Consider a consumer with demographic (and other observable preference related) characteristics  $\mathbf{z}$  and log nominal total expenditures  $x$  that faces the  $J$ -vector of log prices  $\mathbf{p}$ . Assume she chooses a bundle of goods, described by the  $J$ -vector of budget shares  $\mathbf{w}$ , to maximize utility given her linear budget constraint. Hicksian demand functions associated with her utility function, which express  $\mathbf{w}$  as a function of  $\mathbf{p}$ ,  $\mathbf{z}$ , and attained utility level  $u$ , can easily be specified to have many desirable properties. We show that under some conditions, which permit both random utility parameters and arbitrarily complicated Engel curves, utility  $u$  can be expressed as a simple function of the observed variables  $x$ ,  $\mathbf{w}$ ,  $\mathbf{p}$ , and  $\mathbf{z}$ . This function, which we denote  $y$ , can often be interpreted as a measure of log real expenditures. We use these results to directly specify and estimate what we call *implicit Marshallian demands*, which are Hicksian demands after replacing  $u$  with the implicit utility measure  $y$ .

Noting that  $\mathbf{p}'\mathbf{w}$  is the definition of the Stone (1954) log price index, we define the *Exact Affine Stone Index* (EASI) class of cost functions, which have  $y$  equal to an affine transform of Stone index deflated log nominal expenditures,  $x - \mathbf{p}'\mathbf{w}$ . The resulting EASI implicit Marshallian demand functions have the following properties:

1. Like the AID system, EASI budget share demand functions are, apart from the construction

of  $y$ , completely linear in parameters, which facilitates estimation in models with many goods.

2. The AID budget shares are linear in  $\mathbf{p}$ ,  $\mathbf{z}$ , and  $y$ . In contrast, EASI budget shares are linear in  $\mathbf{p}$  and are polynomials of any order in  $\mathbf{z}$  and  $y$ . They can also include interaction terms such as  $\mathbf{p}y$ ,  $\mathbf{z}y$  and  $\mathbf{p}\mathbf{z}'$ , and contain other functions of  $\mathbf{z}$  and  $y$ .

3. EASI Engel curves for each good are almost completely unrestricted. For example, EASI demands can be high order polynomials or splines in  $y$  and  $\mathbf{z}$ , and so can encompass empirically important specifications that most parametric models cannot capture, such as the semiparametric S shaped Engel curves reported by Pendakur (1999) and Blundell, Chen and Kristensen (2007). The AID system is linear in  $y$  and has rank two, and the quadratic AID of Banks, Blundell, and Lewbel (1997) is quadratic in (an approximation of)  $y$  with rank three. In contrast, EASI demands can be polynomials or splines of any order in  $y$ , and can have any rank up to  $J-1$ , where  $J$  is the number of goods. EASI demands are not subject to the Gorman (1981) rank three limit even with polynomial Engel curves (see the Appendix for definitions and details regarding the meaning and nature of rank restrictions).

4. EASI budget share error terms can equal unobserved preference heterogeneity or random utility parameters. The AID and other similar models do not have this property, since in those models unobserved preference heterogeneity requires that additive errors be heteroskedastic (see Brown and Walker 1989 and Lewbel 2001). The EASI unobserved preference heterogeneity is coherent and invertible (see the Appendix).

5. EASI demand functions can be estimated using nonlinear instrumental variables, particularly nonlinear three-stage least-squares (3SLS) and Hansen's (1982) Generalized Method of Moments (GMM). Like the AID system, approximate versions of EASI demands can be estimated by linear regression. We find little empirical difference between exact nonlinear and approximate linear EASI estimates.

6. Since EASI demands are derived from a cost function model, given estimated parameters we have simple closed form expressions for consumer surplus calculations, such as cost-of-living indices for large price changes.

EASI demands thus accommodate an extremely wide class of functional forms. In the empirical work below, we implement a model where, for the budget share  $w_j$  of each good  $j$ , the estimating equations for demand systems have the linear in parameters form

$$w_j = \sum_{r=0}^5 b_{rj} y^r + \sum_{l=1}^L (C_{lj} z_l + D_{lj} z_l y) + \sum_{l=0}^L \sum_{k=1}^J A_{lkj} z_l p_k + \sum_{k=1}^J B_{kj} p_k y + \varepsilon_j,$$

where  $y$  is a measure of real total expenditures. The regressors in this model are a fifth order polynomial in  $y$ , log-prices  $p_k$  of each good  $k$ , and  $L$  different demographic characteristics  $z_l$ , in addition to interaction terms of the forms  $p_k y$ ,  $z_l y$ , and  $z_l p_k$ . Our approximate EASI model

estimates these equations for each good  $j$  by ordinary least squares, letting  $y$  equal  $x - \sum_{j=1}^J p_j w_j$ , while the exact model has  $y$  given by equation (8) below.

We begin with an overview of our approach, defining implicit Marshallian demand functions and general EASI demands. We show how these models accommodate high rank Engel curves and unobserved preference heterogeneity. We describe the specific EASI functional form we will use in our empirical application, and also show how to construct a fully linear approximation for our EASI demand model, analogous to the approximate linear AID model. We show how to apply the EASI model with consumer surplus and compensated elasticity calculations. Then we describe estimators for the EASI model, including consistent, asymptotically normal instrumental variable, three stage least squares, and GMM estimators for the exact model and linear least squares estimators for the approximate model.

We estimate the exact model using Canadian micro-data. We find more complicated Engel curve shapes than those of standard parametric demand systems, and we find that the simple approximate EASI model captures this complexity very well. We apply the model to a cost-of-living experiment, and find that both the increased flexibility of Engel curves and the incorporation of unobserved heterogeneity into the model significantly affect the resulting welfare calculations.

An appendix provides formal theorems and proofs, along with some extensions and additional mathematical properties of EASI and other implicit Marshallian demands, which are relevant for evaluating these models and for other possible applications of our general methodology.

## 2 Methods and Models

Here we define the general idea of implicit Marshallian demands, describe the EASI functional form, and explain how this model can be estimated and applied. Formal theorems regarding these methods and models are provided in the Appendix.

### 2.1 Implicit Marshallian Demand Functions

We specify a cost (expenditure) function and use Shephard's lemma to obtain Hicksian demands that have the desired properties. The usual next step would be to obtain Marshallian demands, which are functions of  $\mathbf{p}$ ,  $\mathbf{z}$  and  $x$ , by solving for indirect utility  $u$  in terms of  $\mathbf{p}$ ,  $\mathbf{z}$  and  $x$ , and substituting this into Hicksian demand functions. We instead construct cost functions that have simple expressions for  $u$  in terms of  $\mathbf{w}$ ,  $\mathbf{p}$ ,  $\mathbf{z}$  and  $x$ . We call this expression  $y$ , implicit utility, and substitute  $y$  for  $u$  in the Hicksian demands to yield what we call implicit Marshallian demand functions. These implicit Marshallian demands circumvent the difficulty of finding simple analytic expressions for indirect utility or Marshallian demands.

We wish to explicitly include both observable and unobservable sources of preference heterogeneity in our models, so let  $\mathbf{z} = (z_1, \dots, z_L)'$  be an  $L$ -vector of observable demographic (or other) characteristics that affect preferences, and let  $\boldsymbol{\varepsilon}$  be a  $J$ -vector of unobserved preference characteristics (taste parameters) satisfying  $\mathbf{1}'_J \boldsymbol{\varepsilon} = 0$  where  $\mathbf{1}_J$  is the  $J$ -vector of ones. Typical elements of  $\mathbf{z}$  would include household size, age, and composition. The log cost or expenditure function is  $x = C(\mathbf{p}, u, \mathbf{z}, \boldsymbol{\varepsilon})$ , which equals the minimum log-expenditure required for an individual with characteristics  $\mathbf{z}, \boldsymbol{\varepsilon}$  to attain utility level  $u$  when facing log prices  $\mathbf{p}$ . Shephard's Lemma expresses Hicksian (compensated) budget-share functions,  $\boldsymbol{\omega}(\mathbf{p}, u, \mathbf{z}, \boldsymbol{\varepsilon})$ , as  $\mathbf{w} = \boldsymbol{\omega}(\mathbf{p}, u, \mathbf{z}, \boldsymbol{\varepsilon}) = \nabla_{\mathbf{p}} C(\mathbf{p}, u, \mathbf{z}, \boldsymbol{\varepsilon})$ . Indirect utility,  $V$ , is the inverse of log-cost with respect to  $u$ :  $u = V(\mathbf{p}, x, \mathbf{z}, \boldsymbol{\varepsilon}) = C^{-1}(\mathbf{p}, \cdot, \mathbf{z}, \boldsymbol{\varepsilon})$ . If an analytic solution for  $V$  is unavailable, it may still be possible to express utility as a function  $g$  of  $\mathbf{w}, \mathbf{p}, x, \mathbf{z}$ . In this case, we can write  $u = g(\boldsymbol{\omega}(\mathbf{p}, u, \mathbf{z}, \boldsymbol{\varepsilon}), \mathbf{p}, x, \mathbf{z})$ , which implicitly defines  $u$ . We may then define *implicit utility*  $y$  by  $y = g(\mathbf{w}, \mathbf{p}, x, \mathbf{z})$ . By construction,  $y = u$  and the function  $g$  that defines  $y$  depends only on observable data, and may have a simple tractable closed form expression even when no closed form expression exists for indirect utility  $V$ . The *implicit Marshallian demand* system is then given by  $\mathbf{w} = \boldsymbol{\omega}(\mathbf{p}, y, \mathbf{z}, \boldsymbol{\varepsilon})$ , which is the Hicksian demand system except for replacing  $u$  with  $y$ .

A similar idea to implicit marshallian demands is Browning's (2001) 'M-demands,' which expresses demand functions in terms of prices and the quantity of one good, instead of in terms of prices and total expenditures. This can be interpreted as a restrictive choice of  $g$ , since it uses only one good.

In this construction, preferences and hence utility remain ordinal. So when we say that implicit utility  $y$  equals the utility level  $u$ , this only means that the observable  $y$  equals one possible cardinalization of  $u$ . In all our examples  $y$  will be linear in  $x$  and will either equal or closely approximate a money metric cardinalization of utility, and hence  $y$  will also be interpretable as a measure of real expenditures. We will sometimes refer to  $y$  as log real expenditures when this interpretation of implicit utility is particularly relevant.

Since  $g$  expresses utility as a function of all the arguments of indirect utility plus allowing for dependence  $\boldsymbol{\omega}$ , it admits more possibilities than are available for explicit marshallian demand systems. Although, in theory, the idea of implicit demand systems opens up an extremely large class of potential demand systems, we find that simple forms for  $g$  allow us to solve a large number of problems facing empirical demand analysis.

Just to illustrate the idea of implicit Marshallian demands, consider the following restrictive example. For simplicity (ignoring the inequality constraints required for demand system regularity) consider a log cost function of the form

$$C(\mathbf{p}, u, \mathbf{z}, \boldsymbol{\varepsilon}) = u + \mathbf{p}' \mathbf{m}(u, \mathbf{z}) + \mathbf{p}' \boldsymbol{\varepsilon} \quad (1)$$

where  $\mathbf{m}(u, \mathbf{z})$  is a  $J$ -vector valued function with  $\mathbf{1}'_J \mathbf{m}(u, \mathbf{z}) = 1$ . By Shephard's lemma, this cost function has Hicksian (compensated) budget shares  $\mathbf{w} = \omega(\mathbf{p}, u, \mathbf{z}, \varepsilon) = \mathbf{m}(u, \mathbf{z}) + \varepsilon$ . Since  $x = u + \mathbf{p}'[\mathbf{m}(u, \mathbf{z}) + \varepsilon]$ , we may substitute  $\mathbf{m}(u, \mathbf{z}) + \varepsilon$  with  $\mathbf{w}$  to derive  $u = x - \mathbf{p}'\mathbf{w}$ . We therefore obtain implicit Marshallian demands with budget shares of the form

$$\begin{aligned}\mathbf{w} &= \mathbf{m}(x - \mathbf{p}'\mathbf{w}, \mathbf{z}) + \varepsilon \\ &= \mathbf{m}(y, \mathbf{z}) + \varepsilon\end{aligned}$$

where  $y = g(\mathbf{w}, \mathbf{p}, x, \mathbf{z}) = x - \mathbf{p}'\mathbf{w}$ . Here implicit utility  $y = x - \mathbf{p}'\mathbf{w}$  equals the log of nominal expenditures  $x$  deflated by the log Stone (1954) price index  $\mathbf{p}'\mathbf{w}$ . Here,  $\exp(y)$  is equal to nominal expenditure deflated by a price index equal to 1 when all prices equal 1. For this reason, we may interpret  $y$  in this example as the log of real expenditures.

Unlike the AID system, where a Stone index is used to approximate the correct deflator for  $x$ , in this model the Stone index is the exact, correct deflator for obtaining real expenditures. This is an example of what we call an Exact Stone Index cost function, which is any cost function for which  $y = u = x - \mathbf{p}'\mathbf{w}$  (the details of this class are provided in section 5.1 of the Appendix). This model does not have a closed form expression for indirect utility or ordinary Marshallian demand functions except for very special choices of  $\mathbf{m}(u, \mathbf{z})$ , but it can still be readily estimated because it expresses budget shares as functions of observables  $\mathbf{p}, \mathbf{z}$  and  $y = x - \mathbf{p}'\mathbf{w}$ . This example also shows how implicit Marshallian demands can easily incorporate unobserved preference heterogeneity, since the error terms in the nonlinear regression of  $\mathbf{w}$  on  $y$  equal the random utility parameters  $\varepsilon$ . This example also shows that the functional form of budget shares over  $y, \mathbf{z}$  is completely unrestricted, and Gorman's rank three limit does not apply. Finally, although  $y$  is endogenous (it depends on  $\mathbf{w}$ ), it can be instrumented (for example, by  $x$  and  $\mathbf{p}$ ), so estimation can take the form of instrumental variables, nonlinear 3SLS or GMM.

This example model is clearly very restrictive, particularly regarding price effects which enter the model only through  $y$ . More generally, in the Appendix we show (by Theorems 1 and 2) that all Exact Stone Index demand systems have some undesirable properties, so we instead propose a generalization in which implicit utility equals an affine transform of the log of Stone-index deflated expenditures. These are the Exact Affine Stone Index (EASI) implicit Marshallian demand functions.

Theorems 3 and 4 in the Appendix describe EASI cost functions, which are those with the property that utility is ordinally equivalent to an affine transformation of  $x - \mathbf{p}'\mathbf{w}$ . Equation (22) in the Appendix describes one such cost function that is particularly convenient for empirical implementation:

$$C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{p}'\mathbf{m}(u, \mathbf{z}) + T(\mathbf{p}, \mathbf{z}) + S(\mathbf{p}, \mathbf{z})u + \mathbf{p}'\varepsilon, \quad (2)$$

This cost function differs from (2) only in the addition of  $T(\mathbf{p}, \mathbf{z}) + S(\mathbf{p}, \mathbf{z})u$ . It is similar to the class of cost functions  $C(\mathbf{p}, u) = u + \mathbf{p}'\mathbf{m}(u) + \frac{1}{2}\mathbf{p}'M(u)\mathbf{p}$  proposed by Pendakur and Sperlich (2005), though they estimate ordinary Marshallian demands for their model by numerically solving for  $u$ . There is also a connection to Gorman's (1976) "Tricks with Utility Functions," though we employ an affine transform of deflated expenditures instead of affine transforms of price or quantity vectors.

By Theorem 4 in the Appendix, this class of cost functions has implicit Marshallian demands  $\mathbf{w} = \boldsymbol{\omega}(\mathbf{p}, y, \mathbf{z}, \varepsilon)$  given by

$$\mathbf{w} = \mathbf{m}(y, \mathbf{z}) + \nabla_{\mathbf{p}}T(\mathbf{p}, \mathbf{z}) + \nabla_{\mathbf{p}}S(\mathbf{p}, \mathbf{z})y + \varepsilon \quad (3)$$

where  $y$  is given by

$$u = y = \frac{x - \mathbf{p}'\mathbf{w} - T(\mathbf{p}, \mathbf{z}) + \mathbf{p}'[\nabla_{\mathbf{p}}T(\mathbf{p}, \mathbf{z})]}{1 + S(\mathbf{p}, \mathbf{z}) - \mathbf{p}'[\nabla_{\mathbf{p}}S(\mathbf{p}, \mathbf{z})]} \quad (4)$$

which is an affine transform of the log of Stone Index deflated nominal expenditures. Properties of this general class of EASI models are provided in the Appendix.

## 2.2 An EASI Demand Model for Estimation

We propose the following parametric EASI cost function as our baseline case for empirical work:

$$C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{p}' \left[ \sum_{r=0}^5 \mathbf{b}_r u^r + \mathbf{Cz} + \mathbf{Dzu} \right] + \frac{1}{2} \sum_{l=0}^L z_l \mathbf{p}' \mathbf{A}_l \mathbf{p} + \frac{1}{2} \mathbf{p}' \mathbf{B} \mathbf{p} u + \mathbf{p}' \varepsilon. \quad (5)$$

Here  $z_0 = 1$  and (for notational convenience inside the  $\mathbf{m}$  function)  $z_0$  is not an element of the vector  $\mathbf{z} = (z_1, \dots, z_L)'$ ; each  $\mathbf{b}_r$  is a  $J$ -vector of parameters with  $\mathbf{1}'_J \mathbf{b}_0 = 1$ ,  $\mathbf{1}'_J \mathbf{b}_r = 0$  for  $r \neq 0$ ;  $\mathbf{A}_l$  for  $l = 0, \dots, L$ , and  $\mathbf{B}$  are  $J \times J$  symmetric matrices with  $\mathbf{1}'_J \mathbf{A}_l = \mathbf{1}'_J \mathbf{B} = \mathbf{0}'_J$ ; and both  $\mathbf{C}$  and  $\mathbf{D}$  are  $L \times J$  matrices with  $\mathbf{1}'_J \mathbf{C} = \mathbf{1}'_J \mathbf{D} = \mathbf{0}_L$ .

This model for empirical application is the special case of the EASI model (2) in which

$$\begin{aligned} T(\mathbf{p}, \mathbf{z}) &= \frac{1}{2} \sum_{l=0}^L z_l \mathbf{p}' \mathbf{A}_l \mathbf{p}, \\ S(\mathbf{p}, \mathbf{z}) &= \frac{1}{2} \mathbf{p}' \mathbf{B} \mathbf{p}, \end{aligned}$$

and

$$\mathbf{m}(u, \mathbf{z}) = \left( \sum_{r=0}^5 \mathbf{b}_r u^r \right) + \mathbf{Cz} + \mathbf{Dzu}. \quad (6)$$

The vector of functions  $\mathbf{m}(u, \mathbf{z})$ , which generates the model's Engel curves, could be replaced with other sets of flexible functions of  $u, \mathbf{z}$  such as splines. We use a polynomial in  $u$  and an affine function of  $\mathbf{z}$  to strike a balance between simple tractability and general Engel curve flexibility. We choose quadratic forms in prices for  $T$  and  $S$  because these have simple gradient vectors  $\nabla_{\mathbf{p}}T(\mathbf{p}, \mathbf{z})$  and  $\nabla_{\mathbf{p}}S(\mathbf{p}, \mathbf{z})$ . Finally, we make  $S$  independent of  $\mathbf{z}$  for the sake of terseness.

By Shephard's lemma, the cost function (5) has Hicksian (compensated) budget shares

$$\mathbf{w} = \sum_{r=0}^5 \mathbf{b}_r u^r + \mathbf{Cz} + \mathbf{Dzu} + \sum_{l=0}^L z_l \mathbf{A}_l \mathbf{p} + \mathbf{Bpu} + \varepsilon \quad (7)$$

It can be readily checked from these formulas that

$$C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{p}'\mathbf{w} - \sum_{l=0}^L z_l \mathbf{p}'\mathbf{A}_l \mathbf{p} / 2 - \mathbf{p}'\mathbf{Bpu} / 2,$$

and solving this expression for  $u$  yields implicit utility  $y$ :

$$y = g(\mathbf{w}, \mathbf{p}, x, \mathbf{z}) = \frac{x - \mathbf{p}'\mathbf{w} + \sum_{l=0}^L z_l \mathbf{p}'\mathbf{A}_l \mathbf{p} / 2}{1 - \mathbf{p}'\mathbf{Bp} / 2}. \quad (8)$$

This  $y$  has many of the properties of log real expenditures. It equals a cardinalization of utility  $u$ , it is affine in nominal expenditures  $x$ , and it equals  $x$  in the base period when all prices equal one (which is when log prices  $\mathbf{p}$  equal zero). Also, when  $\mathbf{B}$  is zero,  $y$  exactly equals the log of nominal expenditures deflated by a price index. We later show empirically that  $y$  is very highly correlated with the log of stone index deflated nominal expenditures, which is a very popular ad hoc measure of real expenditures. Like any money-metric utility measure,  $y$  is just a mathematically convenient representation of utility and need not have any deeper significance as an objective or interpersonally comparable utility measure (see the "Shape Invariance and Equivalence Scales" section in the Appendix for details).

Substituting implicit utility  $y$  into the Hicksian budget shares yields the implicit Marshallian budget shares

$$\mathbf{w} = \sum_{r=0}^5 \mathbf{b}_r y^r + \mathbf{Cz} + \mathbf{Dzy} + \sum_{l=0}^L z_l \mathbf{A}_l \mathbf{p} + \mathbf{Bpy} + \varepsilon. \quad (9)$$

which is the matrix form of the equation provided in the introduction.

Given  $y$ , which is a function of observables  $x, \mathbf{p}, \mathbf{z}$  and the log Stone index  $\mathbf{p}'\mathbf{w}$ , the budget share equations (9) are linear in parameters and so can be easily estimated. An approximate estimator applies ordinary (linear) least squares after replacing  $y$  with  $x - \mathbf{p}'\mathbf{w}$  in (9), just like estimation of the approximate AID model. Better estimators simultaneously estimate the model



(9) with the parameters in  $y$  given by (8). One could use 3SLS or GMM estimation to account for the endogeneity that results from having  $\mathbf{w}$  appear in  $y$ . Use of GMM would allow for possible unknown heteroskedasticity in  $\varepsilon$ .

The EASI budget shares (9) have compensated price effects governed by  $\mathbf{A}_l, l = 0, 1, 2, \dots, L$ , and  $\mathbf{B}$ , which allow for flexible price effects and for flexible interactions of these effects with expenditure and with observable demographic characteristics. The Engel curve terms  $\mathbf{b}_r, r = 0, 1, 2, \dots, 5$  specify budget shares as fifth order polynomials in  $y$ , which (recalling that  $y$  is affine in log nominal expenditures  $x$ ) allows Engel curves to have very complicated shapes. Some empirically popular demand functions have budget shares quadratic in log total expenditures, corresponding to  $r = 0, 1, 2$ . We added higher moments  $r = 3, 4, 5$  because they were statistically significant. The terms  $\mathbf{C}$  and  $\mathbf{D}$  allow demographic characteristics to enter budget shares through both intercept and slope terms on  $y$ . The random utility parameters  $\varepsilon$ , representing unobserved preference heterogeneity, take the form of simple additive errors in the implicit Marshallian demand equations.

