Tricks With Hicks: The EASI Demand System

Arthur Lewbel and Krishna Pendakur
Boston College and Simon Fraser University

Revised October 2005

Abstract

Recent work in consumer demand analysis demonstrates the importance of unobserved preference heterogeneity and the presence of complex shapes for Engel curves. These are difficult features to incorporate into utility derived Marshallian demand systems. We address these key problems by specifying demands in terms of prices and an affine function of stone-index deflated expenditure. This innovation, which we call the Exact Affine Stone Index (EASI) family of demand systems, provides the nice features of Hicksian demands in an empirically practical framework. Like the Almost Ideal Demand (AID) system, EASI budget shares are linear in parameters up to the construction of real expenditures. However, unlike the AID system, EASI Engel curve shapes for each good are almost completely unrestricted; they can have any rank and be polynomials or splines of any order in real expenditures. In addition, EASI error terms equal random utility parameters to account for unobserved preference heterogeneity. EASI demand functions can be estimated using ordinary GMM, and, like the AID system, an approximate EASI model can be estimated by linear regression methods.

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 David Donaldson, Stefan Sperlich, Angus Deaton, and Richard Blundell for helpful comments. Any errors are our own.

1 Introduction

Recent empirical nonparametric and semiparametric work with large consumer expenditure data sets reveals 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 (2003). 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 interpretted as representing unobserved heterogeneity (see, for example, 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 difficult or numerically impossible 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 like those listed above. We show that under some conditions log real expenditures y, which are ordinally equivalent to u, can be expressed as a simple function of $\mathbf{w}$, $\mathbf{p}$, $\mathbf{z}$ and x. We use this result to directly estimate what we call *Pseudo-Marshallian* demands, which are Hicksian demands after replacing u with y. Specifically, we define the *Exact Affine Stone Index* (EASI) class of cost functions, which have y equal to an affine function of Stone index deflated log nominal expenditures, $x - \mathbf{p}'\mathbf{w}$. The resulting EASI Pseudo-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. EASI budget shares are linear in

$\mathbf{p}$ and are polynomials of any order in of $\mathbf{z}$ and y. They can also include interaction terms such as $\mathbf{p}y$, $\mathbf{z}y$ and $\mathbf{pz}'$, 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 z, and so can encompass empirically important specifications that most parametric models cannot capture, such as the semiparametric S shaped Engel curves reported by Blundell, Chen and Kristensen (2005). The AID system is linear in y and has rank two, and the quadratic AID of Banks, Blundell, and Lewbel (1997) is quadratic in y with rank three, but 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 or less restriction even with polynomial Engel curves.

4. EASI budget share error terms can equal unobserved preference heterogeneity or random utility parameters. This unobserved preference heterogeneity is locally coherent and invertible. The AID and other similar models do not have this property, since in those models unobserved preference heterogeneity requires that additive errors depend upon $\mathbf{p}$ or x (see Brown and Walker 1989 and Lewbel 2001).

5. EASI demand functions can be estimated using nonlinear instrumental variables, particularly 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 difference between estimated exact GMM and approximate linear models.

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

The next section describes the model. We begin with an overview of our approach, followed by theorems showing the restrictions on cost functions required to construct Exact Stone Index and Exact Affine Stone Index Pseudo-Marshallian demand functions. We then introduce observed and unobserved preference heterogeneity into EASI demands. Then, we discuss parametric structures which fit into the EASI class, and a fully linear approximation for EASI demands, analogous to the approximate linear AID model, and apply the model consumer surplus and compensated elasticity calculations. Then, we describe model estimators, including consistent, asymptotically normal instrumental variable and GMM based estimators for the exact model and OLS estimators for the approximate model. We estimate the model using Canadian micro-data, and find more complicated Engel curve shapes than those that can be encompassed by parametric demand systems, and find that the approximate linear model captures this complexity very well. In a cost-of-living experiment, we find that both the increased flexibility of Engel curves and the the incorporation of unobserved heterogeneity into the model affect welfare calculations.

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

2 The Model

We specify a cost (expenditure) function and use Shephard's lemma to obtain Hicksian demands that have the desired properties. Standard methods obtain Marshallian demands, which are functions of $\mathbf{p}$, $\mathbf{z}$ and x from Hicksian demands by solving for u in terms of $\mathbf{p}$, $\mathbf{z}$ and x. We instead construct cost functions that have simple expressions for log real-expenditure y in terms of $\mathbf{w}$, $\mathbf{p}$, $\mathbf{z}$ and x, and substitute y for u in the Hicksian demands to yield what we call Pseudo-Marshallian demand functions. These Pseudo-Marshallian demands circumvent the difficulty of finding simple analytic expressions for indirect utility or Marshallian demands. A similar idea 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. 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.

To illustrate the idea of Pseudo-Marshallian demands, consider a simple example. Define the log cost (expenditure) function $x = C(\mathbf{p}, u)$ which equals the minimum log-expenditure required for an individual to attain utility level u when facing log prices $\mathbf{p}$. Consider the log cost function

$$C(\mathbf{p}, u) = u + \mathbf{p}'\mathbf{m}(u)$$

where $\mathbf{m}(u)$ is a $J-$vector valued function with $\mathbf{1}'_J\mathbf{m}(u) = 1$, and $\mathbf{1}_J$ is the $J-$vector of ones. By Shephard's lemma, this cost function has Hicksian (compensated) budget shares $\mathbf{w} = \mathbf{m}(u)$. Since $x = u+\mathbf{p}'\mathbf{m}(u)$ and $\mathbf{w} = \mathbf{m}(u)$, substituting out $\mathbf{m}(u)$ gives $u = x-\mathbf{p}'\mathbf{w}$ and therefore we obtain Pseudo-Marshallian budget shares

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

where y is log real-expenditure which is ordinally equivalent to utility. In this case, $y = x - \mathbf{p}'\mathbf{w}$ is the log of nominal expenditures x deflated by the log Stone price index $\mathbf{p}'\mathbf{w}$. 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. This is an example of what we call an Exact Stone Index demand system, which is any demand system for which $\mathbf{w}$ is a function of $\mathbf{p}$ and $x - \mathbf{p}'\mathbf{w}$. This model does not have a closed form expression for indirect utility or ordinary Marshallian demand functions except for very special choices of $\mathbf{m}(u)$, but it can still be readily estimated because it expresses budget shares as functions of observables $\mathbf{p}$ and $x - \mathbf{p}'\mathbf{w}$.

We show later that all Exact Stone Index demand systems have some undesirable properties, so we propose a generalization in which log real total expenditures are affine transforms of exact Stone index deflated log expenditures. These are Exact Affine Stone Index (EASI) Pseudo-Marshallian demand functions.

We wish to explicitly include both observable and unobservable sources of preference heterogeneity in our models, so in addition to an $L-$vector $\mathbf{z} = (z_1, ..., z_L)'$ of observable demographic (or other) characteristics that affect preferences, let ε be a $J-$vector of unobserved preference characteristics (taste parameters) satisfying $\mathbf{1}'_J\varepsilon = 1$. The log cost or expenditure function is now $x = C(\mathbf{p}, u, \mathbf{z}, \varepsilon)$, which equals the minimum log-expenditure required for an individual with characteristics $\mathbf{z}, \varepsilon$ to attain utility level u when facing log prices $\mathbf{p}$.