The complete set of properties required for the EASI cost function (5) to satisfy cost function regularity are as follows. The adding up constraint and cost function homogeneity conditions are satisfied with  $\mathbf{1}'_J \mathbf{b}_0 = 1, \mathbf{1}'_J \mathbf{b}_r = 0$  for  $r \neq 0, \mathbf{1}'_J \mathbf{A}_l = \mathbf{1}'_J \mathbf{B} = \mathbf{0}'_J, \mathbf{1}'_J \mathbf{C} = \mathbf{1}'_J \mathbf{D} = \mathbf{0}'_L$ , and  $\varepsilon' \mathbf{1}_J = 0$ . Symmetry of  $\mathbf{A}_l$  and  $\mathbf{B}$  ensures Slutsky symmetry. Strict monotonicity of cost requires  $\partial C(\mathbf{p}, u, \mathbf{z}, \varepsilon) / \partial u > 0$ , which implies  $\mathbf{p}' \left[ \left( \sum_{r=0}^5 \mathbf{b}_r r u^{r-1} \right) + \mathbf{Dz} + \mathbf{Bp}/2 \right] > -1$ . Finally, we require concavity of  $\exp[C(\mathbf{p}, u, \mathbf{z}, \varepsilon)]$ , and a sufficient condition for concavity is that  $\sum_{l=0}^L z_l \mathbf{A}_l + \mathbf{B}u$  be negative semidefinite. These constraints are assumed to hold for every value that the variables  $x, \mathbf{p}, \mathbf{z}, \varepsilon$  can take on, and hence every value that  $u, \mathbf{p}, \mathbf{z}, \varepsilon$  can take on. It is shown in the Appendix that this model can be globally regular if the set  $\{x, \mathbf{p}, \mathbf{z}, \varepsilon\}$  has bounded support, and that  $\varepsilon$  can be distributed independently of  $x, \mathbf{p}, \mathbf{z}$ .

Apart from the construction of  $y$ , the implicit Marshallian demand equations (9) are linear in coefficients, which simplifies estimation. In this model the  $\mathbf{D}$  and  $\mathbf{B}$  matrix parameters allow for flexible interactions between  $y$  and both  $\mathbf{z}$  and  $\mathbf{p}$ . Either or both of these matrices could be zero if such interactions are not needed. Note that if  $\mathbf{B}$  were zero then  $y$  in equation (8) would also be linear in parameters.

Our empirical model has Engel curves that are high order polynomials. Unlike Marshallian demands, our implicit Marshallian EASI demands can be polynomials of any degree without being bound by Gorman (1981) and Lewbel (1991) type rank restrictions. Note, however, that monotonicity and concavity of the cost function places inequality constraints on the model, which restricts the range of possible parameter values and the range of values of  $\mathbf{p}, y, \mathbf{z}, \varepsilon$  for which these demand functions satisfy regularity.

Polynomials are simple but are not required, i.e., we can maintain linearity in coefficients by replacing equation (6) with  $\mathbf{m}(y, \mathbf{z}) = \mathbf{Bn}(y, \mathbf{z})$  for some  $J \times K$  matrix of constants  $\mathbf{B}$  and  $K$ -vector

of known functions  $\mathbf{n}(y, \mathbf{z})$ . Our chosen functional form takes  $\mathbf{n}(y, \mathbf{z})$  to be a vector of elements of the form  $u^r$  and  $u^r z_l$ , but other functions could also be chosen. For example, the elements of  $\mathbf{n}(y, \mathbf{z})$  could be splines or bounded functions such as logistic transformations of polynomials (which would automatically bound estimated budget shares). Semiparametric specifications could be obtained by letting  $\mathbf{n}(y, \mathbf{z})$  be basis functions with the number of elements of  $\mathbf{n}$  growing to infinity with sample size.

### 2.3 Approximate Fully Linear Model

The demand functions (9) are linear in parameters except for the terms  $\sum_{l=0}^L z_l \mathbf{p}' \mathbf{A}_l \mathbf{p}$  and  $\mathbf{p}' \mathbf{B} \mathbf{p}$  that appear in the construction of  $y$  in (8). A similar nonlinearity appears in Deaton and Muellbauer's (1980) AID system and Banks, Blundell, and Lewbel's (1997) QUAID system, and can be dealt with in an analogous way, either by nonlinear estimation or by replacing  $y$  with an observable approximation. Consider approximating our real expenditures measure  $y$  with nominal expenditures deflated by a Stone price index, that is, replace  $y$  with  $\tilde{y}$  defined by

$$\tilde{y} = x - \mathbf{p}' \bar{\mathbf{w}} \quad (10)$$

for some set of budget shares  $\bar{\mathbf{w}}$ . Then by comparison with equation (9) we have

$$\mathbf{w} = \sum_{r=0}^5 \mathbf{b}_r \tilde{y}^r + \mathbf{C} \mathbf{z} + \mathbf{D} \mathbf{z} \tilde{y} + \sum_{l=0}^L z_l \mathbf{A}_l \mathbf{p} + \mathbf{B} \mathbf{p} \tilde{y} + \tilde{\varepsilon} \quad (11)$$

where  $\tilde{\varepsilon} \approx \varepsilon$  with  $\tilde{\varepsilon}$  defined to make equations (11) hold. We call the model of equations (10) and (11) the *Approximate EASI* model.

The Approximate EASI nests the model  $\mathbf{w} = \mathbf{b}_0 + \mathbf{b}_1 \tilde{y} + \mathbf{C} \mathbf{z} + \mathbf{A} \mathbf{p} + \tilde{\varepsilon}$ , which is identical to the popular approximate Almost Ideal Demand System (AID) if  $\tilde{y} = x - \mathbf{p}' \bar{\mathbf{w}}$  (that is,  $\bar{\mathbf{w}} = \mathbf{w}$ ). The AID without the approximation has  $y$  equal to deflated  $x$  where the log deflator is quadratic in  $\mathbf{p}$ , while the EASI model without approximation has  $y$  equal to an affine transform of  $x - \mathbf{p}' \bar{\mathbf{w}}$ .

The approximate EASI also nests the model  $\mathbf{w} = \mathbf{b}_0 + \mathbf{b}_1 \tilde{y} + \mathbf{b}_2 \tilde{y}^2 + \mathbf{C} \mathbf{z} + \mathbf{A} \mathbf{p} + \tilde{\varepsilon}$ , which is the model estimated by Blundell, Pashardes, and Weber (1993). Their motivation for this model was by analogy with the Almost Ideal, but if this model was really to be Marshallian then, as they show, utility maximization would require linear rank restrictions on the coefficients  $\mathbf{b}_0$ ,  $\mathbf{b}_1$ , and  $\mathbf{b}_2$ , which they did not impose. The approximate EASI implicit Marshallian demand function therefore provides a rationale for the unrestricted model that Blundell, Pashardes, and Weber actually estimated.

The approximate EASI model, substituting equation (10) into (11), can be estimated by linear

regression methods, with linear cross-equation symmetry restrictions on the  $\mathbf{A}_l$  and  $\mathbf{B}$  coefficients. A natural choice for  $\bar{\mathbf{w}}$  is the sample average of budget shares across consumers. A better approximation to  $y$  would be to let  $\bar{\mathbf{w}}$  be each consumer's own  $\mathbf{w}$ , so each consumer has their own Stone index deflator based on their own budget shares. However, as discussed later, this better approximation introduces endogeneity.

In our empirical application, we estimate the approximate model with  $\bar{\mathbf{w}} = \mathbf{w}$  using seemingly unrelated regressions, and we estimate the true EASI model using nonlinear three stage least squares. As in the approximate AID system, there is no formal theory regarding the quality of the approximation that uses  $\tilde{y}$  in place of  $y$  but we find empirically that approximate model estimates do not differ much from estimates based on the exact  $y$  (most estimated parameters have the same signs and roughly similar magnitudes), and provide good starting values for exact model estimation.

## 2.4 Elasticities and Consumer Surplus

We now show how to evaluate the effects of changing prices or other variables in EASI models. We will give results both for our specific empirical model and for the general EASI model given by equations (2), (3), and (4).

First consider evaluating the cost to an individual of a price change. A consumer surplus measure for the price change from  $\mathbf{p}_0$  to  $\mathbf{p}_1$  is the log cost of living index, which for the cost function (2) is given by

$$C(\mathbf{p}_1, u, \mathbf{z}, \varepsilon) - C(\mathbf{p}_0, u, \mathbf{z}, \varepsilon) = (\mathbf{p}_1 - \mathbf{p}_0)' \mathbf{m}(u, \mathbf{z}) + T(\mathbf{p}_1, \mathbf{z}) - T(\mathbf{p}_0, \mathbf{z}) + S(\mathbf{p}_1, \mathbf{z})u - S(\mathbf{p}_0, \mathbf{z})u + (\mathbf{p}_1 - \mathbf{p}_0)' \varepsilon.$$

If  $C(\mathbf{p}_0, u, \mathbf{z}, \varepsilon)$  is the cost function of a household that has budget shares  $\mathbf{w}_0$  and implicit utility level  $y$  then this expression can be rewritten in terms of observables as

$$C(\mathbf{p}_1, u, \mathbf{z}, \varepsilon) - C(\mathbf{p}_0, u, \mathbf{z}, \varepsilon) = (\mathbf{p}_1 - \mathbf{p}_0)' [\mathbf{w}_0 - \nabla_{\mathbf{p}} T(\mathbf{p}_0, \mathbf{z}) - \nabla_{\mathbf{p}} S(\mathbf{p}_0, \mathbf{z})y] + T(\mathbf{p}_1, \mathbf{z}) - T(\mathbf{p}_0, \mathbf{z}) + S(\mathbf{p}_1, \mathbf{z})y - S(\mathbf{p}_0, \mathbf{z})y.$$

For our base empirical model log cost function (5), this log cost of living index expression simplifies to

$$C(\mathbf{p}_1, u, \mathbf{z}, \varepsilon) - C(\mathbf{p}_0, u, \mathbf{z}, \varepsilon) = (\mathbf{p}_1 - \mathbf{p}_0)' \mathbf{w}_0 + \frac{1}{2} (\mathbf{p}_1 - \mathbf{p}_0)' \left( \sum_{l=0}^L z_l \mathbf{A}_l + \mathbf{B}y \right) (\mathbf{p}_1 - \mathbf{p}_0).$$

The first term in this cost of living index is the Stone index for the price change,  $(\mathbf{p}_1 - \mathbf{p}_0)' \mathbf{w}_0$ . Such indices are commonly used on the grounds that they are appropriate for small price changes and that they allow for unobserved preference heterogeneity across households. In our model, the presence of the second term, which depends upon  $T$  and  $S$ , allows us to explicitly model substitution effects, and so consider large price changes, while also accounting for the behavioral importance of both observed and unobserved heterogeneity.

Define semielasticities to be derivatives of budget shares with respect to log prices  $\mathbf{p}$ , implicit utility  $y$ , log nominal total expenditures  $x$ , and demographic characteristics (or other observed taste shifters)  $\mathbf{z}$ . The semielasticity of a budget share can be converted into an ordinary elasticity of a budget share by dividing by that budget share. We provide semielasticities because they are easier to present algebraically. Hicksian demands are given by

$$\omega(\mathbf{p}, u, \mathbf{z}, \varepsilon) = \mathbf{m}(u, \mathbf{z}) + \nabla_{\mathbf{p}} T(\mathbf{p}, \mathbf{z}) + \nabla_{\mathbf{p}} S(\mathbf{p}, \mathbf{z})u + \varepsilon,$$

so the Hicksian price semielasticities are

$$\nabla_{\mathbf{p}'} \omega(\mathbf{p}, u, \mathbf{z}, \varepsilon) = \nabla_{\mathbf{p}\mathbf{p}'} T(\mathbf{p}, \mathbf{z}) + \nabla_{\mathbf{p}\mathbf{p}'} S(\mathbf{p}, \mathbf{z})u = \nabla_{\mathbf{p}\mathbf{p}'} T(\mathbf{p}, \mathbf{z}) + \nabla_{\mathbf{p}\mathbf{p}'} S(\mathbf{p}, \mathbf{z})y.$$

These are equivalently the price semielasticities holding  $y$  fixed,  $\nabla_{\mathbf{p}'} \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)$ . Similarly, semielasticities with respect to  $y$ , interpretable as real expenditure elasticities, are given by

$$\nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon) = \nabla_y \mathbf{m}(y, \mathbf{z}) + \nabla_{\mathbf{p}} S(\mathbf{p}, \mathbf{z}),$$

and semielasticities with respect to observable demographics  $\mathbf{z}$  are

$$\nabla_{\mathbf{z}} \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon) = \nabla_{\mathbf{z}} \mathbf{m}(y, \mathbf{z}) + \nabla_{\mathbf{p}\mathbf{z}} T(\mathbf{p}, \mathbf{z}) + \nabla_{\mathbf{p}\mathbf{z}} S(\mathbf{p}, \mathbf{z})u$$

Compensated semielasticities for our baseline model, the log cost function (5), are linear apart from the construction of  $y$ . Compensated price semielasticities are given by

$$\nabla_{\mathbf{p}'} \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon) = \sum_{l=0}^L z_l \mathbf{A}_l + \mathbf{B}y, \quad (12)$$

and semielasticities with respect to  $y$  are

$$\nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon) = \sum_{r=1}^5 \mathbf{b}_r r y^{r-1} + \mathbf{D}\mathbf{z} + \mathbf{B}\mathbf{p}, \quad (13)$$

which can vary quite a bit as  $y$  changes, reflecting a high degree of Engel curve flexibility. Demo-

graphic semielasticities are given by

$$\nabla_{z_l} \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon) = \mathbf{c}_l + \mathbf{d}_l y + \mathbf{A}_l \mathbf{p}. \quad (14)$$

where  $\mathbf{c}_l$  and  $\mathbf{d}_l$  are the appropriate rows of  $\mathbf{C}$  and  $\mathbf{D}$ , respectively, which allows for price and  $y$  interactions with demographic effects.

Closed form expressions for Marshallian elasticities are more complicated, and so are provided in the appendix.

### 3 Estimation

#### 3.1 Estimators

We estimate demand systems with  $J$  goods, so as usual we can drop the last equation (the  $J$ 'th good) from the system and just estimate the remaining system of  $J - 1$  equations. The parameters of the the  $J$ 'th good are then recoverable from the adding up constraint that budget shares sum to one. Assume this is done in all of the following discussion. The system of equations to be estimated is (9).

The approximate EASI, equation (11) with  $\tilde{y}$  given by equation (10), is trivial to estimate. If the approximate EASI  $\tilde{\varepsilon}$  is uncorrelated with  $\mathbf{p}$ ,  $\mathbf{z}$ ,  $\mathbf{z}\tilde{y}$ ,  $\mathbf{p}\tilde{y}$ ,  $\mathbf{p}z_l$ , for  $l = 1, \dots, L$  and  $\tilde{y}^r$  for  $r = 0, \dots, 5$ , then, without imposing symmetry of the  $\mathbf{A}_l$  and  $\mathbf{B}$  matrices, estimating each approximate EASI equation separately by linear ordinary least squares is consistent and equivalent to linear seemingly unrelated regressions (SUR). Imposing symmetry of the  $\mathbf{A}_l$  and  $\mathbf{B}$  matrices means imposing linear cross-equation equality constraints on the coefficients, but the resulting approximate EASI model can still be consistently estimated using ordinary linear SUR. Since this model is only an approximation to the EASI model, we should not expect these uncorrelatedness assumptions to hold exactly, but we found that the approximate EASI estimates were generally quite close to the exact model estimates, and can provide useful parameter starting values for exact model estimation.

The exact EASI model (without approximation) that we estimate has equation (8) substituted into equation (9) to give

$$\begin{aligned} \mathbf{w} = & \sum_{r=0}^5 \mathbf{b}_r \left( \frac{x - \mathbf{p}'\mathbf{w} + \sum_{l=0}^L z_l \mathbf{p}'\mathbf{A}_l \mathbf{p}/2}{1 - \mathbf{p}'\mathbf{B}\mathbf{p}/2} \right)^r + \mathbf{C}\mathbf{z} + \sum_{l=0}^L z_l \mathbf{A}_l \mathbf{p} + \\ & (\mathbf{D}\mathbf{z} + \mathbf{B}\mathbf{p}) \left( \frac{x - \mathbf{p}'\mathbf{w} + \sum_{l=0}^L z_l \mathbf{p}'\mathbf{A}_l \mathbf{p}/2}{1 - \mathbf{p}'\mathbf{B}\mathbf{p}/2} \right) + \varepsilon. \end{aligned} \quad (15)$$

Equation (15) is nonlinear in parameters because of the presence of  $\mathbf{A}_l$  and  $\mathbf{B}$  in  $y$ . An additional

complication for estimation is that  $\mathbf{w}$ , which is endogenous, appears on the right side of equation (15) in the Stone index  $\mathbf{p}'\mathbf{w}$ . One possible estimator is to ignore this endogeneity and simply estimate the resulting system of equations by nonlinear least squares. In our empirical application we find that the resulting bias from ignoring this endogeneity is small, because, as we discuss below, the variation in the  $\mathbf{p}'\varepsilon$  component of  $\mathbf{p}'\mathbf{w}$  is tiny relative to the variation in  $x$  and the other components of  $y$ .

To formally account for endogeneity, nonlinearity in parameters, and possible heteroskedasticity of unknown form in  $\varepsilon$ , we use an instrumental variables estimator. Let  $\mathbf{q}$  be an  $M$ -vector of observable variables that are uncorrelated with  $\varepsilon$ , which will be used as instruments for estimation. If  $E(\varepsilon \mid x, \mathbf{p}, \mathbf{z}) = \mathbf{0}_J$  then  $\mathbf{q}$  can include any bounded functions of  $\mathbf{p}$ ,  $\mathbf{z}$ , and  $x$ , for example,  $\mathbf{p}$ ,  $\mathbf{z}$ , and  $\tilde{y} = x - \mathbf{p}'\bar{\mathbf{w}}$  with  $\bar{\mathbf{w}}$  equal to the sample average budget-share vector. However, if unobserved heterogeneity is correlated with some observed characteristics such as  $x$  or elements of  $\mathbf{z}$ , then those elements must be excluded from  $\mathbf{q}$ . Let  $\mathbf{q} = (q_1, \dots, q_M)'$ . Then  $E[\varepsilon'q_m] = \mathbf{0}_J$  for  $m = 1, \dots, M$ , which implies

$$E \left[ \begin{pmatrix} \mathbf{w} - \sum_{r=0}^5 \mathbf{b}_r \left( \frac{x - \mathbf{p}'\mathbf{w} + \sum_{l=0}^L z_l \mathbf{p}'\mathbf{A}_l \mathbf{p}/2}{1 - \mathbf{p}'\mathbf{B}\mathbf{p}/2} \right)^r - \mathbf{C}\mathbf{z} - \sum_{l=0}^L z_l \mathbf{A}_l \mathbf{p} - \\ (\mathbf{D}\mathbf{z} + \mathbf{B}\mathbf{p}) \left( \frac{x - \mathbf{p}'\mathbf{w} + \sum_{l=0}^L z_l \mathbf{p}'\mathbf{A}_l \mathbf{p}/2}{1 - \mathbf{p}'\mathbf{B}\mathbf{p}/2} \right) \end{pmatrix} q_m \right] = \mathbf{0}_J \quad (16)$$

for  $m = 1, \dots, M$ .

Allowing for generally heteroskedastic  $\varepsilon$ , parameters may be estimated by applying Hansen's (1982) Generalized Method of Moments (GMM) to this set of moment conditions. By adding up, the moments associated with the  $J$ 'th good are superfluous, so with  $M$  instruments we will have  $(J - 1)M$  moments. An alternative, simpler estimator (which we found to be numerically better behaved in practice) is homoskedastic nonlinear three stage least squares (3SLS). This 3SLS is, like GMM or heteroskedastic 3SLS, consistent with heteroskedastic  $\varepsilon$ , but is only as asymptotically efficient as these other estimators when  $\varepsilon$  is homoskedastic.

Another possible estimator (suggested by a referee) is a form of iterative linear three stage least squares, as follows. Let  $\theta$  be the vector of all the parameters in the model. If  $y$  were observed, we could estimate  $\theta$  by a linear three stage least squares, that is, we could estimate the system of linear equations (9) using the vector  $\mathbf{q}$  as instruments. Given observations (consumers)  $y_1, \dots, y_n$ , let  $\hat{\theta} = \Theta_n(y_1, \dots, y_n)$  denote this linear three stage least squares estimator, or any other linear model GMM estimator based on the moments  $E[\varepsilon'q_m] = \mathbf{0}_J$  for  $m = 1, \dots, M$  with  $\varepsilon$  given by equation (9). For each consumer  $i$  we have  $\hat{y}_i = g(\mathbf{w}_i, \mathbf{p}_i, x_i, \mathbf{z}_i, \hat{\theta})$  for the function  $g$  defined by equation (8). The estimator would then consist of taking some starting values for  $\hat{\theta}$  (these could be estimates using our approximate model estimator of  $\theta$ ), substituting this  $\hat{\theta}$  into  $g$  to obtain estimates  $\hat{y}_1, \dots, \hat{y}_n$ ,

do linear three stage least squares or linear model GMM using these  $\hat{y}$ 's as data to obtain a new value of  $\theta$  given by  $\hat{\theta} = \Theta_n(\hat{y}_1, \dots, \hat{y}_n)$ , and repeat this process using  $\hat{\theta}$  in place  $\hat{\theta}$ , iterating to convergence. This estimator could be easily implemented in a statistical package like Stata that does linear three stage least squares or linear GMM regression. Note that imposing symmetry restrictions requires imposing cross equation equality restrictions on the estimated coefficients in each iteration.

Formally, let the fixed point  $\hat{\theta} = \Theta_n \left[ (g(\mathbf{w}_1, \mathbf{p}_1, x_1, \mathbf{z}_1, \hat{\theta}), \dots, g(\mathbf{w}_n, \mathbf{p}_n, x_n, \mathbf{z}_n, \hat{\theta})) \right]$  define the estimator  $\hat{\theta}$ . This is a special case of the generic class of fixed point based estimators considered by Dominitz and Sherman (2005), who provide associated consistency and limiting distribution theory assuming that the mapping that defines the estimator  $\hat{\theta}$  is a contraction mapping. If the iterations had been linear least squares seemingly unrelated regressions instead of linear three stage least squares regressions, then this estimator would also be an example of the iterated linear least squares estimator for conditionally linear systems proposed by Blundell and Robin (1999). Blundell and Robin apply the estimator to a demand system context similar to ours, and provide an extension to endogenous regressors, but their extension requires a control function form of endogeneity that our model does not satisfy. We do not provide formal asymptotic theory for  $\hat{\theta}$  in our model, but we found that this estimator performed very well in practice, yielding estimates that are numerically quite close to those of our full nonlinear three stage least squares estimator.

The equality restrictions required for demand system rationality are easily imposed in our context. Adding up and homogeneity constraints on the parameter vectors and matrices hold by omitting the  $J$ 'th good and imposing the linear restrictions that  $\mathbf{1}'_J \mathbf{A}_l = \mathbf{1}'_J \mathbf{B} = \mathbf{0}'$ . Slutsky symmetry is satisfied if and only if  $\mathbf{A}_l$  for  $l = 0, \dots, L$  and  $\mathbf{B}$  are all symmetric matrices. All of these parametric restrictions may be imposed as a set of linear constraints in 3SLS or GMM.

Writing this system linearly as equation (9) suggests that good instruments  $\mathbf{q}$  should be highly correlated with  $\mathbf{p}$ ,  $\mathbf{z}$ ,  $\mathbf{z}_y$ ,  $\mathbf{p}_y$ ,  $\mathbf{p}_{z_1}$ , ...,  $\mathbf{p}_{z_L}$  and  $y^r$ . We assume  $E(\varepsilon | \mathbf{p}, x, \mathbf{z}) = \mathbf{0}_J$  and take  $\mathbf{q}$  to be  $\mathbf{p}$ ,  $\mathbf{z}$ ,  $\mathbf{z}_y$ ,  $\mathbf{p}_y$ ,  $\mathbf{p}_{z_1}$ , ...,  $\mathbf{p}_{z_L}$ , and  $\bar{y}^r$  for  $r = 0, \dots, 5$  with  $\bar{y}$  defined as

$$\bar{y} = \frac{x - \mathbf{p}'\bar{\mathbf{w}} - \sum_{l=0}^L z_l \mathbf{p}'\bar{\mathbf{A}}_l \mathbf{p} / 2}{1 + \mathbf{p}'\bar{\mathbf{B}}\mathbf{p} / 2}$$

where  $\bar{\mathbf{w}}$  is the average budget shares across consumers in our sample, and  $\bar{\mathbf{A}}_l$  and  $\bar{\mathbf{B}}$  are the estimated values of  $\mathbf{A}_l$  and  $\mathbf{B}$  based on linear least squares estimation of the approximate EASI model. Note that use of  $\bar{\mathbf{w}}$  and inconsistency of the estimates of  $\bar{\mathbf{A}}_l$  and  $\bar{\mathbf{B}}$  (due to their coming from the approximate model) in the construction of  $\bar{y}$  only affects the quality of the instruments  $\mathbf{q}$  and hence the efficiency of the 3SLS or GMM estimation, but does not cause inconsistency, because  $\bar{y}$  remains uncorrelated with  $\varepsilon$  given any choice of values of the parameters  $\bar{\mathbf{w}}$ ,  $\bar{\mathbf{A}}_l$  and  $\bar{\mathbf{B}}$  in the

construction of  $\bar{y}$ . When symmetry of  $\mathbf{A}_l$  and  $\mathbf{B}$  is not imposed, this set of moments  $E(\varepsilon' q_m) = 0_J$  for  $m = 1, \dots, M$  exactly identifies the EASI model parameters. Imposing symmetry reduces the number of distinct parameters, yielding overidentification.

Since we have assumed  $E(\varepsilon \mid \mathbf{p}, x, \mathbf{z}) = \mathbf{0}_J$ , we could interpret  $\sum_{r=0}^5 \mathbf{b}_r y^r + \mathbf{Cz} + \mathbf{Dz}y$  as a sieve approximation to a general unknown  $J$ -vector of smooth functions  $\mathbf{n}(y, \mathbf{z})$ , with a spanning basis consisting of functions of the form  $y^s z_k^t$  for integers  $s, t$ . Ai and Chen (2003) provide associated rates of convergence, optimal instrument construction, and limiting distribution theory for general semiparametric sieve GMM estimators of this form. Such estimators can attain the semiparametric efficiency bound for the remaining parameters, in this case  $\mathbf{A}_l$  and  $\mathbf{B}$ .

The parametric 3SLS or GMM estimators can be readily modified to deal with possible measurement error or endogeneity in some regressors, by suitably modifying the set of instruments  $\mathbf{q}$ . For example, if simultaneity with supply is a concern (which is more likely to matter significantly in an aggregate demand context than in our empirical application), then we could replace  $\mathbf{p}$  with  $\bar{\mathbf{p}}$  everywhere  $\mathbf{p}$  appears in the construction of the instruments  $\mathbf{q}$ , where  $\bar{\mathbf{p}}$  are the fits from regressing  $\mathbf{p}$  on supply side instruments.

In many data sets, such as the UK Family Expenditure Survey, consumption is measured over a very short time span and so is subject to considerable infrequency-based measurement error. In our empirical application we use Canadian data where consumption is measured over an entire year, which implies less infrequency-based error, but may suffer from recall-based error. One might deal with this problem in part by including functions of income in place of functions of nominal total expenditures in our list of instruments. However, using an instrument like income, which has a large amount of over-time variation that is smoothed out in consumption decisions, may entail considerable efficiency loss. One might alternatively use measures of household wealth as instruments for total expenditure, but unfortunately, most public-use data sets (including ours) do not have both wealth and consumption data. It should also be noted that measurement errors in total expenditures would affect not only our total expenditures regressor, but also the construction of our dependent variables, the budget shares. Estimators such as Lewbel (1996) could potentially be used to deal with the latter problem.

Similarly, if unobserved preference heterogeneity  $\varepsilon$  is correlated with some observed taste shifters (i.e., elements of  $\mathbf{z}$ ), then those may be excluded from the instrument list and replaced by, e.g. nonlinear functions of  $x$  and of the remaining elements of  $\mathbf{z}$ . However, in these cases one would need to take care in interpreting the resulting model residuals  $\hat{\varepsilon}$ , because they will then contain both unobserved preference heterogeneity and measurement error. With panel data one might separate these two effects by modeling the unobserved preference heterogeneity using standard random or fixed effects methods. The exact model estimators remain consistent regardless of heteroskedasticity in  $\varepsilon$ , so for example the estimates are consistent if  $\varepsilon = \mathbf{N}(x, \mathbf{z})\varepsilon^*$  where  $\varepsilon^*$



are preference parameters that are independent of  $\mathbf{p}$ ,  $x$ ,  $\mathbf{z}$ , and features of the function  $\mathbf{N}(x, \mathbf{z})$  could be estimated based on the estimated conditional variance of residuals  $\widehat{\varepsilon}$ , conditioning on  $x, \mathbf{z}$ .

The above described estimators do not impose the inequality (concavity and monotonicity) constraints implied by regularity of the cost function, or equivalently by utility maximization (global regularity is discussed in the Appendix). In addition to these usual utility derived inequality constraints, interpreting  $\varepsilon$  as preference heterogeneity parameters requires  $\varepsilon$  to be independent of  $\mathbf{p}$ , and having the support of  $\varepsilon$  independent of prices imposes additional inequality constraints (see the Appendix for details). If the model is correctly specified then imposing these inequality constraints on parameter estimates will not be binding asymptotically, and so failing to impose them will not result in a loss of asymptotic efficiency, as long as the true parameter values do not lie on the boundary of the feasible region implied by the inequality constraints. Based on this, a common practice in empirical demand analysis, which we will follow in our application, is to estimate the model without imposing inequality constraints, and then check that they are satisfied for a reasonable range of  $\mathbf{p}$ ,  $x$ , and  $\mathbf{z}$  values. In particular, the inequalities associated with utility function regularity are readily verified using elasticity calculations, e.g., the estimated Slutsky matrix can be checked for negative semidefiniteness.

### 3.2 Data and Model Tests

We first estimate the approximate EASI model, equation (11) with  $\tilde{y}$  given by equation (10), using SUR. We then estimate the exact EASI model given by (15) by nonlinear 3SLS with instruments and starting values as described in the previous section. We use 3SLS because we found the GMM weighting matrix estimates to be numerically unstable. Also, 3SLS is asymptotically efficient under the assumption of independently normally distributed  $\varepsilon$ , and our 3SLS estimator remains consistent (though inefficient) if  $\varepsilon$  is not normal or homoskedastic.

The data used in this paper come from the following public use sources: (1) the Family Expenditure Surveys 1969, 1974, 1978, 1982, 1984, 1986, 1990, 1992 and 1996; (2) the Surveys of Household Spending 1997, 1998 and 1999; and (3) Browning and Thomas (1999), with updates and extensions to rental prices from Pendakur (2002). Price and expenditure data are used from 12 years in 4 regions (Atlantic, Ontario, Prairies and British Columbia) yielding 48 distinct price vectors. Prices are normalised so that the price vector facing residents of Ontario in 1986 is  $(1, \dots, 1)$ , that is, these observations define the base price vector  $\bar{\mathbf{p}} = \mathbf{0}_J$ . Since the model contains a high-order polynomial in  $y$ , we subtract the median value of  $x - \mathbf{p}'\mathbf{w}$  from  $y$ . Translating  $y$  by a constant is absorbed by changes in the  $\mathbf{A}_0$ ,  $\mathbf{b}_r$  and  $\mathbf{C}$  coefficients leaving the fit unchanged, and shifting  $y$  to near zero reduces numerical problems associated with data matrix conditioning by giving the  $\mathbf{b}_r$  coefficients comparable magnitudes.

The empirical analysis uses annual expenditure in  $J = 9$  expenditure categories: food-in, food-out, rent, clothing, household operation, household furnishing & equipment, transportation operation, recreation and personal care. Personal care is the left-out equation, yielding eight expenditure share equations to be estimated. These expenditure categories, which exclude large durables, account for about 85% of the current consumption of the households in the sample.

Our sample for estimation consists of 4847 observations of rental-tenure single-member households who had positive expenditures on rent, recreation and transportation. We use only persons aged 25 to 64 living in cities with at least 30,000 residents. We use residents of English Canada only (so Quebec is excluded). Single-member households are used to avoid issues associated with bargaining or other complications that may be associated with the behavior of collective households. City residents are used to minimise the effects of possible home and farm production. Table 1 gives summary statistics for our estimation sample.

We include  $L = 5$  observable demographic characteristics in our model: (1) the person's age minus 40; (2) the sex dummy equal to 1 for men; (3) a car-nonowner dummy equal to 1 if real gasoline expenditures (at 1986 gasoline prices) are less than \$50; (4) a social assistance dummy equal to 1 if government transfers are greater than 10% of gross income; and (5) a time variable equal to the calendar year minus 1986 (that is, equal to zero in 1986).

We include the time variable in the model to account for possible slow changes over time in tastes, quality, and composition of composite goods. Time trends are also commonly included in demand systems to account for some types of habit formation (see, e.g., Pollak and Wales 1992). The time trend may also help to compensate for average growth in, and hence possible nonstationarity of, real total expenditures over time, noting that our dependent variables, budget shares, are bounded. See, e.g., Lewbel and Ng (2005) for a detailed analysis of the effects of covariate nonstationarity in budget-share demand-system models. However, our data consist of repeated cross sections at the household level, so the scope for possible dependence of errors across observations due to total expenditure growth is limited.

These variables are all zero for a 40 year old car-owning female in 1986 who did not receive much government transfer income. We use a limited set of characteristics like this because each additional element of  $z$  increases the number of additional parameters to estimate in our model by 52 (80 additional parameters minus 28 additional symmetry restrictions). One could introduce additional demographics at a lower cost in degrees of freedom by including a large set of characteristics in the  $\mathbf{m}$  function (the  $\mathbf{C}$  and  $\mathbf{D}$  parameter matrices) and a smaller set interacted with prices (the  $\mathbf{A}_l$  matrices). Restrictiveness of the observed heterogeneity specification is also partly mitigated by the relative demographic homogeneity of the sample, and by the role of  $\varepsilon$  which embodies unobserved preference heterogeneity.

When symmetry is not imposed on  $\mathbf{A}_l$  and  $\mathbf{B}$ , then given the instruments specified above, the

model is exactly identified. Denoting the number of terms in each  $\mathbf{b}_r$  (6 in our baseline model) as  $R$ , the symmetry-unrestricted model has  $[R + L + L + (J - 1)(L + 1) + (J - 1)](J - 1) = 576$  parameters for  $\mathbf{b}_r$ ,  $\mathbf{C}$ ,  $\mathbf{D}$ ,  $\mathbf{A}_l$ , and  $\mathbf{B}$  (72 coefficients in each of 8 equations). When symmetry is imposed on  $\mathbf{A}_l$  and  $\mathbf{B}$ , the model has  $(L + 2)(J - 1)(J - 2)/2 = 196$  restrictions, which are overidentifying restrictions for testing exogeneity and are also symmetry restrictions. Since each estimated model has hundreds of parameters, we do not present all the individual coefficients estimates (they are available from the authors on request). We instead provide many summaries and analyses of the results. Even with parameters all set to zero as starting values, we still obtained convergence of the nonlinear 3SLS exact model estimator in only three iterations, due to the near linearity of the EASI model.

Table 2 gives Wald- and J-tests for various hypotheses concerning the model. Since our sample size is large (8 equations times 4847 observations per equation) we use a 1% critical value for all tests. Consider price effects first. The Wald test of symmetry in the asymmetric model is marginally insignificant with a  $p$ -value of 1.4%, and symmetry is not rejected for either the level of prices ( $\mathbf{A}_l = \mathbf{A}'_l$  for all  $l$ ) or for prices interacted with implicit utility  $y$  ( $\mathbf{B} = \mathbf{B}'$ ). This lack of symmetry rejection is not due to the irrelevance of price effects. In particular, the direct price effects given by the  $\mathbf{A}_0$  matrix are strongly significant, as are many of the interactions between prices and demographics in the  $\mathbf{A}_l$  matrices. The interaction of prices with income are less significant, and we cannot reject the hypothesis that the  $\mathbf{p}y$  interaction given by the  $\mathbf{B}$  matrix is excludable.

Turning to expenditure effects, we first check for adequacy of our fifth order polynomial in  $y$  by adding either a  $y^{-1}$  or  $y^6$  term to the model. The reciprocal of  $y$  is excludable with a  $p$ -value of 4.9%. The 6th order term in  $y$  is marginally significant, but these tests covary, and the joint test of exclusion of both  $y^{-1}$  and  $y^6$  is not significant, with a  $p$ -value of 2.9%. Taken together the above results suggest that symmetry may not be violated and that a 5th order polynomial in  $y$  may be sufficient, so we present further results for a symmetry-restricted model with  $r = 0, 1, \dots, 5$ .

The lower panel of Table 2 shows that one cannot further reduce the level of the polynomial in  $y$  by dropping  $y^5$ . Turning to evidence of complicated Engel curve shapes, we test for whether or not each of the 8 budget share equations can be reduced to a quadratic. The tests suggest that 4 of the 8 budget shares are statistically significantly nonquadratic: food-at-home, rent, household operation, and recreation. These departures from quadratic Engel curves suggest that allowing for complex Engel curves is a useful attribute of our methodology. Later we examine our estimated Engel curve shapes in more detail, considering whether the departures from quadratic are economically important in addition to being statistically significant.

Next consider some model summary statistics. The symmetry-restricted model is overidentified, so we can construct overidentification tests of instrument validity. The 196 symmetry-restrictions result in our model having 196 more moments than parameters, and we pass the re-

sulting overidentification based test of validity of the included moments with a  $p$ -value of 2.1%. However, if we include additional instruments based on  $y^{-1}$  and  $y^6$ , we obtain 16 additional moments (two more instruments times eight budget share equations), and the resulting J-test of overidentifying restrictions becomes marginally significant.

Let  $\mathbf{e}$  denote the sample residuals vector, which is also our estimate of  $\varepsilon$  for each household. The  $R^2$  values for the 8 equations (in the order given in Table 2) are: 0.49, 0.31, 0.65, 0.30, 0.19, 0.43, 0.57, and 0.31, respectively. The model is not estimated by ordinary least squares, so these  $R^2$ 's should only be interpreted as summary measures regarding the size of  $\mathbf{e}$  relative to budget shares. These numbers show that much, even most, of the variation in budget shares is due to unobserved heterogeneity. Our approach formally includes  $\varepsilon$  as random utility parameters, so the relatively large variation in  $\mathbf{e}$  yields large effects in our welfare calculation estimates, as we demonstrate below.

Table 3 gives information on the sources of variation in  $y$ . Since both total expenditures and prices may contain time trends, we detrended all the variables in this table by regressing them on dummies for each time period, and report correlations of the resulting residuals from these regressions. The upper block of Table 3 gives standard deviations and pairwise correlations between three variables. The first is  $y$  itself, detrended. The second, denoted  $y(\mathbf{p}'(\mathbf{w} - \mathbf{e}))$  is identical to  $y$  except that it removes  $\mathbf{p}'\mathbf{e}$ , from the  $\mathbf{p}'\mathbf{w}$  component of  $y$  given in equation (8). The third is the instrument  $\bar{y}$ , which is identical to  $y$ , except that it has  $\mathbf{p}'\bar{\mathbf{w}}$ , where  $\bar{\mathbf{w}}$  is the sample-average budget-share vector, in the numerator instead of  $\mathbf{p}'\mathbf{w}$ . These three variables have almost identical variances and correlations over 0.998 even after detrending. This shows that the instrument  $\bar{y}$  is highly relevant, and that, while our estimator controls for endogeneity in  $y$ , it may not be numerically important to do so in practice, since removing the source of endogeneity from  $y$  leaves it virtually unchanged.

Further examining the components of  $y$ , the second block of Table 3 shows correlations between  $y$  and two measures of the log of stone-deflated expenditures,  $x - \mathbf{p}'\mathbf{w}$  and  $\tilde{y} = x - \mathbf{p}'\bar{\mathbf{w}}$ . Even after detrending these variables are also highly correlated with each other, which explains why the approximate EASI model estimates are close to the exact model estimates, even though the approximate EASI does not include the affine transformation of  $x - \mathbf{p}'\mathbf{w}$  that defines  $y$ , and our approximate EASI estimates do not control for endogeneity.

The third block of Table 3 considers the magnitude of  $\mathbf{p}'\mathbf{e}$  in more detail. Here, we see that the standard deviation of  $\mathbf{p}'\mathbf{e}$  is about one-twentieth as large as that of  $y$  and of the log of stone-deflated expenditures,  $x - \mathbf{p}'\mathbf{w}$ , and that  $\mathbf{p}'\mathbf{e}$  is only weakly correlated with these real expenditure measures. This again shows that the endogenous component of  $y$  or  $\tilde{y}$  is small. It also shows that if the errors in our model were treated as measurement errors in budget shares (or more generally, as errors that are non-behavioral or otherwise not part of the model), then our parameter estimates (including our

finding of significantly nonquadratic Engel curves) would be almost unchanged. However, treating the errors as measurement errors rather than preference heterogeneity parameters substantially affects consumer surplus and social welfare calculations, as we show later in the section on consumer surplus estimates. Formally excluding  $\mathbf{p}'\varepsilon$  from  $y$  requires iterative estimation, following the spirit of the iterative nonparametric approach proposed in Pendakur and Sperlich (2005). We discuss the results of such estimation at the end of section 3.5.

The last block of Table 3 gives the standard deviation and correlations of  $\sum_{l=0}^L z_l \mathbf{p}' \mathbf{A}_l \mathbf{p} / 2$  (denoted  $\mathbf{p}' \mathbf{A} \mathbf{p}$  in the Table) and  $\mathbf{p}' \mathbf{B} \mathbf{p} / 2$ , which are the components of the difference between  $y$  and  $x - \mathbf{p}' \mathbf{w}$ . These components have less than one-tenth the standard deviation of  $\mathbf{p}' \mathbf{e}$  and less than a hundredth the standard deviation of  $y$ . As a result, they only drive a tiny fraction of the variation in  $y$ , which again shows why the approximation that ignores these components is quite good.

### 3.3 Estimated Engel Curves

The easiest way to summarize the income related parameter estimates is to examine the resulting expenditure share equations as functions of  $x$  for particular values of  $\mathbf{p}$ ,  $\mathbf{z}$ , and  $\varepsilon$ . At  $\mathbf{p} = \mathbf{0}_J$ ,  $y$  equals log nominal expenditure  $x$ , so at these base prices we obtain Marshallian Engel curves  $\mathbf{w} = \sum_{r=0}^5 \mathbf{b}_r x^r + \mathbf{Cz} + \mathbf{Dz}x + \varepsilon$ . Figures 1-8 show these estimated Engel curves from our model for a 40 year old car-owning female in 1986 who did not receive much government transfer income, and having  $\varepsilon = \mathbf{0}$ . For her,  $\mathbf{w} = \sum_{r=0}^5 \mathbf{b}_r x^r$ . The base period Engel curves for households with different values of unobserved heterogeneity are identical except for being vertically shifted by  $\varepsilon$ . These base period Engel curves are also informative about the shape of Engel curves in other price regimes, since at other price vectors, Engel curves expressed in terms of our real expenditures measure  $y$  differ from the above only by the addition of the linear function  $\sum_{l=0}^L \mathbf{A}_l z_l \mathbf{p} + \mathbf{B} \mathbf{p} y$ .

In Figures 1 to 8, each figure presents the estimated Engel curve for four models: thick black lines indicate 3SLS estimates of the symmetry-restricted *exact* model; thick grey lines indicate 3SLS estimates of the symmetry-restricted *quadratic* (having a second instead of fifth order polynomial in  $y$ ) exact model; dotted thin black lines indicate *approximate* SUR estimates of the model; and medium black lines indicate exact 3SLS estimates of the *asymmetric* model (*italics* denote shorthand). Estimates are computed at each percentile of log expenditures in the data, and estimated 90% confidence intervals (computed via the delta method) are displayed with small crosses for the exact 3SLS symmetry-restricted model at each fifth percentile of the expenditure distribution.

As noted earlier, our quadratic specification is very similar to Quadratic Almost Ideal (QAI) related models estimated by Blundell, Pashardes, and Weber (1993) and Banks, Blundell and Lewbel (1997). The quadratic EASI differs from these QAI models in that it has a different deflator for

total expenditures, and it does not require the coefficient of squared  $y$  to either depend on prices or on the lower order expenditure coefficients to stay consistent with utility maximization. Our specification also allows for more general demographic effects than is typical in applied QAI models, which usually only have terms analogous to **C**, but not **A<sub>l</sub>** or **D**.

Figures 1 and 2 show Engel curves for food-in (food consumed at home) and food-out. Both these Engel curves are almost linear. Both the approximate and quadratic symmetry-restricted models lie within the 90% pointwise confidence intervals of the exact model. However, in the food-out equation, the asymmetric model estimates are above the confidence intervals in the upper part of the expenditure distribution. The food-in budget share equation is statistically significantly non-quadratic (see Table 2), though the empirical size of the departure is small.

Figure 3 shows the Engel curve for rent. Throughout most of the expenditure distribution, the quadratic estimates lie outside the confidence intervals of the exact estimates. In the exact estimates, the curvature of the Engel curve is near zero in the bottom decile, strongly negative in the next two deciles, and near zero until it becomes positive in the top decile. This kind of complexity cannot be captured in a quadratic model but is readily accommodated in the EASI framework. Further, the magnitude of the quadratic departure from the exact model can be large. At the fifth percentile, the quadratic model overestimates the rent budget share by almost 5 percentage points. At the bottom quintile cutoff, the quadratic model underestimates the rent budget share by about 2 percentage points.