The class of EASI cost functions we develop are characterised by $C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{p}'\mathbf{m}(u, \mathbf{z}, \varepsilon) + T(\mathbf{p}, \mathbf{z}) + S(\mathbf{p}, \mathbf{z})u$, where $T(\mathbf{p}, \mathbf{z})$ and $S(\mathbf{p}, \mathbf{z})$ satisfy certain homogeneity conditions and $\mathbf{1}'_J\mathbf{m}(u, \mathbf{z}, \varepsilon) = 1$. This broad class includes the following parametric model which we take as our baseline case for empirical work:

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

Here, $z_0 = 1$ and (for notational convenience) is not an element of the vector $\mathbf{z}$; 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 $r \neq 0$; and $\mathbf{A}_l$, $l = 0, .., L$, and $\mathbf{B}$ are $J \times J$ symmetric matrices with $\mathbf{1}'_J\mathbf{A}_l = \mathbf{1}'_J\mathbf{B} = \mathbf{0}'_J$. By Shephard's lemma, this cost function has Hicksian (compensated) budget shares

$$\mathbf{w} = \sum_{r=-1}^{5} \mathbf{b}_r u^r + \mathbf{C}\mathbf{z} + \mathbf{D}\mathbf{z}u + \sum_{l=0}^{L} z_l \mathbf{A}_l\mathbf{p} + \mathbf{B}\mathbf{p}u + \varepsilon \quad (2)$$

It can be readily checked from these formulas that $C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{p}'\mathbf{w} - \sum_{l=0}^{L} z_l \mathbf{p}'\mathbf{A}_l\mathbf{p}/2 - \mathbf{p}'\mathbf{B}\mathbf{p}u/2$, and solving this expression for u implies that log real-expenditures y can be written as an affine transform of the log of Stone Index deflated nominal expenditures:

$$y = \frac{x - \mathbf{p}'\mathbf{w} + \sum_{l=0}^{L} z_l \mathbf{p}'\mathbf{A}_l\mathbf{p}/2}{1 - \mathbf{p}'\mathbf{B}\mathbf{p}/2}. \quad (3)$$

Since log real-expenditures are ordinally equivalent to utility, they can be substituted into

Hicksian budget shares to yield Pseudo-Marshallian budget shares

$$\mathbf{w} = \sum_{r=-1}^{5} \mathbf{b}_r y^r + \mathbf{C}\mathbf{z} + \mathbf{D}\mathbf{z}y + \sum_{l=0}^{L} z_l \mathbf{A}_l \mathbf{p} + \mathbf{B}\mathbf{p}y + \varepsilon. \tag{4}$$

Given y, which is a function of observables x, $\mathbf{p}$, $\mathbf{z}$, and the log Stone index $\mathbf{p}'\mathbf{w}$, the budget shares equations (4) are linear in parameters and so can be easily estimated. An approximate estimator can replace y with $x - \mathbf{p}'\mathbf{w}$ in (4), just like estimation of the approximate AID model. Better estimators simultaneously estimate the model (4) with the parameters in y given by (3). GMM estimation is used to account for the endogeneity that results from having $\mathbf{w}$ appear in y, and allows for possible unknown heteroskedasticity in ε.

The EASI budget shares (4) have compensated price effects governed by $\mathbf{A}_l$, $l = 0, 1, 2, ..., L$, 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 $\mathbf{b}_r$, $r = -1, 0, 1, 2, ..., 5$ specify budget shares as high order polynomials in log real expenditure, which allows Engel curves to have very complicated shapes. The terms $\mathbf{C}$ and $\mathbf{D}$ allow demographic characteristics to enter budget shares through intercept and slope terms on log real-expenditure. The random utility parameters ε, representing unobserved preference heterogeneity, take the form of simple additive errors in the Pseudo-Marshallian demand equations.

2.1 Exact Stone Index (ESI) Demands

Ignore preference heterogeneity for now, that is, consider $\mathbf{z}$ and ε 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. In most cases, explicit functions we provide for $C(\mathbf{p}, u)$ will not be globally regular, that is, they may only be regular over some limited range of values of $\mathbf{p}, u$, in which case the reported results will only be valid over these limited ranges. Shephard's lemma relates Hicksian (compensated) budget shares to regular cost functions by

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

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 $u = x - \mathbf{p}'\mathbf{w}$, so real expenditures, which hold utility constant when prices change, are equal to 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 $x - \mathbf{p}'\mathbf{w}$ 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 what we call Pseudo-Marshallian demand functions. The idea is to construct models where utility is ordinally equivalent to some simple function of observables (in this case Stone index deflated nominal total expenditures). Then, instead of solving for Marshallian demands, we can directly estimate these Pseudo-Marshallian demand functions. It is relatively easy to construct Hicksian demand functions that have the desirable properties listed in the introduction, and we exploit this fact by the use of Pseudo-Marshallian demands.

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

$$u = C(\mathbf{p}, u) - \mathbf{p}'[\nabla_{\mathbf{p}} C(\mathbf{p}, u)]. \tag{5}$$

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. $C(\mathbf{p}, u)$ 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 κ.

Proofs are in the Appendix. 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 $\mathbf{p}$, which makes $u = x - \mathbf{p}'\mathbf{w}$. Both homogeneity conditions must hold for a regular cost function to be an ESI cost function.

By ordinality, 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 by ordinality, $C(\mathbf{p}, u) - h(u)$ has the same indifference curves, and hence the same Marshallian and Pseudo-Marshallian demand functions, as $C(\mathbf{p}, h^{-1}(u)) - u = \widetilde{C}(\mathbf{p}, u) - u$. In this representation of the cost function (i.e., in this cardinalization of utility), we would have $x - \mathbf{p}'\mathbf{w} = u$ 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 $\boldsymbol{\omega}(\mathbf{p}, u) = \nabla_{\mathbf{p}} C(\mathbf{p}, u)$ are homogeneous of degree zero in $\mathbf{p}$ and in $\exp(\mathbf{p})$.

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 may bind 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 κ if and only if $d(p_1, p_2, u) = m(u)p_1 + [1 - m(u)]p_2$ for some function $m(u)$.

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

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

where $\mathbf{m}(u)$ is as above and $\mathbf{M}(u)$ is a J by J symmetric matrix-valued function of u with $\mathbf{M}(u)\mathbf{1}_J = \mathbf{0}_J$. Another example is

$$C(\mathbf{p}, u) = u + \mathbf{p}'\mathbf{m}(u) + \left[\mathbf{p}'\mathbf{M}(u)\mathbf{p}\right] / \mathbf{p}'\mathbf{n}(u)$$

where $\mathbf{m}(u)$ and $\mathbf{M}(u)$ are as above and $\mathbf{n}(u)$ is a $J-$vector function of u with $\mathbf{n}(u)'\mathbf{1}_J = 0$. These formulations allow for nonlinearity of log-cost in log-prices, but because of Corollary 1, log-cost cannot be quadratic in log-prices, and therefore ESI budget shares cannot be linear in log prices.

2.2 Exact Affine Stone Index (EASI) Demands

As shown in the previous section, ESI Hicks and Pseudo-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 the ESI. Specifically, instead of imposing the ESI restriction that u be ordinally equivalent $x - \mathbf{p}'\mathbf{w}$, define Exact Affine Stone Index (EASI) cost functions to be cost functions that have the property that u is ordinally equivalent an affine transformation of $x - \mathbf{p}'\mathbf{w}$. Here, u is ordinally equivalent to log real expenditures y defined by

$y = \left(x - \mathbf{p}'\mathbf{w}\right)\widetilde{s}(\mathbf{p}) - \widetilde{t}(\mathbf{p})$ for some functions $\widetilde{s}(\mathbf{p})$ and $\widetilde{t}(\mathbf{p})$. For reasons that will be clear shortly, it will be more convenient to express this relationship as

$$u = y = \frac{x - \mathbf{p}'\mathbf{w} - t(\mathbf{p})}{1 + s(\mathbf{p})} \tag{6}$$

for some functions $s(\mathbf{p})$ and $t(\mathbf{p})$.

This implies

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

so that an EASI cost function $C(\mathbf{p}, u)$ must satisfy

$$C(\mathbf{p}, u) - u - \mathbf{p}'[\nabla_{\mathbf{p}} C(\mathbf{p}, u)] = t(\mathbf{p}) + s(\mathbf{p})u \tag{7}$$

The following two theorems characterize EASI cost functions, and hence the solutions to equation (7), 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 (6) 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 κ. Regularity also requires 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 $\overline{C}(\mathbf{p}, u)$ satisfies the homogeneity conditions required to be an ESI log cost function as given by Theorem 1. Assume $C(\mathbf{p}, u) = \overline{C}(\mathbf{p}, u) + T(\mathbf{p}) + S(\mathbf{p})u$ is a regular log cost function. Then $C(\mathbf{p}, u)$ is an EASI log cost function, so equation (6) holds, with $t(\mathbf{p}) = T(\mathbf{p}) - \mathbf{p}'\left[\nabla_{\mathbf{p}} T(\mathbf{p})\right]$ and $s(\mathbf{p}) = S(\mathbf{p}) - \mathbf{p}'\left[\nabla_{\mathbf{p}} S(\mathbf{p})\right]$.

Based on Theorems 1, 2 and 4, a general class of EASI cost functions is

$$C(\mathbf{p}, u) = u + \mathbf{p}'\mathbf{m}(u) + T(\mathbf{p}) + S(\mathbf{p})u \tag{8}$$

where $\mathbf{m}(u)'\mathbf{1}_J = 1$ and both $\exp T(\mathbf{p})$ and $\exp S(\mathbf{p})$ are homogeneous of degree zero in $\exp(\mathbf{p})$. This is similar to the class of cost functions $C(\mathbf{p}, u) = u + \mathbf{p}'\mathbf{m}(u) + \mathbf{p}'M(u)\mathbf{p}$ proposed by Pendakur and Sperlich (2005), though they estimate ordinary Marshallian demands for their model by numerically solving for u.

2.3 Preference Heterogeneity and Functional Form

For empirical work we wish to explicitly include preference heterogeneity in the model, both from observable sources $\mathbf{z}$ and unobservable sources ε. Typical elements of $\mathbf{z}$ would include household size, age, and composition. A simple way to include these in the cost function without interfering with required price homogeneities is to include both of them in the vector of functions $\mathbf{m}(u)$ and to allow the observable components to enter T and S, giving the general class of EASI cost functions

$$C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{p}'\mathbf{m}(u, \mathbf{z}, \varepsilon) + T(\mathbf{p}, \mathbf{z}) + S(\mathbf{p}, \mathbf{z})u. \tag{9}$$

Based on Theorem 4, this class of cost functions has Pseudo-Marshallian demands $\mathbf{w} = \boldsymbol{\omega}(\mathbf{p}, y, \mathbf{z}, \varepsilon)$ given by

$$\mathbf{w} = \mathbf{m}(y, \mathbf{z}, \varepsilon) + \nabla_{\mathbf{p}}T(\mathbf{p}, \mathbf{z}) + \nabla_{\mathbf{p}}S(\mathbf{p}, \mathbf{z})y \tag{10}$$

where y is given by

$$u = y = \frac{x - \mathbf{p}'\mathbf{w} - T(\mathbf{p}, \mathbf{z}) + \mathbf{p}'\left[\nabla_{\mathbf{p}}T(\mathbf{p}, \mathbf{z})\right]}{1 + S(\mathbf{p}, \mathbf{z}) - \mathbf{p}'\left[\nabla_{\mathbf{p}}S(\mathbf{p}, \mathbf{z})\right]} \tag{11}$$

In our empirical application, we take

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

$$S(\mathbf{p}, \mathbf{z}) = \frac{1}{2}\mathbf{p}'\mathbf{B}\mathbf{p}$$

and

$$\mathbf{m}(u, \mathbf{z}, \varepsilon) = \left(\sum_{r=-1}^{5} \mathbf{b}_r u^r\right) + \mathbf{C}\mathbf{z} + \mathbf{D}\mathbf{z}u + \varepsilon \tag{12}$$

So the resulting log cost function, Hicksian demands, and Pseudo-Marshallian demands are equations (1), (2), and (4), respectively, with y given by equation (3).

The complete set of properties required for the cost function (1) to be EASI and satisfy ordinary cost function regularity properties is as follows. Adding up and required 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 the cost function requires $\partial C(\mathbf{p}, u, \mathbf{z}, \varepsilon)/\partial u > 0$, which implies $\mathbf{p}'\left[-\mathbf{b}_{-1}u^{-2} + \left(\sum_{r=1}^{5}\mathbf{b}_r r u^{r-1}\right) + \mathbf{D}\mathbf{z} + \mathbf{B}\mathbf{p}/2\right] > -1$, and we require concavity of $\exp[C(\mathbf{p}, u)]$. 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 $x, \mathbf{p}, \mathbf{z}, \varepsilon$ 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 $x, \mathbf{p}, \mathbf{z}$ has bounded support.

Apart from the construction of y, the Pseudo-Marshallian demand equations (4) 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 (3) would also be linear in parameters.

This model has Engel curves that are high order polynomials. Unlike Marshallian demands, our Pseudo-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 (12) with $\mathbf{m}(y, \mathbf{z}, \varepsilon) = \mathbf{Bn}(y, \mathbf{z}) + \varepsilon$ 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.