For the rent equation, the approximate model performs very well, lying essentially on top of the exact estimates throughout the distribution of expenditures. As in the food-out equation, the asymmetric model performs relatively poorly (though not as poorly as the quadratic model).

Figures 4, 5, 6 and 7 give the household operation, household furnishing & equipment, clothing and transportation operation Engel curves. All four sets of estimates lie within the 90% confidence intervals of the exact model estimates. However, as shown in Table 2, the household operation equation shows evidence of being statistically significantly nonquadratic. The departure from quadratic can be seen in Figure 4, where the exact model estimate looks more like two nearly linear segments joined by a curved segment in the middle instead of a true quadratic function.

Figure 8 gives the recreation Engel curve. As in Figures 4-7, all four sets of estimates lie within the pointwise confidence intervals of the exact model estimates. However, as in the rent equation, the quadratic model is statistically rejected (see Table 2), and the Engel curve looks nonquadratic, particularly in the bottom decile, where the exact model estimate of the recreation Engel curve flattens out, while the quadratic estimate has a strong negative slope.

We now address the question of whether the complexity we observe in our estimated Engel curves is an artifact of our sample or of our choice of modeled expenditure categories. In particular, much of the nonquadratic behavior we observe is in the tails of the  $x$  distribution, where data

are relatively sparse. In addition, many researchers of consumer demand exclude shelter expenditures (ie., rent) on the grounds that such expenditures cannot be adjusted quickly in the face of budget and price changes. A possible concern is whether the nonlinearity we observe in the rent share artificially induces nonlinearity in other estimated budget share functions. We conduct three robustness exercises aimed at these issues. First, we estimated our EASI model using an extended sample of 18,600 households, which includes multiple-member households and households residing in Quebec. For this model, we included 4 additional demographic variables: the log of the number of household members, and dummies indicating single-parent households, childless couple households and households residing in Quebec. Second, we estimated our EASI model on the 8846 households in our extended sample that did not have transfer income exceeding 10% of gross income and who did own a car (called the ‘nopoor’ extended sample). Third, we estimated our EASI model on a subset of expenditure categories which excludes rent. For this model, we used households of all tenures, not just rental-tenure households, and used only those that did not have transfer income exceeding 10% of gross income and that did own a car. This yielded 32,399 observations available for an estimation sample (called the ‘norent nopoor’ extended sample). For the norent nopoor extended sample, we include the number of rooms in the dwelling and its square as additional demographic characteristics. Detailed results for these models are available on request from the authors.

Figures 9 and 10 show the estimated rent and recreation Engel curves for various samples for the reference type defined above facing the base price vector, plus confidence intervals (shown with  $x$ 's) for the Engel curves of the extended sample. Engel curves are evaluated at each fifth percentile of the single-person household real-expenditure distribution for the sample. Confidence intervals for other samples are similar in size, and are suppressed to reduce clutter. Examination of Figure 9 shows that the non-quadratic curvature we noted in Figure 3 for the rent Engel curve remains evident in the much larger extended sample (thin black line). Of course, the extended nopoor sample (dotted line) is uninformative at the bottom of the distribution because excludes the poor, but it does reveal some departure from the quadratic model at the high end: at the 97th percentile of the nopoor distribution, then EASI Engel curve is 2 percentage points (and statistically significantly) higher than the quadratic model would indicate.

Figure 10 gives Engel curves for the baseline sample, the extended sample, the extended nopoor sample and the extended nopoor norent sample. The figure shows that the curvature in the recreation Engel curve observed at the bottom end of the expenditure distribution in Figure 8 may have overstated the change in curvature for poor households. In particular, the extended sample (thin black line) does not reveal much nonquadratic curvature at the bottom of the distribution. Unsurprisingly, since the curvature was seen mainly at the bottom, the extended nopoor sample (dotted line) does not pick it up either. Also, in the larger sample at the upper end of the expenditure distri-

bution, the two bunched recreation Engel curves lie within each others' 90% confidence intervals. Turning to the extended nopoor norent sample (thick grey line), wherein the rent equation is not modeled and a much larger number of observations is used, apart from a slight upturn at the very bottom of the distribution, we see little evidence of important non-quadratic curvature.

There are two important lessons that we draw from these figures. First, the SUR estimates of the approximate model do quite a good job of approximating the Engel curve estimates of the exact model, even when the underlying Engel curves are quite complex. Second, while the demand functions of some goods are close to linear or quadratic in log total expenditures, other goods such as rent and recreation are not quadratic. This implies a demand system rank (see Gorman 1981 and Lewbel 1991) that is higher than three. Most past empirical studies have found rank to equal three, though Lewbel (2003) finds some empirical evidence for rank four. Figures 1 to 10 suggest that the reason why previous studies failed to find ranks higher than three is because most of the departures from quadratic are either somewhat subtle, as in the household operation equation, or are concentrated in the tails of the expenditure distribution as in the rent equation. Either way, the precision gained by detailed model specification and large sample sizes is needed to confirm the departure from quadratic Engel curves.

### 3.4 Estimated Price Effects

Prices in our model vary only by region and year, so despite having thousands of observations we only have 48 different price regimes. However these data still contain a very substantial amount of relative price variation, both because our sample is spread out over 28 years, and because Canada has historically had considerable regional price variation in some goods and services. As a result, we are able to obtain a reasonable amount of precision in many of our price effect estimates.

In our framework, price effects are most easily evaluated by looking at compensated budget-share semi-elasticities, compensated (good-specific) expenditure elasticities, or compensated quantity derivatives (aka, Slutsky terms). As shown by equation (12), compensated budget-share semi-elasticities with respect to prices are given by the matrix  $\Upsilon \equiv \sum_{l=0}^L \mathbf{A}_l z_l + \mathbf{B}y$ , and so are affine in  $\mathbf{z}$  and  $y$ . Compensated (good-specific) expenditure elasticities with respect to prices is closely related, and is given by  $\mathbf{W}^{-1}(\Upsilon + \mathbf{w}\mathbf{w}')$  where  $\mathbf{W} = \text{diag}(\mathbf{w})$ . The normalised<sup>1</sup> Slutsky matrix,  $\mathbf{S}$ , is related to the compensated semi-elasticity matrix,  $\Upsilon$ , by  $\mathbf{S} = \Upsilon + \mathbf{w}\mathbf{w}' - \mathbf{W}$ . Table 4 assesses all 3 of these measures of price effects. Since  $y = x$  at the base price vector, and since  $x = 0$  at median expenditure, this matrix is equal to  $\mathbf{A}_0$  at median expenditure for the reference type of

<sup>1</sup>The Slutsky matrix,  $\bar{\mathbf{S}}$ , is defined as the matrix of compensated quantity derivatives with respect to (unlogged) prices. The normalised Slutsky matrix,  $\mathbf{S}$ , (see Pollak and Wales 1990) normalises the Slutsky matrix for prices and expenditure:  $\mathbf{S} = \mathbf{P}\bar{\mathbf{S}}/x$ , where  $\mathbf{P} = \text{diag}(\exp(\mathbf{p}))$ . Concavity of cost is necessary and sufficient for negative semidefiniteness of both the Slutsky matrix and the normalised Slutsky matrix.



person, a 40 year old car-owning female in 1986 who did not receive much government transfer income. Table 4 presents summary estimated price effects, with asymptotic standard errors in *italics*. The rightmost block of Table 4 gives compensated price semi-elasticities for a reference person with median expenditure from the symmetry-restricted exact 3SLS estimates.

Consider first the matrix of compensated budget-share semi-elasticities for the reference person at median expenditure given by  $\mathbf{A}_0$ . Several of the own-price effects are large and statistically significant. The own-price compensated semi-elasticity for the rent budget share is 0.063, which implies that a rent price increase of 10% would be associated with a budget share 0.63 percentage points higher when expenditure is raised to equate utility with that in the initial situation. In contrast, if the recreation price rises by 10%, its budget share will be a little more than 1 percentage point lower when expenditure is raised to compensate for the loss in utility.

Several cross-price effects are also large and statistically significant, suggesting that substitution effects are important. For example, the clothing budget share compensated rent cross-price semi-elasticity is -0.066, implying that an increase in the price of rent is associated with a significant decrease in the budget share for clothing even after expenditure is raised to hold utility constant.

For some readers, expenditure elasticities may be more easily interpreted than budget-share semi-elasticities. In order to pin down these elasticities, we additionally need to specify the unobserved heterogeneity terms, which enter  $\mathbf{w}$  (but don't enter  $\Gamma$ ). For the reference type at median expenditures with  $\varepsilon = \mathbf{0}_J$ , we may calculate elasticities using  $\mathbf{A}_0$  and the level of the budget share function at median expenditure. The own-price expenditure elasticities are given in the third column of Table 4, with asymptotic standard errors computed via the delta method in *italics*. (The variance of these elasticities is driven by the variance of the own-price compensated semi-elasticities,  $\Gamma$ , and so reported standard errors are very close to those standard errors deflated by the appropriate budget share.) Unsurprisingly, only those equations with statistically significant own-price compensated budget-share semi-elasticities have statistically significant compensated expenditure elasticities. In particular, compensated rent expenditures have an elasticity of 0.528 and compensated transportation operation expenditure have an elasticity of 0.393. In contrast, compensated recreation expenditures are highly negatively elastic, with a marginally statistically significant own-price expenditure elasticity of  $-1.191$ .

Although some of the own-price elasticities and semi-elasticities in Table 4 are statistically significantly positive, this does not imply that concavity (negative semi-definiteness) is violated. Concavity of cost is satisfied if and only if  $\mathbf{S}$  is negative semi-definite (see, e.g., Pollak and Wales 1992). For the case where  $\varepsilon = \mathbf{0}_J$ , the Slutsky matrix for the reference type with median expenditure facing base prices is fully specified by the matrix  $\mathbf{A}_0$  and the value of the Engel curve functions at median expenditure. The values of the own-price Slutsky terms are reported in the

second column of Table 4, with asymptotic standard errors computed via the delta method in *italics*. A glance down this column reveals that the own-price Slutsky terms are all negative, and most are statistically significant. In addition, the Slutsky matrix evaluated at median expenditure for the reference type facing base prices is negative semidefinite, implying that the cost function is weakly concave at this point in the data.

The leftmost column of estimates in Table 4 contains the estimated own-price elements of  $\mathbf{B}$ , which show the magnitudes of the interaction between own-prices and with log total expenditures. These parameters allow us to assess whether or not compensated semi-elasticities are the same for rich and poor households. As noted in our discussion of Table 2, we cannot strongly reject the joint hypothesis that the entire matrix  $\mathbf{B}$  is zero, so these results must be interpreted with caution. However, the estimated coefficient of the rent own-price compensated semi-elasticity on  $y$  is 0.088, and is by itself marginally statistically significant. Consider the comparison between the rent own-price compensated semi-elasticity for a reference person at the 5th percentile of expenditure ( $x = -0.90$ ) versus that for such a person at the 95th percentile of expenditure ( $x = 0.60$ ). As noted above, its value at the median expenditure of  $x = 0$  is  $-0.066$ . At the 5th percentile, its value is  $-0.066 - 0.90 * 0.088 = -0.016$ , and is insignificantly different from zero. In contrast, at the 95th percentile, its value is  $-0.066 + 0.60 * 0.088 = 0.115$ , and appears highly statistically significant. The corresponding own-price rent Slutsky terms are  $-0.236$  at the fifth percentile and  $-0.080$  at the 95th percentile, and both are statistically significantly negative. These results suggest that poor households substitute much more than do rich households in the face of an increase in the price of rent.

We draw three main conclusions from the analysis of price effects. First, we are able to obtain estimates of compensated elasticities, which given symmetry are second-order derivatives of the log-cost function and capture substitution effects. Uncompensated elasticities can also be calculated (see the Appendix for details). These elasticity estimates suggest that some price effects, and therefore substitution effects, are large in magnitude. Our second conclusion is that the rationality restriction of concavity is not violated, at least for the reference type. Third, there is some evidence that substitution effects are different for rich and poor households, and specifically, that poor households substitute much more in the face of rental price increases than do rich households.

### 3.5 Consumer Surplus Estimates

We assess the economic significance of our models with a cost-of-living experiment. In Canada, rent is not subject to sales taxes, which typically amount to 15% for goods such as food-out and clothing. Consider the cost-of-living index associated with subjecting rent to a 15% sales tax for

people facing the base price vector, so that  $\mathbf{p}_0 = \mathbf{0}_J$  and  $\mathbf{p}_1 = [0 \ 0 \ \ln 0.15 \ 0 \ 0 \ 0 \ 0 \ 0] :$

$$\begin{aligned} C(\mathbf{p}_1, u, \mathbf{z}, \boldsymbol{\varepsilon}) - C(\mathbf{p}_0, u, \mathbf{z}, \boldsymbol{\varepsilon}) &= \ln 0.15 w^{rent} + \ln 0.15^2 \left( \sum_{l=0}^L z_l a_l^{rent,rent} + b^{rent,rent} y \right) / 2 \\ &= \ln 0.15 w^{rent} + \ln 0.15^2 \left( \sum_{l=0}^L z_l a_l^{rent,rent} + b^{rent,rent} x \right) / 2 \end{aligned}$$

where  $a_l^{rent,rent}$  and  $b^{rent,rent}$  are the rent own-price elements of  $\mathbf{A}_l$  and  $\mathbf{B}$ . We choose  $\mathbf{0}_J$  as the comparison price vector because at this price vector,  $y = x$ , and as a consequence only 1 budget share and 6 parameters are needed to estimate the cost-of-living index. Here, unobserved heterogeneity enters only through the level effect on  $w^{rent}$ . We can think of this cost of living index as being comprised of two effects: a first-order effect which is driven by expenditure shares and which incorporates unobserved heterogeneity; and a second-order effect which captures substitution effects. Traditional consumer demand analysis which ignores unobserved heterogeneity would accommodate both first- and second-order effects, but would use  $\widehat{w}^{rent}$ , which contains no ‘error term’, rather than  $w^{rent}$  which contains an unobserved preference heterogeneity component. In contrast, traditional nonparametric approaches to the cost-of-living would use only the first-order term which accommodates unobserved heterogeneity, but would not incorporate the second-order term which captures substitution effects. Our model combines the advantages of both approaches.

Figure 11 shows the estimated values of the cost-of-living index for each household facing the base price vector in our baseline sample incorporating unobserved heterogeneity with empty circles and shows estimated values for each household with unobserved heterogeneity set to zero ( $\boldsymbol{\varepsilon} = \mathbf{0}$ ) using filled circles. In addition, the second-order component capturing substitution effects is shown with filled squares. The reason that the filled circles and filled squares do not each lie on a single line is that variation in demographic characteristics  $\mathbf{z}$  across households affect the surplus measures.

The underlying Engel curve is visible in the estimates which zero out  $\boldsymbol{\varepsilon}$ , but is largely obscured when this unobserved heterogeneity is taken into account. Failure to account for unobserved heterogeneity leads to the erroneous impression that most of the variation across individuals in the cost-of-living impact of a large rent increase is related to expenditure, and only a little is related to other characteristics. The more refined picture is that most of the variation in the impact on cost-of-living is attributable to unobserved characteristics. Even in a model as rich as ours, with many hundreds of parameters, most of the variation in demand is due to unobserved characteristics, and this is reflected in the variation in cost of living responses. As noted earlier, our parameter estimates would be little changed if some or all of the errors were interpreted as ordinary modeling error rather than preference heterogeneity. In this case, the cost of living impacts would lie

somewhere between the filled and empty circles in Figure 11.

Figure 11 also shows the need for highly flexible Engel curves. The first-order term in the consumer surplus calculation is driven by the Engel curve, and as Figure 3 shows, even a quadratic provides a poor approximation to the rent expenditure share equation, so demand systems that only allow for linear or quadratic Engel curves can make substantial errors in policy analyses. These errors would be magnified in a policy experiment that more directly affected the distribution of total expenditures, such as a change in the progressivity of income taxes.

The second-order terms above capture substitution effects across expenditure share equations. These effects are not large in this experiment, but they do have a pronounced pattern, as shown by the filled squares in Figure 11. If consumers substitute greatly in the face of price increases, then the second-order terms will be large and negative; if they substitute little, the the second-order terms will be large and positive. Since the poor substitute more than the rich, second-order effects are positively related to expenditure, so ignoring them would result in underestimating the cost-of-living impact for rich households and over-estimating the impact for poor households. The magnitude of the second-order term is about -0.1 percentage points for households at the 5th percentile (on a total impact of about 6 percentage points), and its magnitude is about +0.1 percentage points for households at the 95th percentile (on a total impact of about 4 percentage points).

As noted above, the treatment of  $\epsilon$  as unobserved preference heterogeneity parameters rather than as measurement error affects consumer surplus measures greatly. However, it turns out not to affect estimated parameters of the model very much. We re-estimated the EASI model assuming  $\epsilon$  is measurement error via an iterative procedure. First, we estimate the exact EASI model, and then compute  $y(\mathbf{p}'(\mathbf{w} - \mathbf{e}))$  as in Table 2, which is identical to  $y$  except that it has  $\mathbf{p}'\mathbf{e}$  subtracted from the numerator in the definition of  $y$  given in equation (8). Then, we estimate the approximate EASI model using  $y(\mathbf{p}'(\mathbf{w} - \mathbf{e}))$  as  $\tilde{y}$ , re-compute  $y(\mathbf{p}'(\mathbf{w} - \mathbf{e}))$ , re-estimate the approximate EASI model, and iterate to convergence (4 iterations). The resulting estimated parameter vector is very close to the exact EASI estimated parameter vector. The maximum absolute distance between parameters of 0.1 standard error (using the exact EASI estimated standard errors), and the average absolute distance is 0.005 standard error. In terms of an economic measure of difference, using the ‘measurement error’ version of the model to assess cost-of-living impacts as in Figure 11 yields estimates visually identical to the filled black circles in the Figure. The maximum absolute distance between the filled black circles in the Figure and the ‘measurement error’ version of cost-of-living estimates is 0.01 percentage points, which is negligible in comparison to the estimated values of around 5 percentage points.

## 4 Conclusions

We invent *implicit Marshallian demand* systems. These are the compensated budget share functions (Hicks demands) associated with cost functions having the property that utility can be represented by a simple function of observables, including budget shares themselves. We show that simple implicit Marshallian demand systems allow us to solve many problems facing consumer demand analysis, including the incorporation of unobserved preference heterogeneity and allowing for flexibility in Engel curves.

We provide the *Exact Affine Stone Index* (EASI) implicit Marshallian demand system, in which utility is ordinally equivalent to an affine function of the log of expenditure deflated by the Stone Index. This EASI demand system is as flexible in price responses, as close to linear in parameters, and as easy to estimate as the Almost Ideal Demand (AID) system. But, in contrast to the AID system, the EASI demand system also allows for flexible interactions between prices and expenditures, permits almost any functional form for Engel curves, and allows error terms in the model to correspond to unobserved preference heterogeneity random utility parameters. Demand functions may be estimated by 3SLS or GMM, and, like the AID system, an approximate model can be estimated by linear regression. We estimate the model, and find some significant departures from linear and quadratic demands. We also find that allowing errors to equal unobserved preference heterogeneity has little effect on model parameter estimates, but substantially affects social welfare calculations such as the cost of living impacts of price changes.

## References

- [1] Ai, Chunrong and Xiaohong Chen (2003), "Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions," *Econometrica*, 71, 1795-1944.
- [2] Banks, James, Richard Blundell, and Arthur Lewbel (1997), "Quadratic Engel curves and Consumer Demand," *Review of Economics and Statistics*, 79, 527-539.
- [3] Beckert, Walter and Richard Blundell (2004), "Invertibility of Nonparametric Stochastic Demand Functions," *Birkbeck Working Papers in Economics and Finance* No. 406.
- [4] Blackorby, Charles, and David Donaldson (1980), "Theoretical Treatment of Indices of Absolute Inequality," *International Economic Review*, 21, 107-136.
- [5] Blackorby, Charles, and David Donaldson (1993), "Adult-Equivalence Scales and the Economic Implementation of Interpersonal Comparisons of Well-Being," *Social Choice and Welfare*, 10, 335-61.
- [6] Blundell, Richard, Xiaohong Chen and Dennis Kristensen (2007), "Nonparametric IV Estimation of Shape-Invariant Engel curves," *Econometrica*, 75, 1613-1669.

- [7] Blundell, Richard, Alan Duncan and Krishna Pendakur (1998), "Semiparametric Estimation of Consumer Demand", *Journal of Applied Econometrics* 13, 435-461.
- [8] Blundell, Richard and Arthur Lewbel (1991), "The Information Content of Equivalence Scales." *Journal of Econometrics* 50, 49-68.
- [9] Blundell, Richard, Panos Pashardes, and Guglielmo Weber (1993), "What do we learn about Consumer Demand Patterns from Micro Data?", *American Economic Review*, vol. 83, 570-597.
- [10] Blundell, Richard and Robin, Jean Marc, (1999), "Estimation in Large and Disaggregated Demand Systems: An Estimator for Conditionally Linear Systems," *Journal of Applied Econometrics*, 14, 209-232.
- [11] Brown, Donald J. and Rosa L. Matzkin, (1998), "Estimation of Nonparametric Functions in Simultaneous Equations Models, with an Application to Consumer Demand," *Cowles Foundation Discussion Papers* 1175, Cowles Foundation, Yale University.
- [12] Brown, Bryan W. and Mary Beth Walker (1989), "The Random Utility Hypothesis and Inference in Demand Systems" *Econometrica* 57, 815-829
- [13] Browning, Martin, Pierre-Andre Chiappori and Arthur Lewbel, (2004) "Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power," *Boston College Working Papers in Economics* 588.
- [14] Browning, Martin and Irena Thomas, (1999) "Prices for the FAMEX," unpublished working paper, Department of Economics, McMaster University.
- [15] Browning, Martin, (2001), *Modelling Commodity Demands and Labour Supply with M-Demands*, unpublished working paper, Institute of Economics, University of Copenhagen.
- [16] Deaton, Angus S., and John Muellbauer (1980), "An Almost Ideal Demand System," *American Economic Review*, vol. 70, 312-326.
- [17] Diewert, Walter Erwin. and Terrence J. Wales (1987), "Flexible Functional Forms and Global Curvature Restrictions," *Econometrica*, 55. 43-68.
- [18] Dominitz, Jeff and Robert P. Sherman (2005), "Some Convergence Theory for Iterative Estimation Procedures with an Application to Semiparametric Estimation," *Econometric Theory*, 21, 838-863.
- [19] Gorman, William M. "Tricks with utility functions", in Artis, M. and A. R. Nobay (eds.), *Essays in Economic Analysis*, Cambridge University Press (1976).
- [20] Gorman, William M. (1981), "Some Engel Curves," in *Essays in the Theory and Measurement of Consumer Behaviour in Honor of Sir Richard Stone*, ed. by Angus Deaton, Cambridge: Cambridge University Press.
- [21] Hansen, Lars, (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029-1054.