2.4 Approximate Fully Linear Model

The demand functions (4) are linear in parameters except for the terms $\frac{1}{2}\sum_{l=0}^{L} z_l \mathbf{p}'\mathbf{A}_l\mathbf{p}$ and $\mathbf{p}'\mathbf{Bp}$ that appear in the construction of y in (3). 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 real expenditures with nominal expenditures deflated by a Stone price index, that is, replace y with $\widetilde{y}$ defined by

$$\widetilde{y} = x - \mathbf{p}'\overline{\mathbf{w}} \tag{13}$$

for some set of budget shares $\overline{\mathbf{w}}$. Then by comparison with equation (4) we have

$$\mathbf{w} = \sum_{r=-1}^{5} \mathbf{b}_r \widetilde{y}^r + \mathbf{Cz} + \mathbf{Dz}\widetilde{y} + \sum_{l=0}^{L} z_l \mathbf{A}_l \mathbf{p} + \mathbf{Bp}\widetilde{y} + \widetilde{\varepsilon} \tag{14}$$

where $\widetilde{\varepsilon} \approx \varepsilon$ with $\widetilde{\varepsilon}$ defined to make equations (14) hold. We call the model of equations (13) and (14) the *Approximate EASI* model.

The Approximate EASI nests the model $\mathbf{w} = \mathbf{b_0} + \mathbf{b}_1\widetilde{y} + \mathbf{Cz} + \mathbf{Ap} + \widetilde{\varepsilon}$, which is identical to the popular approximate Almost Ideal Demand System (AID). Note, however, that 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 $x - \mathbf{p}'\mathbf{w}$.

The approximate EASI also nests the model $\mathbf{w} = \mathbf{b_0} + \mathbf{b}_1\widetilde{y} + \mathbf{b}_2\widetilde{y}^2 + \mathbf{Cz} + \mathbf{Ap} + \widetilde{\varepsilon}$, 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 Pseudo-Marshallian demand function therefore provides a rationale for the unrestricted model that Blundell, Pashardes, and Weber actually estimated.

The approximate EASI model, substituting equation (13) into (14), 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 $\overline{\mathbf{w}}$ is the sample average of budget shares across consumers. A better approximation to y would be to let $\overline{\mathbf{w}}$ be each consumer's own $\mathbf{w}$, so each consumer has their own Stone index deflator based on their own budget shares. As discussed later, this introduces endogeneity.

In our empirical application, we estimate the approximate model with $\overline{\mathbf{w}} = \mathbf{w}$ using seemingly unrelated regressions, and we estimate the true EASI model using the generalized method of moments. As in the approximate AID system, there is no formal theory regarding the quality of the approximation that uses $\widetilde{y}$ in place of y but we find empirically that approximate model estimates do not differ much from estimates based on the exact y, and provide good starting values for exact model estimation.

2.5 Elasticities and Consumer Surplus

Since y is ordinally equivalent to utility, social welfare functions or inequality in welfare measures can be directly evaluated using estimates of y. We now show how to evaluate the effects of changing prices or other variables in EASI models. We will give results using the general EASI cost function

$$C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{p}'\mathbf{m}(u, \mathbf{z}, \varepsilon) + T(\mathbf{p}, \mathbf{z}) + S(\mathbf{p}, \mathbf{z})u \tag{15}$$

which has Pseudo-Marshallian demands $\mathbf{w} = \mathbf{m}(y, \mathbf{z}, \varepsilon) + \nabla_{\mathbf{p}} T(\mathbf{p}, \mathbf{z}) + \nabla_{\mathbf{p}} S(\mathbf{p}, \mathbf{z})y$ where y is defined by (11).

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 (15) is given by

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

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

$$\begin{aligned} C(\mathbf{p}_1, u, \mathbf{z}, \varepsilon) - C(\mathbf{p}_0, u, \mathbf{z}, \varepsilon) &= (\mathbf{p}_1 - \mathbf{p}_0)' \left[\mathbf{w} - \boldsymbol{\nabla}_{\mathbf{p}} T(\mathbf{p}_0, \mathbf{z}) - \boldsymbol{\nabla}_{\mathbf{p}} S(\mathbf{p}_0, \mathbf{z})y\right] + \\ & \quad T(\mathbf{p}_1, \mathbf{z}) - T(\mathbf{p}_0, \mathbf{z}) + S(\mathbf{p}_1, \mathbf{z})y - S(\mathbf{p}_0, \mathbf{z})y. \end{aligned}$$

For our base empirical model log cost function (1), this log cost of living index expression simplifies to

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

The first term in this cost of living index is the Stone index for the price change, $(\mathbf{p}_1 - \mathbf{p}_0)'\mathbf{w}$. 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 behavioural importance of observed heterogeneity.

Define semielasticities to be derivatives of budget shares with respect to log prices $\mathbf{p}$, log real total expenditures 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 elasticitiy by dividing by that budget share. We provide semi-elasticities because they are easier to present algebraically. Hicksian demands are given by

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

so the Hicksian price semielasticities are

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

These are equivalently the price semielasticities holding real expenditures y fixed, $\nabla_{\mathbf{p}'}\omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)$. Similarly, real-expenditure semielasticities for (15) are given by

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

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

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

Compensated semi-elasticities for our baseline model, the log cost function (1), are linear apart from the construction of y. Compensated price semielasticities are given by

$$\nabla_{\mathbf{p}'}\omega(\mathbf{p}, y, \mathbf{z}, \varepsilon) = \sum_{l=0}^{L} z_l \mathbf{A}_l + \mathbf{B}y, \tag{16}$$

and real-expenditure y semi-elasticities are

$$\nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon) = -\mathbf{b}_{-1} y^{-2} + \left(\sum_{r=1}^{5} \mathbf{b}_r r y^{r-1} \right) + \mathbf{Dz} + \mathbf{Bp}, \tag{17}$$

which can vary quite a bit as y changes. Demographic semi-elasticities are given by

$$\nabla_{z_l} \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon) = \mathbf{c}_l + \mathbf{d}_l \mathbf{y} + \mathbf{A}_l \mathbf{p}. \tag{18}$$

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

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

3 Estimation

3.1 Estimators

We are estimating demand systems with J goods, so as usual we can drop the last equation (the J'th good) from the system and just do estimation on 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 (4).

The approximate EASI, equation (14) with $\widetilde{y}$ given by equation (13), is simple to estimate. If the approximate EASI $\widetilde{\varepsilon}$ is homoskedastic and uncorrelated with $\mathbf{p}$, $\mathbf{z}$, $\mathbf{z}\widetilde{y}$, $\mathbf{p}\widetilde{y}$, $\mathbf{p}z_l$, for $l = 1, ..., L$ and $\widetilde{y}^r$ for $r = -1, ..., 5$, then the approximate EASI fits the form of linear seemingly unrelated regressions (SUR). Without imposing symmetry of the $\mathbf{A}_l$ and $\mathbf{B}$ matrices, estimating each equation separately by ordinary least squares is consistent and equivalent to SUR. Symmetry of the $\mathbf{A}_l$ and $\mathbf{B}$ matrices imposes linear cross-equation equality constraints on the coefficients which requires a system estimator such as SUR. If the above homoskedasticity and uncorrelatedness assumptions hold and $\widetilde{\varepsilon}$ is normal then the SUR estimator is asymptotically equivalent to maximum likelihood. If $\widetilde{\varepsilon}$ is not normal or is heteroskedastic the SUR remains consistent, but in that case it will be more efficient to use the generalized method of moments (GMM), based on the moments defined by $\mathbf{p}$, $\mathbf{z}$, $\mathbf{z}\widetilde{y}$, $\mathbf{p}\widetilde{y}$, $\mathbf{p}z_l$, and $\widetilde{y}^r$ uncorrelated with $\widetilde{\varepsilon}$ and by $\widetilde{\varepsilon}$ having mean zero. 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 approximate EASI estimates provide useful parameter starting values for exact model estimation.

The EASI model without approximation that we estimate has equation (3) substituted into equation (4) to give

$$\begin{aligned} \mathbf{w} \;=\; & \sum_{r=-1}^{5} \mathbf{b}_r \left(\frac{x - \mathbf{p}'\mathbf{w} + \sum_{l=0}^{L} z_l \mathbf{p}'\mathbf{A}_l\mathbf{p}/2}{1 - \mathbf{p}'\mathbf{B}\mathbf{p}/2} \right)^r + \mathbf{C}\mathbf{z} + \sum_{l=0}^{L} z_l \mathbf{A}_l \mathbf{p} + \\ & (\mathbf{D}\mathbf{z}+\mathbf{B}\mathbf{p}) \left(\frac{x - \mathbf{p}'\mathbf{w} + \sum_{l=0}^{L} z_l \mathbf{p}'\mathbf{A}_l\mathbf{p}/2}{1 - \mathbf{p}'\mathbf{B}\mathbf{p}/2} \right) + \varepsilon. \end{aligned} \tag{19}$$

Equation (19) 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 (19) 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, probably because variation in the endogenous portion of y, that is, variation in the $\mathbf{p}'\varepsilon$ component of $\mathbf{p}'\mathbf{w}$, is small 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 ε, we use an instrumental variables estimator. Let $\mathbf{r}$ be an $M-$vector of observable variables that are uncorrelated with ε, which will be used as instruments for estimation. If $E(\varepsilon \mid x, \mathbf{p}, \mathbf{z}) = \mathbf{0}_J$ then $\mathbf{r}$ can include any bounded functions of $\mathbf{p}$, $\mathbf{z}$, and x, for example, $\mathbf{p}$, $\mathbf{z}$, and $\widetilde{y}$. 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 $\mathbf{r}$. Let $\mathbf{r} = (r_1,...,r_M)'$. It then follows from $E(\varepsilon \mathbf{r}_m) = \mathbf{0}_J$ for $m = 1,...,M$ so

$$E\left[\left(\begin{array}{c} \mathbf{w} - \sum_{r=-1}^{5} \mathbf{b}_r \left(\frac{x-\mathbf{p}'\mathbf{w}+\sum_{l=0}^{L} z_l\mathbf{p}'\mathbf{A}_l\mathbf{p}/2}{1-\mathbf{p}'\mathbf{B}\mathbf{p}/2} \right)^r - \mathbf{C}\mathbf{z} - \sum_{l=0}^{L} z_l \mathbf{A}_l\mathbf{p} - \\ (\mathbf{D}\mathbf{z}+\mathbf{B}\mathbf{p}) \left(\frac{x-\mathbf{p}'\mathbf{w}+\sum_{l=0}^{L} z_l\mathbf{p}'\mathbf{A}_l\mathbf{p}/2}{1-\mathbf{p}'\mathbf{B}\mathbf{p}/2} \right) \end{array}\right) r_m \right] = \mathbf{0}_J \tag{20}$$

for $m = 1, ..., M$. If ε is homoskedastic, then one possible estimator based on these moment conditions is nonlinear three stage least squares. Otherwise, allowing for generally heteroskedastic ε, 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.

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 $l = 0, ..., L$ and $\mathbf{B}$ are all symmetric matrices. All of these parametric restrictions may imposed as a set of linear constraints on the GMM problem.

Writing this system linearly as equation (4) suggests that good instruments $\mathbf{r}$ 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 $E(\varepsilon \mid \mathbf{p}, x, \mathbf{z}) = \mathbf{0}_J$ and take $\mathbf{r}$ to be $\mathbf{p}$, $\mathbf{z}$, $\mathbf{z}\overline{y}$, $\mathbf{p}\overline{y}$, $\mathbf{p}z_1$, ...,$\mathbf{p}z_L$, and $\overline{y}^r$ for $r = -1, ..., 5$ with $\overline{y}$ defined as

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

where $\overline{\mathbf{w}}$ is the average budget shares across consumers in our sample, and $\overline{\mathbf{A}}_l$ and $\overline{\mathbf{B}}$ 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 $\overline{\mathbf{w}}$ and inconsistency of the estimates of $\overline{\mathbf{A}}_l$ and $\overline{\mathbf{B}}$ (due to their coming from the approximate model) in the construction of $\overline{y}$ only affects the quality of the instruments $\mathbf{r}$ and hence the efficiency of the GMM estimation, but does cause inconsistency, because $\mathbf{r}$ remains uncorrelated with ε given any choice of values of the parameters $\overline{\mathbf{w}}$, $\overline{\mathbf{A}}_l$ and $\overline{\mathbf{B}}$ in the construction of $\mathbf{r}$. When symmetry of $\mathbf{A}_l$ and $\mathbf{B}$ is not imposed, this set of moments $E(\varepsilon r_m) = 0$ exactly identifies the EASI model parameters. Imposing symmetry reduces the number of distinct parameters, yielding overidentifcation.

Since we have assumed $E(\varepsilon \mid \mathbf{p}, x, \mathbf{z}) = \mathbf{0}_J$, we could interpret $\sum_{r=-1}^{5} \mathbf{b}_r y^r + \mathbf{C}\mathbf{z} + \mathbf{D}\mathbf{z}y$ 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 $y^r z_k^s$ for integers r, s. Ai and Chen (2003) provide associated rates, 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 GMM estimator can be readily modified to deal with possible measurement error or endogeneity in some regressors, by suitably modifying $\mathbf{r}$. 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 $\overline{\mathbf{p}}$ everywhere $\mathbf{p}$ appears in the construction of the instuments $\mathbf{r}$, where $\overline{\mathbf{p}}$ are the fits from regressing $\mathbf{p}$ on supply side instruments. If log total nominal expenditure x suffers from measurement error then one might use functions of income (eg, log income) instead of functions of x in $\mathbf{r}$ to instrument the problem. Similarly, if unobserved preference heterogeneity ε 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 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 $\widehat{\varepsilon}$, 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 parametric GMM also remains consistent regardless of heteroskedasticity in ε, so for example the estimates are consistent if $\varepsilon = \mathbf{N}(x, \mathbf{z})\varepsilon^*$ where ε^* 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). If the model is correctly specified then imposing these constraints would not be binding asymptotically. A common practice in empirical demand analysis is to estimate without imposing inequality constraints, and then check that they are satistifed for a reasonable range of $\mathbf{p}$, x, and $\mathbf{z}$ values. These inequalities can be readily checked 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 (14) with $\widetilde{y}$ given by equation (13), using SUR. We then consistently estimate the full EASI model using GMM with instruments and starting values as described in the previous section. This is the model that has equation (3) substituted into equation (4) to give equation (19).