- [22] Hildenbrand, Werner, (1994), *Market Demand*, Princeton University Press.
- [23] Lewbel, Arthur (1989). "Household Equivalence Scales and Welfare Comparisons," *Journal of Public Economics*, 39, 377-91.
- [24] Lewbel, Arthur (1991), "The Rank of Demand Systems: Theory and Nonparametric Estimation," *Econometrica*, 59, 711-730.
- [25] Lewbel, Arthur (1996), "Demand Estimation with Expenditure Measurement Errors on the Left and Right Hand Side," *The Review of Economics and Statistics*, 78(4), 718-725.
- [26] Lewbel, Arthur (2001), "Demand Systems with and without Errors," *American Economic Review*, 91, 611-618
- [27] Lewbel, Arthur (2003), "A Rational Rank Four Demand System," *Journal of Applied Econometrics*, 18, 127-135.
- [28] Lewbel, Arthur and Serena Ng, (2005), "Demand Systems with Nonstationary Prices," *The Review of Economics and Statistics*, 87(3), 479-494.
- [29] Matzkin, Rosa L. (2005), "Heterogeneous Choice," for *Advances in Economics and Econometrics*, edited by Richard Blundell, Whitney Newey, and Torsten Persson, vol. 3, chap. 4, 75-110, Cambridge University Press.
- [30] McFadden, Daniel. and Marcel K. Richter (1990), "Stochastic Rationality and Revealed Stochastic Preference," in J. Chipman, D. McFadden, M. K. Richter (eds) *Preferences, Uncertainty, and Optimality, Essays in Honor of Leo Hurwicz*, 161-186, Westview Press: Boulder, CO.
- [31] Pendakur, Krishna (1999) "Estimates and Tests of Base-Independent Equivalence Scales," *Journal of Econometrics*, 88, 1-40.
- [32] Pendakur, Krishna (2002) "Taking Prices Seriously in the Measurement of Inequality," *Journal of Public Economics*, 86, 47-69.
- [33] Pendakur, Krishna and Stefan Sperlich, (2005), "Semiparametric Estimation of Consumer Demand Systems with Almost Partially Linear Price Effects," unpublished working paper, Department of Economics, Simon Fraser University.
- [34] Pollak, Robert A and Wales, Terence J, (1980) "Comparison of the Quadratic Expenditure System and Translog Demand Systems with Alternative Specifications of Demographic Effects," *Econometrica*, 48, 595-612.
- [35] Pollak, Robert A and Wales, Terence J, (1992) *Demand System Specification and Estimation*, Oxford: Oxford University Press.
- [36] Ryan, David and Terry Wales, (1998) "A Simple Method for Imposing Local Curvature in Some Flexible Consumer Demand Systems," *Journal of Business and Economic Statistics*, 16(3), 331-338.

- [37] Stone, John Richard Nicholas (1954), *The Measurement of Consumers' Expenditure and Behaviour in the United Kingdom, 1920-1938*. Cambridge: Cambridge University Press.
- [38] van Soest, Arthur, Arie Kapteyn, and Peter Kooreman, (1993), "Coherency and Regularity of Demand Systems with Equality and Inequality Constraints." *Journal of Econometrics*, 57, 161-188.

## 5 Appendix

This appendix provides Theorems associated with our model and with implicit Marshallian demands in general. First we provide theorems that lead up to the EASI class of demand systems. This is followed by extensions and additional mathematical properties of our models.

### 5.1 Exact Stone Index (ESI) Demands

Ignore preference heterogeneity for now, that is, consider  $\mathbf{z}$  and  $\varepsilon$  as fixed. We will reintroduce them later. Define preferences to be *regular* if they can be represented by a log cost function  $C(\mathbf{p}, u)$  where  $\exp[C(\mathbf{p}, u)]$  is concave, increasing, differentiable and homogeneous of degree one in (unlogged) prices  $\exp(\mathbf{p})$  and is monotonically increasing and differentiable in  $u$ . For a given  $C(\mathbf{p}, u)$  function it may be necessary to restrict the domain of  $\mathbf{p}, u$  to values where  $C(\mathbf{p}, u)$  is regular. Shephard's lemma relates Hicksian (compensated) budget shares to regular cost functions by

$$\mathbf{w} = \boldsymbol{\omega}(\mathbf{p}, u) = \nabla_{\mathbf{p}} C(\mathbf{p}, u).$$

Consider log Stone index deflated expenditures  $x - \mathbf{p}'\mathbf{w}$ . Suppose preferences are represented by a cost function  $C(\mathbf{p}, u)$  that makes  $u = y$  where implicit utility  $y = x - \mathbf{p}'\mathbf{w}$ , which is the log of Stone index deflated expenditures. We call this an Exact Stone Index (ESI) cost function. If we have an ESI cost function, then we can substitute out  $u$  in the Hicks demand functions  $\mathbf{w} = \boldsymbol{\omega}(\mathbf{p}, u)$  to obtain  $\mathbf{w} = \boldsymbol{\omega}(\mathbf{p}, x - \mathbf{p}'\mathbf{w})$ . The name Exact Stone Index is in contrast with the approximate Almost Ideal demand system, which uses  $y = x - \mathbf{p}'\mathbf{w}$  as an approximation to deflating  $x$  by a certain quadratic function of  $\mathbf{p}$ . In an ESI cost function, the Stone index is not an approximation to some true deflator. Instead, the Stone index is the exact correct deflator for  $x$ .

Given an ESI cost function,  $\mathbf{w} = \boldsymbol{\omega}(\mathbf{p}, x - \mathbf{p}'\mathbf{w})$  is an example of an implicit Marshallian demand functions. Using  $x = C(\mathbf{p}, u)$  and Shephard's lemma, we obtain  $u = x - \mathbf{p}'\mathbf{w}$ , and therefore have an ESI cost function, if and only if

$$u = C(\mathbf{p}, u) - \mathbf{p}'[\nabla_{\mathbf{p}} C(\mathbf{p}, u)]. \quad (17)$$

Theorem 1 characterizes the solutions to this equation.

**Theorem 1:** Define  $d(\mathbf{p}, u) \equiv C(\mathbf{p}, u) - u$  and assume  $C(\mathbf{p}, u)$  is a regular cost function. Then  $C(\mathbf{p}, u)$  is an Exact Stone Index (ESI) cost function if and only if  $d(\lambda\mathbf{p} + \mathbf{1}_J\kappa, u) = \lambda d(\mathbf{p}, u) + \kappa$  for any scalars  $\lambda > 0$  and  $\kappa$ .

**Proof of Theorem 1:**  $d(\lambda\mathbf{p}, u) = \lambda d(\mathbf{p}, u)$  is equivalent to  $d(\mathbf{p}, u)$  linearly homogeneous in  $\mathbf{p}$  and  $d(\mathbf{p} + \mathbf{1}_J\kappa, u) = d(\mathbf{p}, u) + \kappa$  is equivalent to  $\exp[d(\mathbf{p}, u)]$  linearly homogeneous in  $\exp(\mathbf{p})$ .



The latter condition is required by cost function regularity. For the former, given equation (17) we have  $C(\mathbf{p}, u) - u = \mathbf{p}' [\nabla_{\mathbf{p}} C(\mathbf{p}, u)]$ . This is equivalent to  $d(\mathbf{p}, u) = \mathbf{p}' [\nabla_{\mathbf{p}} d(\mathbf{p}, u)]$ , which by the Euler Theorem shows that  $d(\mathbf{p}, u)$  is linearly homogeneous in  $\mathbf{p}$ .

Theorem 1 can be equivalently stated as the requirement that  $\exp[d(\mathbf{p}, u)]$  be linearly homogeneous in  $\exp(\mathbf{p})$ , which follows from ordinary cost function homogeneity, and also that  $d(\mathbf{p}, u)$  be linearly homogeneous in  $\mathbf{p}$ , which makes  $u = x - \mathbf{p}'\mathbf{w}$ . A function  $d(\mathbf{p})$  is defined to be, "translatable of degree  $c$ " if  $d(\mathbf{p} + \mathbf{1}_J \kappa) = d(\mathbf{p}, u) + c\kappa$  (see, e.g., Blackorby and Donaldson (1980)), so another way to state Theorem 1 is that  $C(\mathbf{p}, u)$  is an ESI cost function if and only if  $C(\mathbf{p}, u) - u$  is both linearly homogeneous in  $\mathbf{p}$  and translatable of degree one in  $\mathbf{p}$ .

By ordinality of utility, instead of  $x - \mathbf{p}'\mathbf{w} = u$ , we could have defined Stone Index exactness by  $x - \mathbf{p}'\mathbf{w} = h(u)$  for any strictly monotonically increasing function  $h$ , and Theorem 1 would then hold with  $d(\mathbf{p}, u)$  defined by  $C(\mathbf{p}, u) - h(u)$ . There is no gain in generality from doing so, because again by ordinality,  $C(\mathbf{p}, u) - h(u)$  has the same indifference curves, and hence the same Marshallian and implicit Marshallian demand functions, as  $C(\mathbf{p}, h^{-1}(u)) - u = \tilde{C}(\mathbf{p}, u) - u$ . In this representation of the cost function (i.e., in this cardinalization of utility), we then obtain  $x - \mathbf{p}'\mathbf{w} = u$  as before.

The following Corollary and theorem illustrates the restrictiveness of ESI demands.

**Corollary 1:** Assume  $C(\mathbf{p}, u)$  is an ESI cost function. Then Hicksian demands  $\omega(\mathbf{p}, u) = \nabla_{\mathbf{p}} C(\mathbf{p}, u)$  are homogeneous of degree zero in  $\mathbf{p}$  and in  $\exp(\mathbf{p})$ .

**Proof of Corollary 1:** Homogeneity of degree zero in  $\exp(\mathbf{p})$  follows from ordinary cost function regularity, and homogeneity of degree zero in  $\mathbf{p}$  follows from the property that the derivative of a linearly homogeneous function, in this case  $d(\mathbf{p}, u)$ , must be homogeneous of degree zero.

Cost function regularity implies that Hicksian demands do not change when all prices are scaled by a constant factor. Corollary 1 shows that ESI requires that Hicksian demands also not change when all prices undergo the same power transformation, for example, when all prices are squared, and also rules out Hicksian demands that are linear in  $\mathbf{p}$ . Relative prices can change dramatically when all prices are squared, so this additional homogeneity condition binds in economically implausible ways.

**Theorem 2:** Consider the case where  $J = 2$ . Then  $d(\lambda p_1 + \kappa, \lambda p_2 + \kappa, u) = \lambda d(p_1, p_2, u) + \kappa$  for any scalars  $\lambda > 0$  and  $\kappa$  if and only if  $d(p_1, p_2, u) = m(u)p_1 + [1 - m(u)]p_2$  for some function  $m(u)$ .

**Proof of Theorem 2:** Since  $d(\lambda p_1 + \kappa, \lambda p_2 + \kappa, u) = \lambda d(p_1, p_2, u) + \kappa$  holds for any  $\lambda$  and  $\kappa$ , let  $\lambda = 1/(p_2 - p_1)$  and  $\kappa = -p_1/(p_2 - p_1)$  to obtain  $d(0, 1, u) = [d(p_1, p_2, u)/(p_2 - p_1)] - [p_1/(p_2 - p_1)]$ . The theorem holds with  $m(u) = 1 - d(0, 1, u)$ .

Theorem 2 shows that with  $J = 2$  goods, ESI cost functions must have the linear in  $\mathbf{p}$  form  $C(\mathbf{p}, u) = u + \mathbf{p}'\mathbf{m}(u)$  where  $\mathbf{m}(u)$  is a  $J$ -vector of functions that satisfy  $\mathbf{m}(u)'\mathbf{1}_J = 1$ . This is an ESI cost function for any  $J$ , but has the unattractive feature that it has Hicksian budget shares  $\mathbf{m}(u)$  that are independent of  $\mathbf{p}$ .

ESI cost functions having more than  $J = 2$  goods can be nonlinear in  $\mathbf{p}$ . One example is

$$C(\mathbf{p}, u) = u + \mathbf{p}'\mathbf{m}(u) + [\mathbf{p}'\mathbf{M}(u)\mathbf{p}]^{1/2}$$

where  $\mathbf{m}(u)$  is as above and  $\mathbf{M}(u)$  is a  $J$  by  $J$  symmetric matrix-valued function of  $u$  with  $\mathbf{M}(u)\mathbf{1}_J = \mathbf{0}_J$ . As this example shows, ESI log cost functions can be nonlinear in log prices, however, by Corollary 1 ESI log cost functions cannot be quadratic in log prices, and therefore ESI budget shares cannot be linear in log prices.

## 5.2 Exact Affine Stone Index (EASI) Demands

As shown in the previous section, ESI Hicks and implicit Marshallian budget shares must possess the unattractive feature of not changing when all prices are squared, and also must either be independent of  $\mathbf{p}$  or nonlinear in  $\mathbf{p}$ . We can avoid these drawbacks by generalizing the expression for real expenditures, thereby relaxing the homogeneity restrictions required by ESI demands. Specifically, instead of imposing the ESI restriction that some cardinalization of  $u$  equal  $x - \mathbf{p}'\mathbf{w}$ , define Exact Affine Stone Index (EASI) cost functions to be cost functions that have the property that a cardinalization of  $u$  exists that equals an affine transformation of  $x - \mathbf{p}'\mathbf{w}$ . In this cardinalization  $u = y$  where implicit utility  $y$  is defined by  $y = (x - \mathbf{p}'\mathbf{w})\tilde{s}(\mathbf{p}) - \tilde{t}(\mathbf{p})$  for some functions  $\tilde{s}(\mathbf{p})$  and  $\tilde{t}(\mathbf{p})$ . For reasons that will be clear shortly, it is more convenient to express this relationship as

$$u = y = \frac{x - \mathbf{p}'\mathbf{w} - t(\mathbf{p})}{1 + s(\mathbf{p})} \quad (18)$$

for some functions  $s(\mathbf{p})$  and  $t(\mathbf{p})$ . We choose affine transforms for their simplicity, in particular, for now we wish to avoid having to solve nonlinear transformations to obtain cost functions (the next section in this appendix discusses more general implicit Marshallian demand functions that may require nonlinear transformation).

Equation (18) implies

$$x - u - \mathbf{p}'\mathbf{w} = t(\mathbf{p}) + s(\mathbf{p})u,$$

and therefore any EASI cost function  $C(\mathbf{p}, u)$  must satisfy

$$C(\mathbf{p}, u) - u - \mathbf{p}'[\nabla_{\mathbf{p}}C(\mathbf{p}, u)] = t(\mathbf{p}) + s(\mathbf{p})u \quad (19)$$

The following two theorems characterize EASI cost functions, and hence the solutions to equation (19), and provide a convenient way to construct such functions.

**Theorem 3:** Assume  $C(\mathbf{p}, u)$  is regular. Let  $d(\mathbf{p}, u) \equiv C(\mathbf{p}, u) - u$ . There exists functions  $t(\mathbf{p})$  and  $s(\mathbf{p})$  that make equation (18) hold, and hence make  $C(\mathbf{p}, u)$  an EASI cost function, if and only if  $d(\mathbf{p} + \mathbf{1}_J\kappa, u) = d(\mathbf{p}, u) + \kappa$  and  $\nabla_u^2 d(\lambda\mathbf{p}, u) = \lambda \nabla_u^2 d(\mathbf{p}, u)$  for any scalars  $\lambda > 0$  and  $\kappa$ . When equation (18) holds, the functions  $t(\mathbf{p})$  and  $s(\mathbf{p})$  satisfy  $t(\mathbf{p} + \mathbf{1}_J\kappa) = t(\mathbf{p}) + \kappa$  and  $s(\mathbf{p} + \mathbf{1}_J\kappa) = s(\mathbf{p}) + \kappa$ .

**Proof of Theorem 3:** The condition that  $d(\mathbf{p} + \mathbf{1}_J\kappa, u) = d(\mathbf{p}, u) + \kappa$  is equivalent to  $\exp[d(\mathbf{p}, u)]$  linearly homogeneous in  $\exp(\mathbf{p})$ , which is required by cost function regularity. Equation (19) is equivalent to  $d(\mathbf{p}, u) - t(\mathbf{p}) - s(\mathbf{p})u = \mathbf{p}'[\nabla_{\mathbf{p}}d(\mathbf{p}, u)]$ . Taking derivatives of this expres-

sion with respect to  $u$  gives  $\nabla_u d(\mathbf{p}, u) - s(\mathbf{p}) = \mathbf{p}' [\nabla_{\mathbf{p}} \nabla_u d(\mathbf{p}, u)]$ , so  $\nabla_u^2 d(\mathbf{p}, u) = \mathbf{p}' [\nabla_{\mathbf{p}} \nabla_u^2 d(\mathbf{p}, u)]$ , which by the Euler Theorem shows that  $\nabla_u^2 d(\mathbf{p}, u)$  is linearly homogeneous in  $\mathbf{p}$ , that is,  $\nabla_u^2 d(\lambda \mathbf{p}, u) = \lambda \nabla_u^2 d(\mathbf{p}, u)$ . In the other direction, if  $\nabla_u^2 d(\mathbf{p}, u)$  is linearly homogeneous then  $\nabla_u^2 d(\mathbf{p}, u) = \mathbf{p}' [\nabla_{\mathbf{p}} \nabla_u^2 d(\mathbf{p}, u)]$ . Integrate this equation twice with respect to  $u$  (defining  $s(\mathbf{p})$  be the constant of integration for the first integration and  $t(\mathbf{p})$  to be the constant of integration for the second integration) and continue following the above steps in reverse to obtain equation (19). Given exactness, Equation (19) implies  $t(\mathbf{p}) + s(\mathbf{p})u = u - C(\mathbf{p}, u) + \mathbf{p}'\omega(\mathbf{p}, u)$ . Ordinary homogeneity of cost functions and Hicksian demands requires that  $\exp[u - C(\mathbf{p}, u) + \mathbf{p}'\omega(\mathbf{p}, u)]$  be linearly homogeneous in  $\exp(\mathbf{p})$ , so  $t(\mathbf{p}) + s(\mathbf{p})u$  must be linearly homogeneous in  $\exp(\mathbf{p})$ . Since  $u$  can vary holding  $\mathbf{p}$  fixed, this requires that  $t(\mathbf{p})$  and  $s(\mathbf{p})$  must each be linearly homogeneous in  $\exp(\mathbf{p})$ , and therefore that  $t(\mathbf{p} + \mathbf{1}_J \kappa) = t(\mathbf{p}) + \kappa$  and  $s(\mathbf{p} + \mathbf{1}_J \kappa) = s(\mathbf{p}) + \kappa$ .

**Theorem 4:** Assume the function  $\bar{C}(\mathbf{p}, u)$  satisfies the homogeneity conditions required to be an ESI log cost function as given by Theorem 1. Assume  $C(\mathbf{p}, u) = \bar{C}(\mathbf{p}, u) + T(\mathbf{p}) + S(\mathbf{p})u$  is a regular log cost function. Then  $C(\mathbf{p}, u)$  is an EASI log cost function, so equation (18) holds, with  $t(\mathbf{p}) = T(\mathbf{p}) - \mathbf{p}' [\nabla_{\mathbf{p}} T(\mathbf{p})]$  and  $s(\mathbf{p}) = S(\mathbf{p}) - \mathbf{p}' [\nabla_{\mathbf{p}} S(\mathbf{p})]$ .

**Proof of Theorem 4:**  $C(\mathbf{p}, u) = \bar{C}(\mathbf{p}, u) + T(\mathbf{p}) + S(\mathbf{p})u$ , so

$$\begin{aligned} C(\mathbf{p}, u) - \mathbf{p}' [\nabla_{\mathbf{p}} C(\mathbf{p}, u)] &= \bar{C}(\mathbf{p}, u) - \mathbf{p}' [\nabla_{\mathbf{p}} \bar{C}(\mathbf{p}, u)] + T(\mathbf{p}) - \mathbf{p}' [\nabla_{\mathbf{p}} T(\mathbf{p})] + \\ &\quad S(\mathbf{p})u - \mathbf{p}' [\nabla_{\mathbf{p}} S(\mathbf{p})] u \\ &= u + T(\mathbf{p}) - \mathbf{p}' [\nabla_{\mathbf{p}} T(\mathbf{p})] + S(\mathbf{p})u - \mathbf{p}' [\nabla_{\mathbf{p}} S(\mathbf{p})] u \end{aligned}$$

where the right side of the second equality follows from Stone index exactness for  $\bar{C}(\mathbf{p}, u)$  from Theorem 1. Comparing this equation to (19) proves the result.

Now consider how observed variables  $\mathbf{z}$  that affect preferences and how random unobserved preference heterogeneity parameters  $\varepsilon$  can be incorporated into the ESI and EASI cost functions. These cost functions are formulated above in terms of the functions  $\bar{C}(\mathbf{p}, u)$ ,  $T(\mathbf{p})$  and  $S(\mathbf{p})$ . Incorporation of observed heterogeneity is completely general in that there are no restrictions on how observed characteristics  $\mathbf{z}$  enter the ESI cost function  $\bar{C}(\mathbf{p}, u, \mathbf{z})$  or the price functions  $T(\mathbf{p}, \mathbf{z})$  and  $S(\mathbf{p}, \mathbf{z})$ . However, the incorporation of unobserved heterogeneity is restricted by the fact that implicit utility  $y$  must be a function of observables only. This can be accomplished by letting  $\bar{C}$  depend arbitrarily on  $\varepsilon$  and having the price functions  $T(\mathbf{p}, \mathbf{z})$  and  $S(\mathbf{p}, \mathbf{z})$  be independent of  $\varepsilon$ . The incorporation of both forms of heterogeneity then results in the general class of EASI cost functions

$$C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = \bar{C}(\mathbf{p}, u, \mathbf{z}, \varepsilon) + T(\mathbf{p}, \mathbf{z}) + S(\mathbf{p}, \mathbf{z})u, \quad (20)$$

where  $\bar{C}$  is a regular ESI cost function and  $\exp(T(\mathbf{p}, \mathbf{z}))$  and  $\exp(S(\mathbf{p}, \mathbf{z}))$  are both homogeneous of degree zero in  $\exp(\mathbf{p})$ . Based on equation (19) and Theorems 3 and 4, one very flexible choice for  $\bar{C}$  is  $\bar{C}(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{p}' \mathbf{m}(u, \mathbf{z}, \varepsilon)$ , which allows for nonseparability with respect to the unobserved heterogeneity parameters  $\varepsilon$ , and results in an EASI form

$$C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{p}' \mathbf{m}(u, \mathbf{z}, \varepsilon) + T(\mathbf{p}, \mathbf{z}) + S(\mathbf{p}, \mathbf{z})u. \quad (21)$$

We will use this specification for illustration in Theorems 6 through 8.

Replacing  $\mathbf{m}(u, \mathbf{z}, \varepsilon)$  with  $\mathbf{m}(u, \mathbf{z}) + \varepsilon$  in equation (21) yields Hicksian (and hence implicit

Marshallian) budget shares that are additively separable in  $\varepsilon$ , which is convenient for estimation. This results in the EASI cost function form

$$C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{p}'\mathbf{m}(u, \mathbf{z}) + T(\mathbf{p}, \mathbf{z}) + S(\mathbf{p}, \mathbf{z})u + \mathbf{p}'\varepsilon, \quad (22)$$

where  $\mathbf{1}'\mathbf{m}(u, \mathbf{z}) = 1$  and  $\mathbf{1}'\varepsilon = 0$ . We use this form as the basis for our empirical investigation and for most discussion in the text.