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 available for 12 years in 5 regions (Atlantic, Quebec, Ontario, Prairies and British Columbia) yielding 60 distinct price vectors. Prices are normalised so that the price vector facing residents of Ontario in 1986 is (1, ..., 1), that is, these observations define the base price vector $\overline{\mathbf{p}} = \mathbf{0}_J$. Since the model contains a high-order polynomial in y, we subtract 7.5 from y. Translating y by a constant is completely absorbed by changes in the $\mathbf{b}_r$ 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.

Table 1 gives summary statistics for 19,276 observations of rental-tenure households located in cities with at least 30,000 residents. The empirical analysis uses annual expenditure in nine expenditure categories: food-in, food-out, rent, clothing, household operation, household furnishing & equipment, private transportation operation, public transportation and personal care. Personal care is the left-out equation, yielding eight expenditure share equations which depend on 9 log-prices and log-expenditure. These expenditure categories account for about three quarters of the current consumption of the households in the sample.

We use $L = 4$ observable demographic characteristics in our model: (1) a childless couple dummy; (2) the age of the household head minus 40; (3) the log of the number of household members; and (4) a single parent dummy. These variables are all zero for a single childless adult aged 40. When symmetry is not imposed on $\mathbf{A}_l$ and $\mathbf{B}$, then given the instruments specified above, the model is exactly identified. In this case, the model has $[7 + L + L + (J - 1) * (L + 1) + (J - 1)] (J - 1) = 504$ parameters for $\mathbf{b}$, $\mathbf{C}$, $\mathbf{D}$, $\mathbf{A}_l$, and $\mathbf{B}$ (63 coefficients in each of 8 equations). When symmetry is imposed on $\mathbf{A}_l$ and $\mathbf{B}$, the model has $(L + 2) (J - 1) (J - 2) /2 = 168$ restrictions, which are overidentifying restrictions for testing exogeneity and are also symmetry restrictions. With starting values obtained by seemingly unrelated regressions estimation of the approximate EASI model, we obtained convergence of the GMM estimator in only three iterations, due to the near linearity of the EASI model.

Table 2 gives R^2 values for each equation for various models, showing that most of the variation in budget shares is unexplained heterogeneity. A very important feature of our approach is that this unexplained heterogeneity is interpreted as individual variation in utility function parameters, and as such has measurable consequences in welfare calculations (as we demonstrate below). Table 3 gives tests for various hypotheses. Although Table 3 uses some tests from a symmetry-unrestricted model, to save space we do not present the 504 individual parameter estimates of the unrestricted model. However, Tables A1-A4 in the appendix do provide all 336 estimated coefficients for our symmetry-restricted baseline models. The left side of each table gives GMM estimates for the exact model, and the right side of the table gives estimates for the linear approximate model estimated using SUR, as discussed in the previous section.

We note that this is a model with an extremely large number of parameters, so despite the large sample size one must be very cautious regarding the following statistical test results, since it is difficult to assess the quality of asymptotic approximations in models of this size.

Define Q_n as the value of the GMM criterion function, and let $\widehat{Q}_n$ and $\overline{Q}_n$ denote the unrestricted and symmetry-restricted minimised values of Q_n, respectively. Exact identification of the unrestricted model implies $\widehat{Q}_n = 0$. We therefore obtain a test statistic for overidentifying restrictions of $N\overline{Q}_n = 2171$, distributed asymptotically under the null as a χ^2_{168}. The 1% critical value of the χ^2_{168} is 214. A score test of Slutsky symmetry, and hence of the overidentifying restrictions, is given by N times the gradient of Q_n in the metric of its Hessian evaluated at the symmetry-restricted estimates. The value of this test statistic is 1621, and it is also asymptotically distributed as a χ^2_{168} under the null. Finally, a Wald test of the same hypothesis, constructed from the unrestricted estimates, yields a test statistic of 2669 with the same asymptotic distribution under the null. In all these cases, rejection could be due to endogeneity or other invalidity of some instruments, violation of Slutsky symmetry, some other model misspecification, or poor asymptotic approximations (this applies to all later tests as well), given our caveat regarding the size of the model.

The p-values in Table 3 show that we reject almost every joint hypothesis regarding model restrictions or simplifications. Slutsky symmetry requires that $\mathbf{A}_l = \mathbf{A}_l'$ and $\mathbf{B} = \mathbf{B}'$, which is rejected as we noted above. In addition, we see in the table that this rejection is driven neither solely by $\mathbf{B} \neq \mathbf{B}'$ nor $\mathbf{A}_l \neq \mathbf{A}_l'$, since we reject the hypothesis that $\mathbf{A}_l = \mathbf{A}_l'$ for all l and we separately reject the hypothesis that $\mathbf{B} = \mathbf{B}'$. The second group of tests considers the various interactions between $\mathbf{p}$, $\mathbf{z}$ and $\mathbf{y}$ that we include in our model. Testing indicates that all of the two-way interactions that we include in our baseline model are statistically important. One such interaction test considers if $\mathbf{D} = \mathbf{0}$, which is of economic interest because of Shape Invariance (see the appendix for details). Another interaction test looks at whether $\mathbf{B} = \mathbf{0}$. If this holds then y simplies into log stone-index deflated expenditures plus a function of $\mathbf{p}, \mathbf{z}$. In this case the demand functions are closer to pure polynomials (if $\mathbf{b}_{-1}$ is also zero then the demand functions are pure polynomials in $\mathbf{p}$, $\mathbf{z}$ and the Stone index deflated expenditures $\mathbf{x} - \mathbf{p}'\mathbf{w}$). As one would expect in a model that is closer to linear in regressors, when $\mathbf{B} = \mathbf{0}$ is imposed we find that the SUR approximation improves, giving estimates that are closer to GMM.

We also consider testing down the complexity of the Engel curve functions in Table 3. Engel curves in our baseline model are characterised primarily by the seven $\mathbf{b}_r$ parameter vectors, and we consider successive deletions down to a quadratic specification. We find that all the y^5 terms are together only significant at the ten percent level, and we strongly reject the hypothesis that the model is a lower than fourth order polynomial (reestimation excluding the y^5 terms does not change our reported results in any meaningful way).

3.3 Estimated Engel Curves

Table A1 in the appendix gives estimated $\mathbf{b}_r$, $\mathbf{C}$, and $\mathbf{D}$ coefficients and standard errors that determine the Engel curves implied by our estimated model. The easiest way to summarize these parameter estimates is to examine the resulting expenditure share equations as functions of x for particular values of $\mathbf{p}$, $\mathbf{z}$, and ε. At $\mathbf{p} = \mathbf{0}_J$ log real expenditure y equals log nominal expenditure x, so at these base prices we obtain Marshallian Engel curves $\mathbf{w} = \sum_{r=-1}^{5} \mathbf{b}_r x^r + \mathbf{C}\mathbf{z} + \mathbf{D}\mathbf{z}x + \varepsilon$. Figures 1-8 show these estimated Engel curves from our model for a reference household consisting of a single childless adult aged 40 with $\varepsilon = \mathbf{0}$, and for a household comprised of 2 adults and 2 children with the head aged 40 with $\varepsilon = \mathbf{0}$. The base-period Engel curves for households with different values of unobserved heterogeneity are identical except for being vertically shifted by ε. These base-period Engel curves are also informative about the shape of real expenditure Engel curves in other price regimes, since at other price vectors, real expenditure Engel curves differ only by the addition of the linear function $\sum_{l=0}^{L} \mathbf{A}_l z_l \mathbf{p} + \mathbf{B}\mathbf{p}y$.

Estimates are computed at each percentile of real expenditures for each household type, and estimated 90% confidence intervals (computed via the delta method) are displayed with small crosses for the GMM estimates at each decile of the real expenditure distribution. Each figure has 4 lines: thick black lines indicate GMM estimates of the exact model for the reference type; thin black lines indicate GMM estimates of the exact model for families of 4, thick grey lines indicate SUR estimates of the approximate model for the reference type, and dotted black lines indicate GMM estimates of the exact model constrained so that Engel curves are quadratic.

Our quadratic specification is very similar to models estimated by Blundell, Pashardes, and Weber (1993) and Banks, Blundell and Lewbel (1997). Our quadratic model differs from the Quadratic Almost Ideal (QAI) model of Banks, Blundell, and Lewbel (1997) in that it has a different deflator for total expenditures, and does not require the coefficient of squared log real-expenditures to depend on prices. Our specification also allows for more general demographic effects than is typical in applied QAI models, which usually only have terms similar to $\mathbf{C}$, but not $\mathbf{A}_l$ or $\mathbf{D}$.

Figures 1 and 2 show Engel curves for food-in (food consumed at home) and food-out. Both these Engel curves are almost linear, so it is not surprising that there is hardly any variation across the specifications. Indeed, for both goods, the SUR approximate model and the GMM quadratic model lie within the 90% pointwise confidence intervals of the exact GMM model.

Figure 3 shows the Engel curve for rent. For the bottom half of the real expenditure distribution of reference households, the quadratic estimates lie outside the confidence intervals of the EASI estimates. The Engel curve for rent is roughly linear above the 30th percentile for single childless adults and roughly quadratic below that percentile. The Engel curve for couples with 2 children looks somewhat different: its curvature flattens out and even reverses above the 80th percentile. This kind of complexity cannot be captured in a quadratic model but is readily accomodated in the EASI framework. The approximate

SUR estimates are not entirely within the 90% confidence intervals around the exact model GMM estimates as was the case in Figures 1 and 2. However, the SUR approximation does capture the changes in curvature in various parts of the distribution quite well, as is especially obvious in the bottom decile of the distribution.

Figures 4, 5 and 6 give the household operation, household furnishing & equipment and clothing Engel curves. In these cases, the quadratic model does reasonably well. The SUR estimates of the approximate model also do reasonably well, although they fall somewhat below the GMM estimates of the exact model in the lower half of the expenditure distribution in the household furnishing & equipment equation. Although it may be of dubious statistical significance, we note that the clothing equation looks more like two linear segments joined together with a curve than a quadratic function.

Figure 7 gives the private transportation operation Engel curve. As in the rent equation, the underlying complexity of the Engel curve is too great for a quadratic model to capture. Below the median, the curvature is positive, but above the median it is negative. Thus, we have the GMM quadratic estimates outside the 90% confidence intervals of the GMM exact model for the bottom half of the distribution of expenditure for the reference household. The Engel curve for couples with 2 children is also too complex for a quadratic function to capture, because its curvature changes too much over the expenditure distribution. The SUR estimates of the approximate model, however, fit very well in this case, lying almost on top of the GMM estimates throughout the distribution of expenditure.

Figure 8 gives the public transportation Engel curve. Here, the Engel curve given by the GMM exact model exhibits an S-shape, but it is clear from the pointwise confidence intervals that the quadratic model does a tolerable job in a statistical sense. Here again the SUR estimates of the approximate model lie almost on top of the GMM estimates of the exact model.

There are two important lessons that we draw from these figures. First, while the demand functions of some goods are close to linear or quadratic in real expenditure, other goods such as rent and private transportation are not quadratic, requiring higher order elements of polynomials in real expenditure. This implies demand rank higher than three. Second, the SUR estimates of the approximate model do quite a good job of approximating the GMM estimates of the exact model, even when the underlying Engel curves are quite complex.

3.4 Estimated Price Effects

Table 4 presents compensated price semi-elasticities for the same household types as in the Figures, evaluated at the medians of their respective real expenditure distributions. Assuming Slutsky symmetry, compensated semi-elasticities are given by the matrix $\mathbf{A}_0 + \mathbf{B}y$

for the reference household type, and by $\sum_{l=0}^{L} \mathbf{A}_l z_l + \mathbf{B}y$ for other types of households. These matrices may also be calculated from parameters estimated without imposing symmetry, though in that case the notion of compensated demands is not well defined. To summarize the precision of estimated price effects and the magnitude of symmetry violations, Table 4 presents estimates for symmetry-restricted and unrestricted models. The restricted estimates are compensated semi-elasticities, and the unrestricted estimates are differences between semi-elasticities that would be equal under symmetry.

Considering first the symmetry-restricted estimates of compensated semi-elasticities, we see that many of these price effects are large and statistically significant. For example, for reference households with median expenditure, the cross-price semi-elasticity between food-in and food-out is 0.058, which implies that a 20% increase in the price of food-in would result in an increase of the food-out share of 1.2 percentage points. Given that (from Figure 2) the expenditure share at base prices for this type of household is 9.5%, this is a fairly large effect. Further, it is estimated quite precisely with a t value of 5.12. There are many other examples of large and significant cross price effects, especially those involving the rent price.

Table 4 also allows us to directly assess the importance of the symmetry restrictions. For example, for single childless adults with median expenditure, the food-in, food-out cross-price semi-elasticity is 0.058 in the symmetry-restricted model. However, in the unrestricted model the difference with its symmetric partner is 0.072 which is larger than the symmetry-restricted estimate. Thus, for this cross-price semi-elasticity, the violation of symmetry is both large in an economic sense and statistically significant (with a t-statistic of 2.48).

Considering other rows of Table 4, we see similar evidence that the violation of symmetry may be both statistically and economically significant. This violation could be due to several factors. Unobserved preference heterogeneity has often been invoked to explain symmetry violations (see, e.g., Lewbel 2001). However, our model incorporates additive unobserved heterogeneity effects in the Hicksian (Pseudo-Marshallian) share equations, so this explanation would require some more elaborate model of unobserved heterogeneity. A second explanation for the symmetry violation could be unobserved price heterogeneity, that is, violations of the law of one price. A third explanation is that households might engage in bargaining or otherwise fail to maximize a simple joint utility function (as in, e.g., Browning, Chiapporri and Lewbel 2004), which results in weaker symmetry restrictions. However, this does not help with the rejection of symmetry we observe for single-member households. Other explanations could include poor asymptotic approximations due to the size of the model, measurement errors in prices or other variables, violations of separability across goods or over time, functional form misspecification, endogeneity of prices or other instruments, or the failure of consumers to maximise regular utility functions. We nevertheless use symmetry restricted estimates for our policy analyses, since basic consumer surplus and cost of living analyses assume utility maximization and hence depend upon symmetry.