### 5.3 General Cost Functions For Implicit Marshallian Demands

Instead of Stone index related constructions, we could more generally define implicit utility by

$$y = \frac{x - G(\mathbf{w}, \mathbf{p}, \mathbf{z})}{1 + s(\mathbf{p}, \mathbf{z})} \quad (23)$$

for any deflator function  $G$ . Setting  $y = u$  will then require

$$C(\mathbf{p}, u, \mathbf{z}) = [1 + s(\mathbf{p}, \mathbf{z})]u + G(\nabla_{\mathbf{p}}C(\mathbf{p}, u, \mathbf{z}), \mathbf{p}, \mathbf{z}) \quad (24)$$

and the resulting implicit Marshallian demand functions are

$$\mathbf{w} = \omega \left[ \mathbf{p}, \mathbf{z}, \frac{x - G(\mathbf{w}, \mathbf{p}, \mathbf{z})}{1 + s(\mathbf{p}, \mathbf{z})} \right]. \quad (25)$$

where  $\omega(\mathbf{p}, \mathbf{z}, u) = \nabla_{\mathbf{p}}C(\mathbf{p}, u, \mathbf{z})$ . At this level of generality,  $u = y$  is possible for almost any log cost function  $C(\mathbf{p}, u, \mathbf{z})$  even with  $s(\mathbf{p}, \mathbf{z}) = 0$  as the following Theorem shows.

**Theorem 5:** Assume  $C(\mathbf{p}, u, \mathbf{z})$  is regular. A sufficient condition for existence of a function  $G(\mathbf{w}, \mathbf{p}, \mathbf{z})$  such that  $u = x - G(\mathbf{w}, \mathbf{p}, \mathbf{z})$  is the existence of some scalar valued function of  $\nabla_{\mathbf{p}}C(\mathbf{p}, u, \mathbf{z})$  and  $\mathbf{p}$  that is strictly monotonic in  $u$ .

**Proof of Theorem 5:** Let  $R$  be a scalar valued function of  $\nabla_{\mathbf{p}}C(\mathbf{p}, u, \mathbf{z})$  and  $\mathbf{p}$  that is strictly monotonic in  $u$ . Since  $R(\mathbf{p}, \mathbf{z}, \omega(\mathbf{p}, \mathbf{z}, u))$  is strictly monotonic in  $u$ , it can be inverted to obtain  $u = S(\mathbf{p}, \mathbf{z}, R)$ , and we may define a function  $G$  by  $G(\mathbf{w}, \mathbf{p}, \mathbf{z}) = C(\mathbf{p}, S(\mathbf{p}, \mathbf{z}, R(\mathbf{w}, \mathbf{p}, \mathbf{z})), \mathbf{z}) - S(\mathbf{p}, \mathbf{z}, R(\mathbf{w}, \mathbf{p}, \mathbf{z}))$ .

Theorem 5 provides a very weak sufficient condition for existence of implicit Marshallian demands  $\mathbf{w} = \omega[\mathbf{p}, \mathbf{z}, x - G(\mathbf{w}, \mathbf{p}, \mathbf{z})]$  for some function  $G$ . For example, this condition is satisfied if any good or combination of goods has a budget share that is strictly increasing or strictly decreasing in total expenditures (and hence in utility). Moreover, this condition is sufficient but not necessary, e.g., if preferences are homothetic so  $C(\mathbf{p}, u, \mathbf{z}) = u + t(\mathbf{p}, \mathbf{z})$  for some function  $t$ , then even though the sufficient condition need not hold now, we can define  $G(\mathbf{w}, \mathbf{p}, \mathbf{z}) = -t(\mathbf{p}, \mathbf{z})$  to obtain implicit marshallian demands. So the only kinds of situations where  $G$  might fail to exist is when all budget shares are independent of  $u$  for some, but not all, ranges of values of  $u$ .

To give an example of Theorem 5, suppose that the Hicksian budget share function for good 1,  $w_1 = \omega_1(\mathbf{p}, \mathbf{z}, u)$ , is invertible in  $u$ , so  $u = \omega_1^{-1}(\mathbf{p}, \mathbf{z}, w_1)$ . Then a deflator  $G$  that makes  $u = y$  is  $G(\mathbf{w}, \mathbf{p}, \mathbf{z}) = C(\mathbf{p}, \omega_1^{-1}(\mathbf{p}, \mathbf{z}, w_1), \mathbf{z}) - \omega_1^{-1}(\mathbf{p}, \mathbf{z}, w_1)$ .

The function  $G$  satisfying equation (24) is not unique, e.g., it can be freely translated by an ordinal transformation of  $u$ , and in the previous example, different invertible budget shares could be used to obtain different expressions for  $G$ . For applications, the main point of implicit Marshallian demands is not existence or uniqueness, but rather convenience for demand estimation, which is why this paper focuses on examples where  $G$  has a simple closed form, namely, a translation of the log Stone index.

## 5.4 Generalized EASI cost functions

A generalization of the EASI class of cost functions (22) is

$$x = C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{c}(\mathbf{p}, \mathbf{z})' \mathbf{m}(u, \mathbf{z}, \varepsilon) + T(\mathbf{p}, \mathbf{z}) + S(\mathbf{p}, \mathbf{z})u \quad (26)$$

where  $\mathbf{c}(\mathbf{p}, \mathbf{z})$  is a  $J$ -vector valued function. The Hicksian budget shares for this class are

$$\mathbf{w} = \omega(\mathbf{p}, u, \mathbf{z}, \varepsilon) = \nabla_{\mathbf{p}} \mathbf{c}(\mathbf{p}, \mathbf{z})' \mathbf{m}(u, \mathbf{z}, \varepsilon) + \nabla_{\mathbf{p}} T(\mathbf{p}, \mathbf{z}) + \nabla_{\mathbf{p}} S(\mathbf{p}, \mathbf{z})u$$

Solving this expression for  $\mathbf{m}(u, \mathbf{z}, \varepsilon)$  and substituting the result into (26) gives

$$x = u + \mathbf{c}(\mathbf{p}, \mathbf{z})' [\nabla_{\mathbf{p}} \mathbf{c}(\mathbf{p}, \mathbf{z})']^{-1} [\mathbf{w} - \nabla_{\mathbf{p}} T(\mathbf{p}, \mathbf{z}) - \nabla_{\mathbf{p}} S(\mathbf{p}, \mathbf{z})u] + T(\mathbf{p}, \mathbf{z}) + S(\mathbf{p}, \mathbf{z})u$$

Now solve this expression for  $u$  and call the result  $y$  to get

$$y = \frac{x - T(\mathbf{p}, \mathbf{z}) - \mathbf{c}(\mathbf{p}, \mathbf{z})' [\nabla_{\mathbf{p}} \mathbf{c}(\mathbf{p}, \mathbf{z})']^{-1} [\mathbf{w} - \nabla_{\mathbf{p}} T(\mathbf{p}, \mathbf{z})]}{1 + S(\mathbf{p}, \mathbf{z}) - \mathbf{c}(\mathbf{p}, \mathbf{z})' [\nabla_{\mathbf{p}} \mathbf{c}(\mathbf{p}, \mathbf{z})']^{-1} \nabla_{\mathbf{p}} S(\mathbf{p}, \mathbf{z})}$$

which we can write as

$$y = \frac{x - \bar{\mathbf{c}}(\mathbf{p}, \mathbf{z})' \mathbf{w} - \bar{t}(\mathbf{p}, \mathbf{z})}{1 + \bar{s}(\mathbf{p}, \mathbf{z})}$$

for appropriately defined functions  $\bar{t}$ ,  $\bar{s}$ , and  $J$ -vector  $\bar{\mathbf{c}}$ , and with this definition of  $y$  we obtain implicit Marshallian demands

$$\mathbf{w} = \nabla_{\mathbf{p}} \mathbf{c}(\mathbf{p}, \mathbf{z})' \mathbf{m}(y) + \nabla_{\mathbf{p}} T(\mathbf{p}, \mathbf{z}) + \nabla_{\mathbf{p}} S(\mathbf{p}, \mathbf{z})y$$

This is a generalization of EASI demands where  $y$  is an affine transform of  $x - \bar{\mathbf{c}}(\mathbf{p}, \mathbf{z})' \mathbf{w}$  instead of an affine transform of  $x - \mathbf{p}' \mathbf{w}$ . This generalization is useful for the closure under unit scaling property discussed later, and could be used to introduce additional interactions among  $y$ ,  $\mathbf{p}$ , and  $\mathbf{z}$  if required.

## 5.5 Global Regularity of EASI Demands

**Theorem 6:** Assume a log cost function in the general EASI form of equation (21). Sufficient conditions for the cost function to be regular everywhere on the support of the data are  $\mathbf{1}'_J \mathbf{m}(u, \mathbf{z}, \varepsilon) = 1$ ,  $\exp [T(\mathbf{p}, \mathbf{z})]$  and  $\exp [S(\mathbf{p}, \mathbf{z})]$  homogeneous of degree zero in  $\exp(\mathbf{p})$ ,  $T(\mathbf{p}, \mathbf{z})$  and  $S(\mathbf{p}, \mathbf{z})$  concave and differentiable in  $\mathbf{p}$ ,  $1 + \inf [\mathbf{p}' \nabla_u \mathbf{m}(u, \mathbf{z}, \varepsilon)] > \inf [S(\mathbf{p}, \mathbf{z})]$  and  $\inf(u) > 0$ , where these infima are over the supports of  $\mathbf{p}$ ,  $u$ ,  $\mathbf{z}$ ,  $\varepsilon$ .

**Proof of Theorem 6:** Homogeneity of the cost function follows from  $\mathbf{1}'_j \mathbf{m}(u, \mathbf{z}, \varepsilon) = 1$  and homogeneity of  $\exp T(\mathbf{p}, \mathbf{z})$  and  $\exp S(\mathbf{p}, \mathbf{z})$ . Concavity of  $T(\mathbf{p}, \mathbf{z})$  and  $S(\mathbf{p}, \mathbf{z})$  and having  $u$  positive means that  $T(\mathbf{p}, \mathbf{z}) + S(\mathbf{p}, \mathbf{z})u$  is concave, which suffices for cost function concavity. For monotonicity  $\nabla_u C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = 1 + \mathbf{p}' \nabla_u \mathbf{m}(u, \mathbf{z}, \varepsilon) + S(\mathbf{p}, \mathbf{z})$ , which is positive given  $1 + \inf[\mathbf{p}' \nabla_u \mathbf{m}(u, \mathbf{z}, \varepsilon)] > -\inf[S(\mathbf{p}, \mathbf{z})]$ .

By Theorem 6, EASI demands can be globally (i.e., over the entire supports of the data) regular if the functions  $T$  and  $S$  satisfy homogeneity and concavity conditions, and if  $T$ ,  $S$ ,  $\mathbf{m}$  and the supports of the data are suitably bounded. The sufficient, though stronger than necessary, conditions in Theorem 6 include  $\inf(u) > 0$  (recall that  $u$  here is expressed in the cardinalization that makes  $u = y$ ). This condition can always be satisfied without loss of generality by suitably adjusting the units used to measure expenditures. To see this, define  $K_* = \inf[\mathbf{p}' \mathbf{m}(\tilde{u}, \mathbf{z}, \varepsilon) + T(\mathbf{p}, \mathbf{z})]$  where this infimum is over all  $\mathbf{p}, \mathbf{z}, \varepsilon$  on their support, and over all  $\tilde{u} \geq 0$ . If  $K_*$  exists,  $\inf[S(\mathbf{p}, \mathbf{z})] > -1$  and  $\inf(x) > 0$ , then dividing the units expenditures are measured in by  $\exp(K_*)$  makes the log of these rescaled expenditures  $x$  satisfy  $\inf(x) > K_*$ . This then suffices to make  $\inf(u) > 0$  since, using equation (2),  $u = [x - \mathbf{p}' \mathbf{m}(u, \mathbf{z}, \varepsilon) + T(\mathbf{p}, \mathbf{z})] / [1 + S(\mathbf{p}, \mathbf{z})] \geq [\inf(x) - K_*] / [1 + \inf(S(\mathbf{p}, \mathbf{z}))] > 0$ .

In our empirical model based on the cost function (5),  $T(\mathbf{p}, \mathbf{z})$  and  $S(\mathbf{p}, \mathbf{z})$  satisfy homogeneity, are bounded given boundedness of prices and are concave if  $\sum_{l=0}^L z_l \mathbf{A}_l$  and  $\mathbf{B}$  are negative semidefinite. Let  $c_r = \mathbf{p}' \mathbf{b}_r$ , so boundedness of the support of  $\mathbf{p}$  means that each  $c_r$  lies in some interval. For monotonicity in our empirical model, assume  $y$  is everywhere positive, let  $\lambda_0 = [\inf(x) - \sup(\mathbf{p}' \mathbf{1}_j - \sum_{l=0}^L z_l \mathbf{p}' \mathbf{A}_l \mathbf{p} / 2)] / [1 + \sup(-\mathbf{p}' \mathbf{B} \mathbf{p})]$  and let  $\lambda_1 = \sup(x)$ . It follows from equation (8) that  $\lambda_0 \leq y \leq \lambda_1$ . Let  $\tau = -\inf[1 + \mathbf{p}' \mathbf{D} \mathbf{z} + \mathbf{p}' \mathbf{B} \mathbf{p} / 2]$ . Since  $c_r = \mathbf{p}' \mathbf{b}_r$ , boundedness of the support of  $\mathbf{p}$  means that each  $c_r$  lies in some interval. Then the sufficient monotonicity condition in Theorem 6 is satisfied if  $\sum_{r=0}^5 c_r r y^{r-1} > \tau$  holds for all  $c_r$  on their supports and all  $y$  in the interval  $\lambda_0 \leq y \leq \lambda_1$ . In particular, if quantities are scaled so  $\mathbf{p}$  is nonnegative, and recalling that  $\lambda_0 > 0$ , a sufficient condition is existence of a finite lower bound on the slope of every Engel curve at every point. More generally, monotonicity requires that downward sloping portions of Engel curves for any good need to be sufficiently offset by upward slopes of other Engel curves.

In practice, imposing the restrictions on parameter values that are sufficient but not necessary to guarantee global regularity of demand systems can induce undesirable restrictions on flexibility. See, e.g., Diewert and Wales (1987) and Ryan and Wales (1998). In our empirical application we therefore follow the common practice of estimating without imposing the inequality constraints associated with regularity, and afterwards check for regularity in the neighborhood of the data.

## 5.6 Coherency, Invertibility, and the Distribution of Errors

Coherency of a structural model is defined as the property that, for each value of the exogenous variables and errors, there exists a unique corresponding value of the endogenous variables. Invertibility is the property that a unique value of errors is associated with each value of the endogenous and exogenous variables. Coherency is essentially the condition that the model is fully specified, in the sense that the joint distribution of the endogenous variables is completely determined by the joint distribution of the errors and exogenous variables. Invertibility is the condition that permits

uniquely recovering estimates of the errors given estimates of the model parameters. Since our model takes the errors to be random utility parameters, we will require invertibility to do policy analyses that depend upon these random utility parameters.

In the demand system context (see, e.g., van Soest, Kapteyn, and Kooreman (1993), Brown and Matzkin (1998) and Beckert and Blundell (2004)), coherency requires that a unique value of budget shares  $\mathbf{w}$  be associated with each possible value of  $\mathbf{p}, x, \mathbf{z}, \varepsilon$ , and invertibility requires that a unique value of errors  $\varepsilon$  be associated with each possible value of  $\mathbf{p}, x, \mathbf{z}, \mathbf{w}$ .

**Theorem 7:** Assume a log cost function in the EASI class (21) that is regular. Assume  $\tilde{\mathbf{m}} = \mathbf{m}(u, \mathbf{z}, \varepsilon)$  is invertible in  $\varepsilon$ , so we may write  $\varepsilon = \mathbf{m}^{-1}(u, \mathbf{z}, \tilde{\mathbf{m}})$ . Then the resulting budget share demand functions are coherent and invertible.

**Proof of Theorem 7:** Invertibility of this class is established by observing that by equation (3)

$$\varepsilon = \mathbf{m}^{-1}(y, \mathbf{z}, \mathbf{w} - [\nabla_{\mathbf{p}}T(\mathbf{p})] - [\nabla_{\mathbf{p}}S(\mathbf{p})]y)$$

where  $y$  is given by equation (4). This uniquely defines  $\varepsilon$  as a function of  $\mathbf{p}, x, \mathbf{z}, \mathbf{w}$ . Coherency follows from  $C(\mathbf{p}, u, \mathbf{z}, \varepsilon)$  having all of the properties of a regular cost function, which ensures existence of Marshallian demands (treating  $\varepsilon$  as preference parameters) and hence of a unique value of  $\mathbf{w}$  associated with each possible value of  $\mathbf{p}, x, \mathbf{z}, \varepsilon$ , even though we do not have a closed form analytic expression for it.

Our functional form for empirical work, the cost function (5), satisfies the conditions of Theorem 7 assuming regularity holds on the support of  $\mathbf{p}, x, \mathbf{z}, \varepsilon$  (which, e.g., is satisfied given Theorem 6). For that model  $\mathbf{m}$  is invertible with, by equation (6),  $\mathbf{m}^{-1}(u, \mathbf{z}, m) = \mathbf{m} - \left(\sum_{r=0}^5 \mathbf{b}_r u^r\right) - \mathbf{Cz} - \mathbf{Dzu}$ . More generally, invertibility is satisfied if  $\mathbf{m}(y, \mathbf{z}, \varepsilon) = \mathbf{Bn}(y, \mathbf{z}) + \varepsilon$  for any parameter matrix  $\mathbf{B}$  and functions  $\mathbf{n}$ .

For consistency of our estimator it was assumed that the distribution of  $\varepsilon$  is mean independent of  $\mathbf{p}, x, \mathbf{z}$ , and our 3SLS estimator is efficient (relative to GMM) if  $\varepsilon$  is homoskedastic. A sufficient condition for these to hold is that  $\varepsilon$  be distributed independently of  $\mathbf{p}, x, \mathbf{z}$ . Having  $\varepsilon$  independent of  $\mathbf{p}$  is also appropriate for interpreting  $\varepsilon$  as preference heterogeneity parameters. Having  $\varepsilon$  be distributed independently of  $\mathbf{p}, x, \mathbf{z}$  requires that the support of  $\varepsilon$  not depend on  $\mathbf{p}, x, \mathbf{z}$ . Theorem 7 does not impose this condition.

**Theorem 8:** Assume a log cost function in the EASI class (2). Then cost function regularity, the constraint that budget shares lie between zero and one, coherency, invertibility, and  $\varepsilon$  distributed independently of  $\mathbf{p}, x, \mathbf{z}$  can all hold simultaneously.

**Proof of Theorem 8:** Consider equation (2) cost functions with  $\mathbf{m}(u, \mathbf{z}, \varepsilon) = \mathbf{Bn}(y, \mathbf{z}) + \varepsilon$  for some matrix of parameters  $\mathbf{B}$  and vector of functions  $\mathbf{n}(y, \mathbf{z})$ . Our empirical model (5) is an example of this form. Then

$$\mathbf{w} = \mathbf{Bn}(y, \mathbf{z}) + \nabla_{\mathbf{p}}T(\mathbf{p}, \mathbf{z}) + \nabla_{\mathbf{p}}S(\mathbf{p}, \mathbf{z})y + \varepsilon. \quad (27)$$

Consider bounding each element of  $\mathbf{Bn}(y, \mathbf{z}) + \nabla_{\mathbf{p}}T(\mathbf{p}, \mathbf{z}) + \nabla_{\mathbf{p}}S(\mathbf{p}, \mathbf{z})y$  to lie inside some interval  $[\tau, 1 - \tau]$  for every value of  $y$  in some positive interval  $[\lambda_0, \lambda_1]$  and for every value of  $\mathbf{p}, \mathbf{z}$  on their

support, where  $0 < \tau < 1/2$ . The interval  $[\lambda_0, \lambda_1]$  can be made arbitrarily large by restricting the range of values  $\mathbf{n}(y, \mathbf{z})$  and  $\nabla_{\mathbf{p}}S(\mathbf{p}, \mathbf{z})$  can take on. Let  $\varepsilon$  be drawn from any distribution that satisfies  $\varepsilon' \mathbf{1}_J = 0$  and has the absolute value of each element of  $\varepsilon$  be less than  $\tau$ . Let  $x$  be drawn from any distribution that has support sufficiently bounded to make  $\tilde{y}$  defined by

$$\tilde{y} = \frac{x - \mathbf{p}'\tilde{\mathbf{w}} - T(\mathbf{p}, \mathbf{z}) + \mathbf{p}' [\nabla_{\mathbf{p}}T(\mathbf{p}, \mathbf{z})]}{1 + S(\mathbf{p}, \mathbf{z}) - \mathbf{p}' [\nabla_{\mathbf{p}}S(\mathbf{p}, \mathbf{z})]} \quad (28)$$

lie in  $[\lambda_0, \lambda_1]$  for every value of  $\mathbf{p}, \mathbf{z}$  on their support and for every value of  $\tilde{\mathbf{w}}$  on the unit simplex. Note that, as in the second part of Theorem 6, the units  $x$  is measured in can be adjusted to move the location of the support of  $x$  and hence of  $\tilde{y}$ . It then follows that  $y$  given by equation (4) will lie in  $[\lambda_0, \lambda_1]$ , and then by construction  $\mathbf{w}$  defined by equation (27) will lie in the unit simplex.

Given the cost function (21) with  $\mathbf{m}(u, \mathbf{z}, \varepsilon) = \mathbf{B}\mathbf{n}(u, \mathbf{z}) + \varepsilon$ , and the assumptions already made above, the remaining restrictions required for global regularity in Theorem 6 are  $\mathbf{1}'_J \mathbf{B}\mathbf{n}(u, \mathbf{z}) = 1$ ,  $\exp T(\mathbf{p}, \mathbf{z})$  and  $\exp S(\mathbf{p}, \mathbf{z})$  homogeneous of degree zero in  $\exp(\mathbf{p})$ ,  $T(\mathbf{p}, \mathbf{z})$  and  $S(\mathbf{p}, \mathbf{z})$  concave and differentiable in  $\mathbf{p}$ , and  $1 + \inf[\mathbf{p}'\mathbf{B}\nabla_u \mathbf{n}(u, \mathbf{z})] > \inf[S(\mathbf{p}, \mathbf{z})]$ . Given that the support of  $u$  lies inside an interval  $[\lambda_0, \lambda_1]$  for any value  $\varepsilon$  can take on, we obtain global regularity for any  $\varepsilon$  on its support as long as these conditions (which do not depend directly on  $\varepsilon$ ) hold for every  $\mathbf{p}, \mathbf{z}$  on their support and every  $u$  in the interval  $[\lambda_0, \lambda_1]$ . We then also obtain coherency and invertibility, because the conditions of Theorem 7 all hold for this example.

None of the conditions imposed in the construction of the cost function and the distribution of  $\varepsilon$  in Theorem 8 conflict with assumptions imposed for consistency of the the 3SLS estimates of our exact, symmetry restricted empirically estimated EASI demand functions. For example, the proof of Theorem 8 allows each  $\varepsilon$  to have conditional or unconditional mean zero, because it permits  $\varepsilon$  to be drawn from any distribution that satisfies  $\varepsilon' \mathbf{1}_J = 0$  and has the absolute value of each element of  $\varepsilon$  be less than  $\tau$ . However, having  $\varepsilon$  be distributed independently of  $\mathbf{p}, x, \mathbf{z}$  means that for any value of  $\mathbf{p}, x, \mathbf{z}$ , the range of possible values of  $\mathbf{w}$  is limited by the support of  $\varepsilon$ . For example, if the support of  $\varepsilon$  is such that each element of  $\varepsilon$  is less than .1 in absolute value, and if for a given  $\mathbf{p}, x, \mathbf{z}$  we have each element of  $\mathbf{B}\mathbf{n}(y, \mathbf{z}) + \nabla_{\mathbf{p}}T(\mathbf{p}, \mathbf{z}) + \nabla_{\mathbf{p}}S(\mathbf{p}, \mathbf{z})y$  lying between .2 and .8 for any value of  $\varepsilon$  (which, along with  $\mathbf{p}, x, \mathbf{z}$  determines  $y$ ), then each element of  $\mathbf{w}$  for consumers having this observed value of  $\mathbf{p}, x, \mathbf{z}$  must lie between .1 and .9 (though consumers with different values of  $\mathbf{p}, x, \mathbf{z}$  could have smaller or larger budget shares). This restriction on the conditional support of  $\mathbf{w}$  is a general feature of virtually all nontrivial demand systems that map demand errors into unobserved preference heterogeneity parameters, such as Brown and Matzkin (1998) and Matzkin (2005).

Having  $\varepsilon$  distributed independently of  $\mathbf{p}, x, \mathbf{z}$  implies conditional homoskedasticity of  $\varepsilon$ . Brown and Walker (1989) and Lewbel (2001) show that additive errors in Marshallian budget share demand systems must be heteroskedastic except under special circumstances (like homotheticity). Their results do not apply to the EASI model, because the EASI model has additive errors in implicit Marshallian (and hence in the Hicksian) budget share demand functions, rather than in the Marshallian budget share demand functions. Still, one might expect heteroskedastic errors, e.g., empirical demands may have the property noted by Hildenbrand (1994) of variance increasing in  $x$ . This need not conflict with the EASI model, or with regularity, coherence, or invertibility since all these conditions can be satisfied if  $\mathbf{m}(y, \mathbf{z}, \varepsilon) = \mathbf{B}\mathbf{n}(y, \mathbf{z}) + \mathbf{N}(y, \mathbf{z}, \varepsilon)$  where  $\mathbf{N}$  is mean zero and



invertible in  $\varepsilon$ , which would then generate implicit Marshallian budget errors given by  $\mathbf{N}$  instead of  $\varepsilon$ .

## 5.7 The Rank of EASI Demand Systems

For clarity, in this section the heterogeneity parameters  $\mathbf{z}$ ,  $\varepsilon$  are held fixed, and so without loss of generality are omitted from the equations here. A demand system is defined to be exactly aggregable if it has Marshallian demands of the form  $\mathbf{w} = \mathbf{G}(\mathbf{p})\mathbf{f}(x)$  for some matrix valued function  $\mathbf{G}(\mathbf{p})$  and some vector valued function  $\mathbf{f}(x)$ . Gorman (1981) proved that for any exactly aggregable demand system that is derived from utility maximization, the maximum possible rank of the matrix  $\mathbf{G}(\mathbf{p})$  is three. Generalizing Gorman, Lewbel (1991) defined the rank of any demand system to be dimension of the space spanned by its Engel curves. This definition coincides with Gorman's rank of  $\mathbf{G}(\mathbf{p})$  for the special case of exactly aggregable demands. Lewbel (1991) shows that rank of the demand system corresponding to any cost function  $C(\mathbf{p}, u)$  equals the smallest integer  $\ell$  such that there exist functions  $g, \theta_1, \dots, \theta_\ell$ , where  $C(\mathbf{p}, u) = g[\theta_1(\mathbf{p}), \dots, \theta_\ell(\mathbf{p}), u]$ . The maximum possible rank of any demand system is the number of goods in the system minus one.

It follows by inspection of equations (2) and (5) that EASI demands are not constrained to have rank less than three, and can in fact have any rank, even when the EASI demands are linear in  $y$ . For example, for the reference type household the EASI demands we use for our empirical work has the cost function  $C(\mathbf{p}, u) = \mathbf{p}' \left[ \sum_{r=0}^5 \mathbf{b}_r u^r \right] + \frac{1}{2} \mathbf{p}' \mathbf{A}_0 \mathbf{p} + \frac{1}{2} \mathbf{p}' \mathbf{B} \mathbf{p} u$ , which to minimize  $\ell$  can be written as  $C(\mathbf{p}, u) = (\mathbf{p}' \mathbf{b}_0 + \frac{1}{2} \mathbf{p}' \mathbf{A}_0 \mathbf{p}) + (\mathbf{p}' \mathbf{b}_1 + \frac{1}{2} \mathbf{p}' \mathbf{B} \mathbf{p}) u + \sum_{r=2}^5 (\mathbf{p}' \mathbf{b}_r) u^r$ , and so has rank  $\ell = 6$ .

## 5.8 Shape Invariance and Equivalence Scales

Shape-invariance is a property of demand functions that is relevant for the construction of equivalence scales, is convenient for semiparametric demand modelling, and has been found to at least approximately hold empirically in some data sets. See, e.g., Blundell and Lewbel (1991), and Blundell, Duncan and Pendakur (1998), Pendakur (1999), and Blundell, Chen, and Kristensen, (2003). Shape-invariance is satisfied if and only if Marshallian budget shares are identical across household types except for equation specific vertical translations and a horizontal translation that is common across equations.

In our notation, shape-invariance is satisfied if and only if the log-cost function may be written as  $C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = f[\mathbf{p}, h(u, \mathbf{z}, \varepsilon)] + G(\mathbf{p}, \mathbf{z}, \varepsilon)$ . In this case, Hicks demands are given by  $\omega(\mathbf{p}, u, \mathbf{z}, \varepsilon) = \nabla_{\mathbf{p}} f[\mathbf{p}, h(u, \mathbf{z}, \varepsilon)] + \nabla_{\mathbf{p}} G(\mathbf{p}, \mathbf{z}, \varepsilon)$ . We may then invert  $C$  with respect to utility to obtain indirect utility  $h(\cdot, u, \mathbf{z}, \varepsilon) = f^{-1}[\cdot, x - G(\mathbf{p}, \mathbf{z}, \varepsilon)]$  and substituting this expression into Hicks demands yields  $\mathbf{w} = \nabla_{\mathbf{p}} f[\mathbf{p}, x - G(\mathbf{p}, \mathbf{z}, \varepsilon)] + \nabla_{\mathbf{p}} G(\mathbf{p}, \mathbf{z}, \varepsilon)$ . This structure for demands may be very complex over  $x$ , but the characteristics  $\mathbf{z}$  and  $\varepsilon$  enter demands in a very simple way: they translate expenditure shares vertically by the  $J$ -vector  $\nabla_{\mathbf{p}} G(\mathbf{p}, \mathbf{z}, \varepsilon)$  and horizontally by the scalar  $G(\mathbf{p}, \mathbf{z}, \varepsilon)$ .

Define an equivalence scale  $E$  as the ratio of costs of a household with characteristics  $\mathbf{z}$ ,  $\varepsilon$  and a household with reference characteristics  $\mathbf{0}_L, \mathbf{0}_J$ , so that  $\ln E(\mathbf{p}, u, \mathbf{z}, \varepsilon) = C(\mathbf{p}, u, \mathbf{z}, \varepsilon) - C(\mathbf{p}, u, \mathbf{0}_L, \mathbf{0}_J)$ . If demands are shape invariant and the untestable restriction that  $h(u, \mathbf{z}, \varepsilon) = \bar{h}(u)$

holds then  $C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = f[\mathbf{p}, \bar{h}(u)] + G(\mathbf{p}, \mathbf{z}, \varepsilon)$  so

$$\ln E(\mathbf{p}, u, \mathbf{z}, \varepsilon) = G(\mathbf{p}, \mathbf{z}, \varepsilon) \quad (29)$$

and therefore the equivalence scale depends only on prices and characteristics, but not on the utility level  $u$ . This property for  $E$  is called equivalence scale exactness (ESE) or independence-of-base (IB) by Blackorby and Donaldson (1993) and Lewbel (1989), respectively. Shape invariance is a necessary condition for IB/ESE.

EASI models can be shape invariant and can satisfy IB/ESE. The cost function (21) satisfies shape-invariance if and only if  $S$  is independent of  $\mathbf{z}$  and the vector-function  $\mathbf{m}$  is additively separable into a vector-function  $\mathbf{m}_1$  that depends on utility only and a vector-function  $\mathbf{m}_2$  that depends on characteristics  $\mathbf{z}, \varepsilon$  only. In this case, we have

$$\begin{aligned} C(\mathbf{p}, u, \mathbf{z}, \varepsilon) &= u + \mathbf{p}'\mathbf{m}_1(h(u, \mathbf{z}, \varepsilon)) + \mathbf{p}'\mathbf{m}_2(\mathbf{z}, \varepsilon) + T(\mathbf{p}, \mathbf{z}) + S(\mathbf{p})h(u, \mathbf{z}, \varepsilon) \\ &= [u + \mathbf{p}'\mathbf{m}_1(h(u, \mathbf{z}, \varepsilon)) + S(\mathbf{p})h(u, \mathbf{z}, \varepsilon)] + [\mathbf{p}'\mathbf{m}_2(\mathbf{z}, \varepsilon) + T(\mathbf{p}, \mathbf{z})]. \end{aligned}$$

This specification also satisfies IB/ESE if it satisfies the additional, untestable restriction that  $h(u, \mathbf{z}, \varepsilon) = \bar{h}(u)$ .

Shape-invariance is easily imposed on our empirically estimated model given by (5). In particular, that model has shape invariant demands if  $\mathbf{D} = \mathbf{0}$ . However, with  $\mathbf{D} = \mathbf{0}$  the log equivalence-scale is given by  $\ln E(\mathbf{p}, u, \mathbf{z}, \varepsilon) = \mathbf{p}'(\mathbf{Cz} + \varepsilon) + \frac{1}{2} \sum_{l=1}^L z_l \mathbf{p}'\mathbf{A}_l \mathbf{p}$ , which takes on a fixed value at the base price vector, so  $\ln E(\mathbf{0}_J, u, \mathbf{z}, \varepsilon) = 0$ . To relax this implausible restriction, one could add a term linear in  $\mathbf{z}$  to the log-cost function with  $\mathbf{D} = \mathbf{0}$ , so that

$$C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{d}'\mathbf{z} + \mathbf{p}' \left[ \sum_{r=0}^5 \mathbf{b}_r u^r + \mathbf{Cz} + \varepsilon \right] + \frac{1}{2} \sum_{l=0}^L z_l \mathbf{p}'\mathbf{A}_l \mathbf{p} + \frac{1}{2} \mathbf{p}'\mathbf{B}\mathbf{p}u, \quad (30)$$

where  $\mathbf{d}$  is a  $T$ -vector of parameters. The equivalence scale is then given by  $\ln E(\mathbf{p}, u, \mathbf{z}, \varepsilon) = \mathbf{d}'\mathbf{z} + \mathbf{p}'(\mathbf{Cz} + \varepsilon) + \frac{1}{2} \sum_{l=1}^L z_l \mathbf{p}'\mathbf{A}_l \mathbf{p}$ , which takes on the value  $\mathbf{d}'\mathbf{z}$  when evaluated at the base price vector. In this version of the model, implicit utility  $y$  includes the constant term  $\mathbf{d}'\mathbf{z}$  and is given by

$$y = \frac{x - \mathbf{p}'\mathbf{w} - \mathbf{d}'\mathbf{z} + \sum_{l=0}^L z_l \mathbf{p}'\mathbf{A}_l \mathbf{p}/2}{1 - \mathbf{p}'\mathbf{B}\mathbf{p}/2}.$$

Other than this change in  $y$ , the demand functions are same as before, so the parameter vector  $\mathbf{d}$  enters the model only through  $y$ . Estimates of the model in this form are available on request from the authors.

## 5.9 Closure Under Unit Scaling

A desirable feature of demand models is that they be closed under unit scaling, that is, that a change in the units that goods are measured in (or equivalently, a change in the base year or region where prices are normalized to equal one) only changes the values of the parameters or functions that define the model, leaving predicted values and estimated elasticities unchanged. See, e.g., Pollak and Wales (1980), especially footnote 15. The AID system is closed under unit scaling. The Quadratic AID of Banks, Blundell, and Lewbel (1997) is closed if the constant scalar parameter

$a_0$  in that model is estimated, though in practice that parameter is usually fixed at some convenient value.

The parametric models proposed in this paper are not closed under unit scaling. To close them, we could replace  $\mathbf{p}$  with  $\mathbf{p} + \mathbf{k}$  everywhere that  $\mathbf{p}$  appears in equations (2), (3), and (4), or in our empirically estimated model defined by equations (5), (8), and (9), where  $\mathbf{k}$  is an additional  $J$ -vector of parameters to be estimated, with the free normalization  $\mathbf{k}'\mathbf{1}_J = 0$ . With the addition of  $\mathbf{z}$  and  $\varepsilon$  terms, these are examples of equation (26) with  $\bar{\mathbf{c}}(\mathbf{p}) = \mathbf{c}(\mathbf{p}) = \mathbf{p} + \mathbf{k}$ .

To check possible sensitivity of our empirical results to unit scaling, we tried to reestimate our empirical model including the additional parameter vector  $\mathbf{k}$ , but in every attempt  $\mathbf{k}$  was either completely insignificant or the model failed to converge. Also, reestimating the model (without  $\mathbf{k}$ ) after changing the base year and region had little effect on the estimates, leaving Engel curve shapes unchanged and altering elasticities by a few percent at most. Like the Quadratic AID model with  $a_0$  fixed, the parametric EASI models without  $\mathbf{k}$  are approximately, though not exactly, closed under unit scaling. Formally, the objective function used for parameter estimation is relatively flat in directions corresponding to changes in  $\mathbf{k}$ .

To see why the parametric EASI models as estimated are almost closed, consider the general cost function

$$C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{p}'\mathbf{m}(u, \mathbf{z}, \varepsilon) + \mathbf{p}'[\mathbf{A}_1 + \mathbf{A}_2h(u, \mathbf{z}, \varepsilon)]\mathbf{p}/2 \quad (31)$$

where  $h$  and the  $J$ -vector  $\mathbf{m}$  are nonparametric functions. This cost function has demands that are closed under unit scaling (up to possible inequality constraints on the nonparametric functions) because

$$\begin{aligned} C(\mathbf{p} + \mathbf{k}, u, \mathbf{z}, \varepsilon) &= u + (\mathbf{p} + \mathbf{k})'\mathbf{m}(u, \mathbf{z}, \varepsilon) + (\mathbf{p} + \mathbf{k})'[\mathbf{A}_1 + \mathbf{A}_2h(u, \mathbf{z}, \varepsilon)](\mathbf{p} + \mathbf{k})/2 \\ &= u^* + \mathbf{p}'\mathbf{m}^*(u^*, \mathbf{z}, \varepsilon) + \mathbf{p}'[\mathbf{A}_1 + \mathbf{A}_2h^*(u^*, \mathbf{z}, \varepsilon)]\mathbf{p}/2 \end{aligned}$$

where  $u^* = u + \mathbf{k}'\mathbf{m}(u, \mathbf{z}, \varepsilon) + \mathbf{k}'[\mathbf{A}_1 + \mathbf{A}_2h(u, \mathbf{z}, \varepsilon)]\mathbf{k}/2$ ,  $\mathbf{m}^*(u^*, \mathbf{z}, \varepsilon) = \mathbf{m}(u, \mathbf{z}, \varepsilon) + [\mathbf{A}_1 + \mathbf{A}_2h(u, \mathbf{z}, \varepsilon)]\mathbf{k}$ , and  $h^*(u^*, \mathbf{z}, \varepsilon) = h(u, \mathbf{z}, \varepsilon)$ , so by suitably redefining the functions  $\mathbf{m}^*$  and  $h^*$ , adding  $\mathbf{k}$  to the log price vector is equivalent to ordinally transforming  $u$ , which leaves the resulting demand functions unchanged. This paper's parametric EASI models are special cases of the cost function (31), so they fail to be closed under unit scaling only because  $\mathbf{m}^*$  and  $h^*$  need not be contained in the same family of functional forms that are assumed for  $\mathbf{m}$  and  $h$ . However, our flexible choice of these functions, particularly of  $\mathbf{m}$ , means that  $\mathbf{m}^*$  and  $h^*$  can be closely approximated by suitable choice of parameterization of  $\mathbf{m}$  and  $h$ , which explains the empirical finding that the numerical effects of violation of closure under unit scaling are very small.

## 5.10 Marshallian Elasticity Calculations

Cost functions in the class of equation (21) have  $y$  given equation (8), so Marshallian demand functions  $\mathbf{w}(\mathbf{p}, x, \mathbf{z}, \varepsilon)$  for these EASI models are implicitly given by

$$\mathbf{w}(\mathbf{p}, x, \mathbf{z}, \varepsilon) = \omega \left[ \mathbf{p}, \frac{x - \mathbf{p}'\mathbf{w}(\mathbf{p}, x, \mathbf{z}, \varepsilon) - \mathbf{p}'\mathbf{A}_1\mathbf{p}/2}{1 + \mathbf{p}'\mathbf{A}_2\mathbf{p}/2}, \mathbf{z}, \varepsilon \right]$$

Taking the total derivative of this expression with respect to any variable  $v$  gives

$$\begin{aligned} \nabla_v \mathbf{w}(\mathbf{p}, x, \mathbf{z}, \varepsilon) &= \nabla_v \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon) \\ &+ [\nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)] \left[ \nabla_v \left( \frac{x - \mathbf{p}'\mathbf{w} - \mathbf{p}'\mathbf{A}_1\mathbf{p}/2}{1 + \mathbf{p}'\mathbf{A}_2\mathbf{p}/2} \right) - \left( \frac{x - \mathbf{p}'\nabla_v \mathbf{w}(\mathbf{p}, x, \mathbf{z}, \varepsilon)}{1 + \mathbf{p}'\mathbf{A}_2\mathbf{p}/2} \right) \right] \end{aligned}$$

and solving for the Marshallian semielasticity  $\nabla_v \mathbf{w}(\mathbf{p}, x, \mathbf{z}, \varepsilon)$  yields

$$\begin{aligned} \nabla_v \mathbf{w}(\mathbf{p}, x, \mathbf{z}, \varepsilon) &= \left[ I_J - \frac{[\nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)] \mathbf{p}'}{1 + \mathbf{p}'\mathbf{A}_2\mathbf{p}/2} \right]^{-1} \\ &\left[ \nabla_v \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon) + [\nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)] \left[ \nabla_v \left( \frac{x - \mathbf{p}'\mathbf{w} - \frac{\mathbf{p}'\mathbf{A}_1\mathbf{p}}{2}}{1 + \mathbf{p}'\mathbf{A}_2\mathbf{p}/2} \right) - \frac{x}{1 + \frac{\mathbf{p}'\mathbf{A}_2\mathbf{p}}{2}} \right] \right] \end{aligned} \quad (32)$$

where  $I_J$  is the  $J$  by  $J$  identity matrix. In particular, taking  $v$  to be  $x$  above shows that, after algebraic simplification, the Marshallian semielasticity with respect to nominal expenditures  $x$  is

$$\nabla_x \mathbf{w}(\mathbf{p}, x, \mathbf{z}, \varepsilon) = \left( I_J - \frac{[\nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)] \mathbf{p}'}{1 + \mathbf{p}'\mathbf{A}_2\mathbf{p}/2} \right)^{-1} \left( \frac{(1-x) \nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)}{1 + \mathbf{p}'\mathbf{A}_2\mathbf{p}/2} \right)$$

where  $\nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)$  is given by equation (13). Equation (32) could also be evaluated taking  $v$  to be  $\mathbf{p}$  to obtain Marshallian price elasticities, but it is simpler to recover them from the Hicksian  $\mathbf{p}$  elasticities (12) and the above Marshallian  $x$  elasticities using the Slutsky matrix

$$\nabla_{\mathbf{p}'} \mathbf{w}(\mathbf{p}, x, \mathbf{z}, \varepsilon) = \nabla_{\mathbf{p}'} \omega(\mathbf{p}, u, \mathbf{z}, \varepsilon) - [\nabla_x \mathbf{w}(\mathbf{p}, x, \mathbf{z}, \varepsilon)] \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)'$$

Finally, again using equation (32), the Marshallian semielasticity with respect to  $\mathbf{z}$  is

$$\nabla_{\mathbf{z}} \mathbf{w}(\mathbf{p}, x, \mathbf{z}, \varepsilon) = \left( I_J - \frac{[\nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)] \mathbf{p}'}{1 + \mathbf{p}'\mathbf{A}_2\mathbf{p}/2} \right)^{-1} \left( \nabla_{\mathbf{z}} \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon) - \frac{x \nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)}{1 + \mathbf{p}'\mathbf{A}_2\mathbf{p}/2} \right)$$

where  $\nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)$  and  $\nabla_{\mathbf{z}} \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)$  are given by equations (13) and (14).

Some of the above elasticity expressions depend on  $\varepsilon$  either directly or through  $\mathbf{w}$ . Given estimated model parameters,  $\varepsilon$  for each consumer can be estimated as the model residuals (the difference between fitted and observed  $\mathbf{w}$ ). We may therefore estimate mean elasticities in the population by calculating the estimated elasticities for each individual in the sample, plugging in their observed  $\mathbf{w}$  or estimated  $\varepsilon$  where needed, and averaging the result. Other features of the population distribution of elasticities such as its median or variance can likewise be estimated from the corresponding empirical distribution.

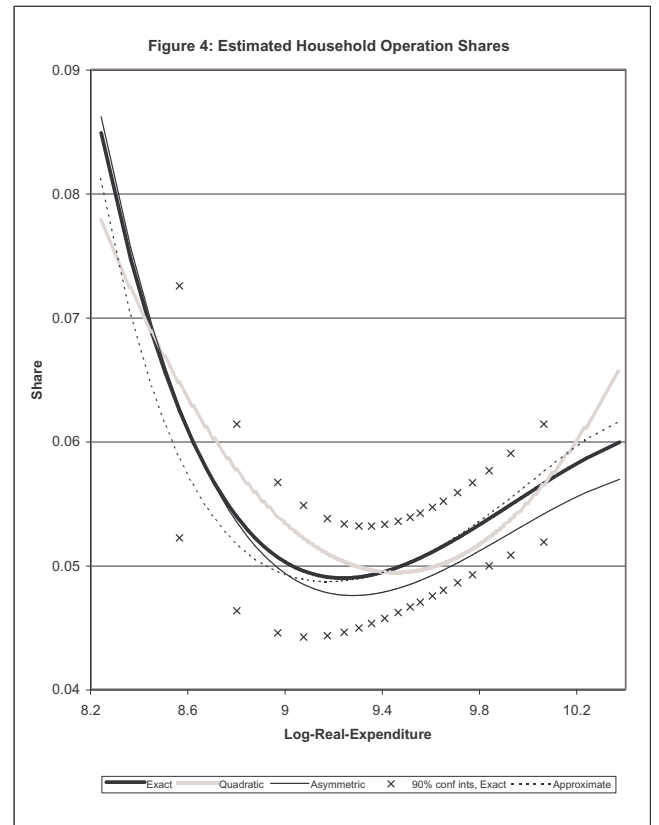
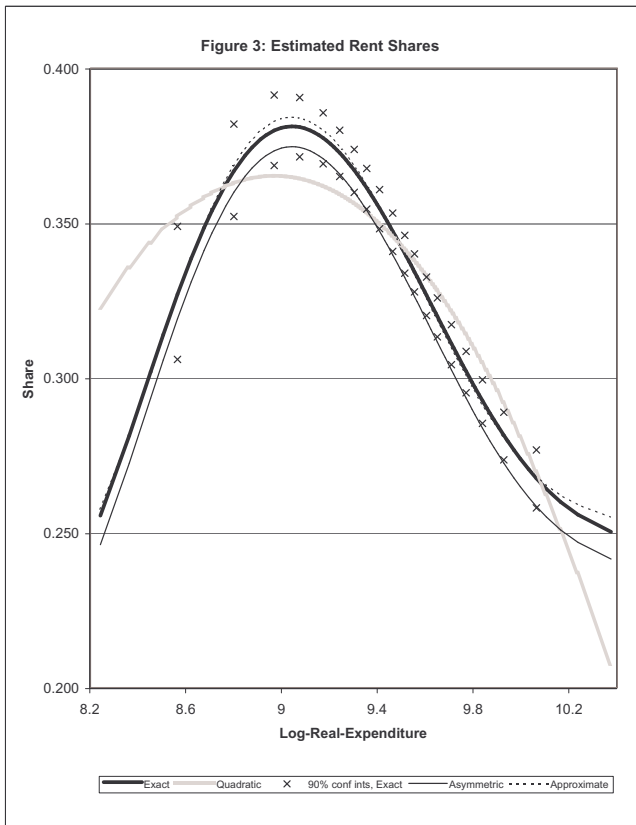
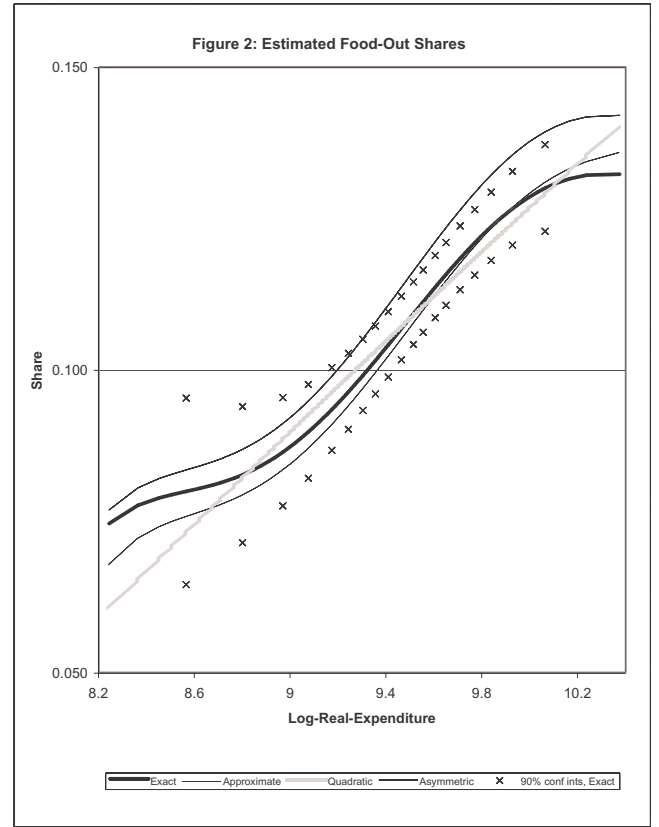
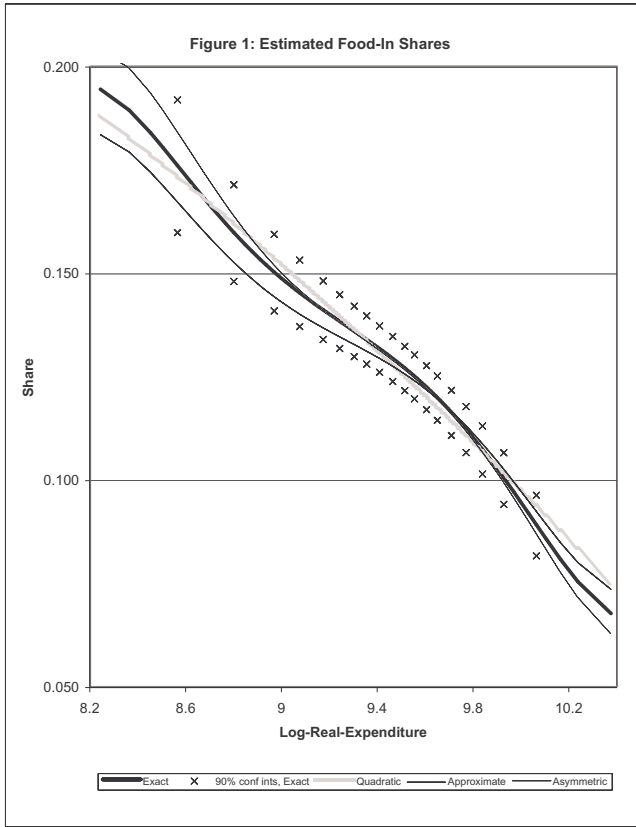
If Marshallian demands are of direct interest, evaluated at points other than those observed in the sample, they can be obtained numerically by, e.g., numerically solving  $x = C(\mathbf{p}, u, \mathbf{z}, \varepsilon)$  for  $u$  and substituting the result into the Hicksian demand functions. However, as the above equations show, this will often not be necessary for evaluating the effects on demand or welfare of price, expenditure, or demographic changes.

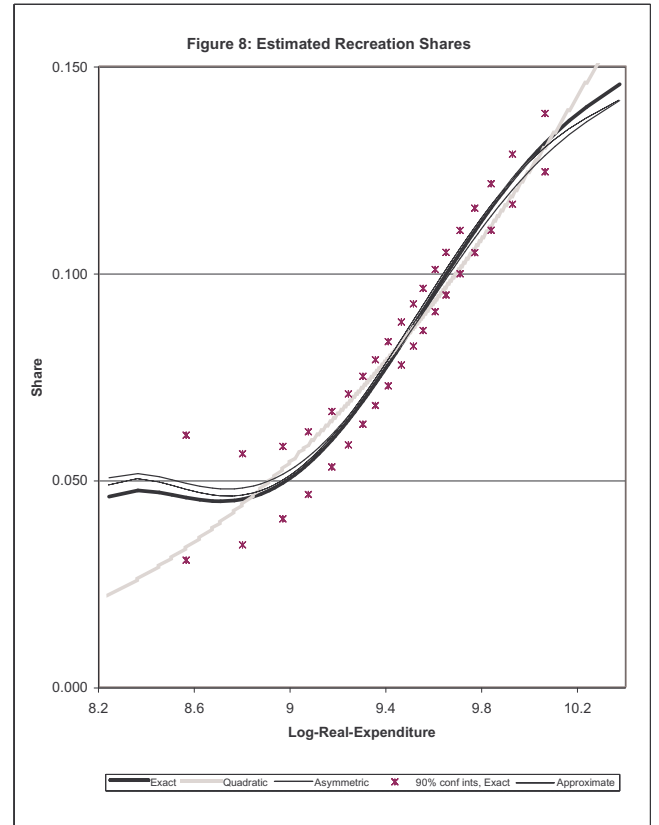
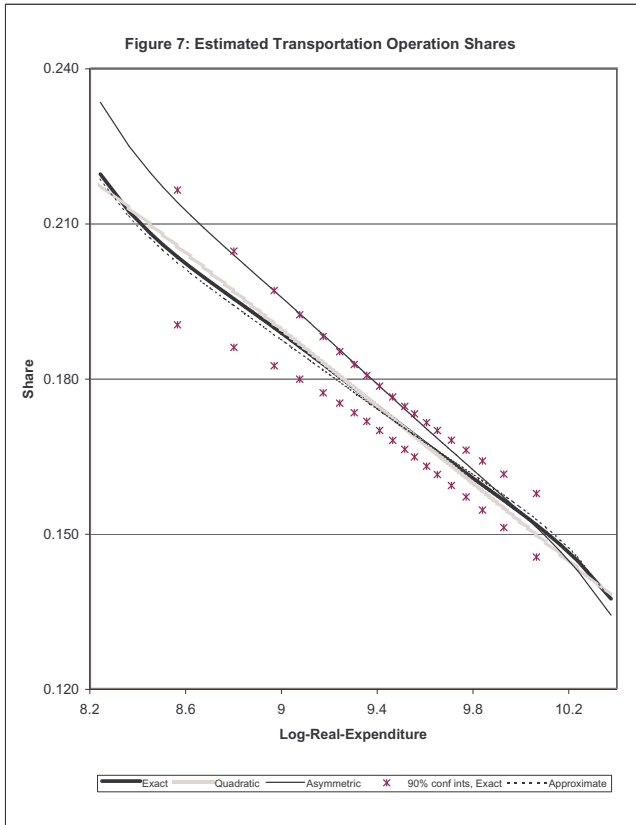
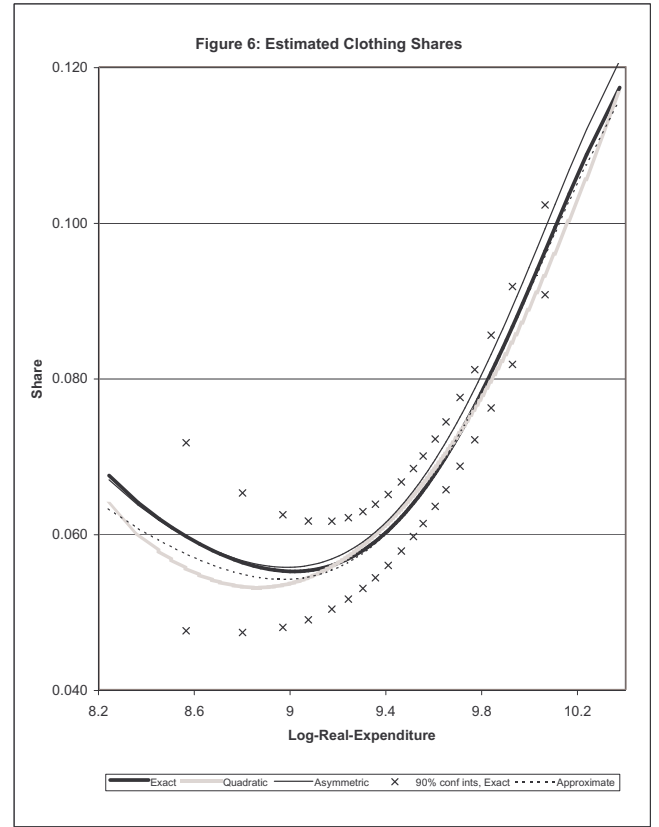
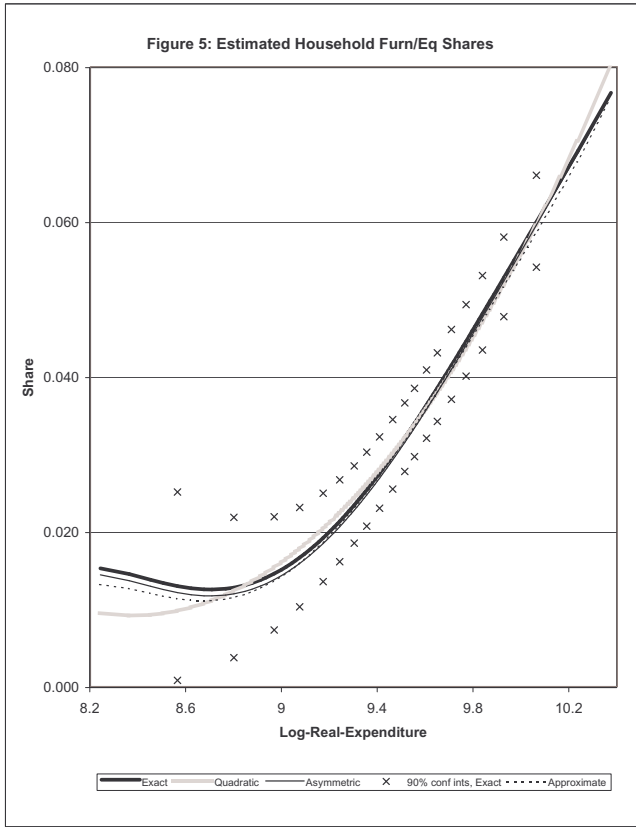
Variable		Mean	Std Dev	Minimum	Maximum
budget shares	Food-at-Home	0.14	0.08	0.00	0.61
	Food-Out	0.08	0.07	0.00	0.63
	Rent	0.37	0.12	0.01	0.89
	Household Oper	0.07	0.04	0.00	0.61
	Household Furneq	0.04	0.05	0.00	0.52
	Clothing	0.08	0.06	0.00	0.48
	Transport Oper	0.11	0.08	0.00	0.60
	Recreation	0.08	0.07	0.00	0.59
	Personal Care	0.03	0.02	0.00	0.21
log-prices	Food-at-Home	-0.05	0.43	-1.41	0.34
	Food-Out	0.04	0.49	-1.46	0.53
	Rent	-0.08	0.40	-1.27	0.37
	Household Oper	-0.06	0.45	-1.40	0.32
	Household Furneq	-0.05	0.32	-0.94	0.20
	Clothing	0.04	0.35	-0.94	0.34
	Transport Oper	-0.07	0.57	-1.53	0.57
	Recreation	0.01	0.40	-1.04	0.42
	Personal Care	-0.03	0.38	-1.11	0.29
demographics	age-40	0.71	11.89	-15.00	24.00
	male	0.51	0.50	0.00	1.00
	car-owner	0.42	0.49	0.00	1.00
	social asst	0.27	0.44	0.00	1.00
	time	88.99	8.73	69.00	99.00
log-expenditure (median-norm'd)	$x$	-0.11	0.59	-2.75	1.66
	$x \cdot p \cdot w$	-0.07	0.44	-1.70	1.44

model	test of	parameters	df	Test Stat	p-value
asymmetric with with $y^{-1}$ , $y^{\circ}$	symmetry	$A_i = A_i'$ for all $i$ ; $B = B'$	196	241.8	0.014
	symmetry	$A_i = A_i'$ for all $i$	168	194.1	0.082
	symmetry	$B = B'$	28	24.4	0.661
	exclusion	$B = 0$	64	82.5	0.060
	exclusion	$A_i = 0$ for all $i$	384	636.0	0.000
	exclusion	$B = 0, A_i = 0$ for all $i$	448	999.7	0.000
	exclusion	$y^{-1}$	8	15.6	0.049
	exclusion	$y^6$	8	20.6	0.008
	exclusion	$y^{\circ}$	8	21.0	0.007
	exclusion	$y^{-1}, y^{\circ}$	16	28.3	0.029
	exclusion	$y^{-1}, y^{\circ}, y^{\circ}$	24	75.5	0.000
	symmetric with without $y^{-1}, y^{\circ}$	exclusion	$B = 0$	36	51.7
exclusion		$A_i = 0$ for all $i$	216	646.2	0.000
exclusion		$B = 0, A_i = 0$ for all $i$	252	710.6	0.000
exclusion		$y^5$	8	45.8	0.000
non-quadratic		Food-at-Home	3	16.9	0.001
non-quadratic		Food-Out	3	10.3	0.016
non-quadratic		Rent	3	94.4	0.000
non-quadratic		Household Oper	3	22.7	0.000
non-quadratic		Household Furneq	3	3.4	0.340
non-quadratic		Clothing	3	9.4	0.025
non-quadratic		Transport Oper	3	6.8	0.079
non-quadratic		Recreation	3	28.9	0.000
overidentification		J-test (with sym $pz$ )	196	238.4	0.021
overidentification		J-test (with sym. $pz, y^{-1}, y^{\circ}$ )	212	273.4	0.003

Variable	std dev	$y$	$y(p'(w-e))$	$y\text{-bar}$	$x-p'w$	$x-p'w\text{-bar}$	$p'e$	$p'(w-e)$	$p'Ap/2$	$p'Bp/2$
$y$	0.4393	1.0000								
$y(p'(w-e))$	0.4390	0.9988	1.0000							
$y\text{-bar}$	0.4459	0.9984	0.9996	1.0000						
$x-p'w$	0.4387	1.0000	0.9988	0.9984	1.0000					
$x-p'w\text{-bar}$	0.4452	0.9984	0.9996	1.0000	0.9984	1.0000				
$p'e$	0.0213	-0.0445	0.0041	0.0035	-0.0444	0.0035	1.0000			
$p'(w-e)$	0.0716	0.0678	0.0672	0.0741	0.0685	0.0748	-0.0141	1.0000		
$p'Ap/2$	0.0018	0.0173	0.0165	0.0234	0.0134	0.0196	-0.0172	-0.2243	1.0000	
$p'Bp/2$	0.0011	-0.0123	-0.0141	-0.0122	-0.0140	-0.0140	-0.0371	-0.6439	0.4676	1.0000

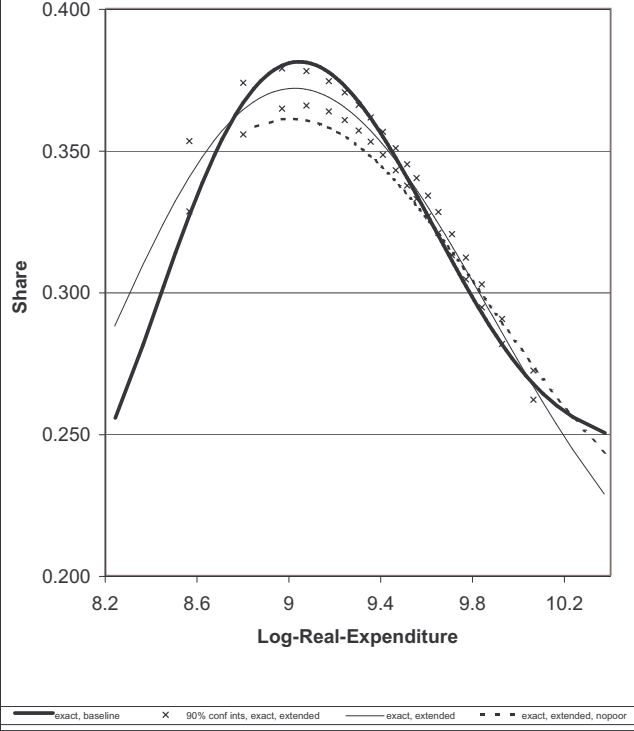
	Own-Price	Own-Price	Own-Price	Budget-share Semi-Elasticities							
	B element	Slutsky Terms	Quant Elast	Food-in	Food-Out	Rent	HH Oper	HH Furneq	Clothing	Tran Oper	Recr
Food-at-Home	0.047	-0.137	-0.064	-0.025							
<i>Std Err</i>	0.063	0.045	0.347	0.045							
Food-Out	-0.009	-0.120	-0.126	0.051	-0.025						
<i>Std Err</i>	0.068	0.037	0.346	0.033	0.037						
Rent	0.088	-0.164	0.528	-0.026	0.037	0.063					
<i>Std Err</i>	0.047	0.029	0.083	0.021	0.021	0.029					
Household Oper	0.044	-0.035	0.292	0.026	-0.002	-0.002	0.012				
<i>Std Err</i>	0.039	0.026	0.517	0.026	0.022	0.014	0.026				
Household Furneq	0.037	-0.034	-0.150	0.110	-0.043	0.028	-0.018	-0.005			
<i>Std Err</i>	0.114	0.058	1.937	0.035	0.032	0.021	0.026	0.058			
Clothing	-0.050	-0.066	-0.063	-0.050	-0.001	-0.066	-0.025	-0.023	-0.008		
<i>Std Err</i>	0.060	0.046	0.734	0.030	0.032	0.020	0.024	0.038	0.046		
Transport Oper	-0.006	-0.105	0.393	-0.023	-0.030	-0.019	-0.013	-0.034	0.053	0.038	
<i>Std Err</i>	0.035	0.020	0.115	0.021	0.019	0.016	0.017	0.021	0.019	0.020	
Recreation	-0.030	-0.182	-1.191	-0.044	0.032	-0.021	0.034	-0.010	0.087	0.033	-0.106
<i>Std Err</i>	0.093	0.053	0.635	0.034	0.032	0.023	0.025	0.045	0.037	0.022	0.053



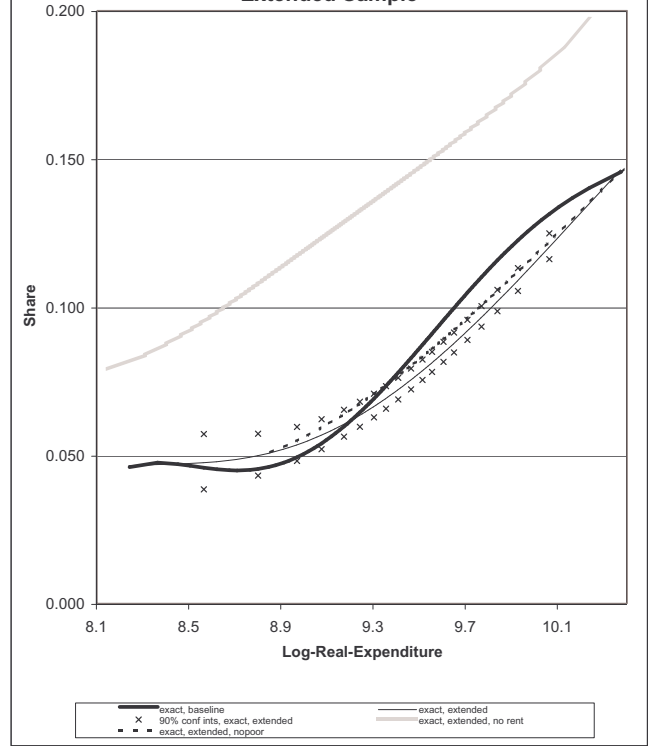




**Figure 9: Rent Shares: Baseline vs Extended Sample**



**Figure 10: Recreation Shares, Baseline vs Extended Sample**



**Figure 11: Cost-of-Living Experiment: 15% tax on rent**