Table 5 gives Slutsky terms from the symmetry restricted model for our two household types at median expenditure. Denoting the matrix of compensated semi-elasticities as Φ, the matrix of Slutsky terms which give the Hessian of (unlogged) cost with respect to (unlogged) prices is given by $\Phi + \mathbf{w}\mathbf{w}' - diag(\mathbf{w})$ where $diag(\mathbf{w})$ is a diagonal matrix with $\mathbf{w}$ on the main diagonal. Concavity of cost in prices requires that this matrix be negative semidefinite, and hence that the matrix without the omitted good be negative definite, including negative own price elements (the diagonal of this matrix). The Slutsky matrices reported in Table 5 for these two household types at median expenditure without the omitted good are indeed negative definite, with significantly negative own-price elements.

An interesting feature of our model is that it allows us to assess how substitution effects (compensated semi-elasticities) may differ between rich and poor. In our model, that dependence is governed by the matrix **B**. Table 6 gives the estimated elements of the matrix **B** along with t-statistics in italics below each estimate. Here, we see some important dependence of substitution effects on log real-expenditure y. Substitution effects for rent depend strongly on expenditure. In Table 4, we saw that at median expenditure the compensated rent own-price semi-elasticity for reference households is 0.70. Given that the 90-10 percentile difference in log real-expenditure for single childless adults is about 1, the estimate of 0.15 for the corresponding element of **B** suggests that the compensated rent own-price semi-elasticity is about 0.15 larger for the richest than the poorest households, which is a large difference given the magnitude of the semi-elasticity at the median.

We draw three main conclusions from the analysis of price effects. First, we are able to obtain precise estimates of compensated semi-elasticities, which given symmetry are second-order derivatives of the log-cost function and capture substitution effects. These estimates suggests that price effects, and therefore substitution effects, are large in magnitude. Second, we are somewhat cautious about interpreting the symmetry restricted estimates in light of the rejection of symmetry, though we require symmetry restricted estimates for consumer surplus and related policy analyses. Third, the concavity restrictions seem to present less trouble than symmetry restrictions, at least in the neighborhood of moderate expenditure values.

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:

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

where $a_l^{rent,rent}$ and $b^{rent,rent}$ are the rent own-price elements of $\mathbf{A}_l$ and $\mathbf{B}$. This function can be evaluated for each observation facing the base price vector. Here, unobserved heterogeneity enters only through the level effect on w^{rent} (because at the base price vector, $y = x$ is independent of unobserved heterogeneity). 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 and is not affected by unobserved heterogeneity. Traditional consumer demand analysis which ignores unobserved heterogeneity would accomodate 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 accomodates unobserved heterogeneity, but would not incorporate the second-order term which captures substitution effects. The EASI model combines the advantages of both approaches.

Figure 9 shows the estimated values of the cost-of-living index for each household incorporating unobserved heterogeneity with empty circles and estimated values for each household with unobserved heterogeneity set to zero ($\varepsilon = \mathbf{0}$) with filled circles. In addition, the second-order component capturing substitution effects is shown with filled squares. The reason that the filled circles do not lie on a single line is that variation in demographic characteristics $\mathbf{z}$ across households affects the surplus measures. Figure 9 covers a larger span of expenditure than Figures 1-8 because other demographic types, including rich childless couples, are included in Figure 9 but not in Figures 1-8.

The underlying Engel curve is visible in the estimates which zero out ε, 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 in cost-of-living impacts is related to expenditure, and only a little is related to other characteristics. The more refined picture is that only a little is related to *observed* characteristics, with most of the variation attributable to unobserved characteristics.

The unobserved hetereogeneity distribution includes estimation error, and could also be due in part to measurement errors rather than true preference heterogeneity. Given our model estimates, alternative distributions of consumer surplus could be obtained by drawing from alternative simulated ε distributions.

In addition to showing the importance of accounting for unobserved heterogeneity, Figure 9 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.

The second-order terms capture substitution 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 9. Substitution 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 this under-estimate is about 0.3 percentage points for the richest households (on an impact of about 4 percentage points), and the magnitude of the over-estimate is about 0.1 percentage points for the poorest households (on an impact of about 6 percentage points).

4 Conclusions

We have provided a utility derived consumer demand system, the EASI demand system, that is as flexible in price responses and as close to linear in parameters as the Almost Ideal Demand System, but also allows for flexible interactions between prices and expenditures, and permits almost any functional form for Engel curves. Error terms in the model correspond to unobserved preference heterogeneity random utility parameters. Demand functions are estimated by GMM, and, like the AID system, an approximate model can be estimated by linear regression. We empirically estimate the model allowing Engel curves to be high order polynomials, and find some significant departures from linear and quadratic demands.

The model is an application of a new general methodology we propose for constructing demand systems, which we call pseudo-Marshallian demands. These are essentially the Hicks demands associated with cost functions having the property that utility can be represented by a simple function of observables. In EASI demands this function is an affine transform of total expenditures deflated by a Stone index.

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 Economic Studies, 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 (1993), "Adult-Equivalence Scales and the Economic Implementation of Interpersonal Comparisons of Well-Being," Social Choice and Welfare, 10, 335-61.

[5] Blundell, Richard, Xiaohong Chen and Dennis Kristensen (2003), Semi-Nonparametric IV Estimation of Shape-Invariant Engel Curves, Centre for Microdata Methods and Practice, working paper CWP15/03.

[6] Blundell, Richard, Alan Duncan and Krishna Pendakur (1998), "Semiparametric Estimation of Consumer Demand", Journal of Applied Econometrics 13, 435-461.

[7] Blundell, Richard and Arthur Lewbel (1991), "The Information Content of Equivalence Scales." Journal of Econometrics 50, 49-68.

[8] Blundell, Richard. W., Panos Pashardes, and Guglielmo Weber (1993), ”What do we learn about Consumer Demand Patterns from Micro Data?”, American Economic Review, vol. 83, 570-597.

[9] 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.

[10] Brown, Bryan W. and Mary Beth Walker (1989), “The Random Utility Hypothesis and Inference in Demand Systems” Econometrica 57, 815-829

[11] 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.

[12] Browning, Martin and Irena Thomas, (1999) "Prices for the FAMEX," unpublished working paper, Department of Economics, McMaster University.

[13] Browning, Martin, (2001), Modelling Commodity Demands and Labour Supply with M-Demands, unpublished working paper, Institute of Economics, University of Copenhagen.

[14] Deaton, Angus S., and John Muellbauer (1980), ”An Almost Ideal Demand System,” American Economic Review, vol. 70, 312-326.

[15] Diewert, W. E. and T. J. Wales (1987), "Flexible Functional Forms and Global Curvature Restrictions," Econometrica, 55. 43-68.

[16] Gorman, William M. "Tricks with utility functions", in Artis, M. and A. R. Nobay (eds.), Essays in Economic Analysis, Cambridge University Press (1976).

[17] 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.

[18] Hansen, Lars, (1982), “Large Sample Properties of Generalized Method of Moments Estimators,” Econometrica, 50, 1029-1054.

[19] Hildenbrand, Werner, (1994), Market Demand, Princeton University Press.

[20] Lewbel, Arthur (1989). "Household Equivalence Scales and Welfare Comparisons," Journal of Public Economics, 39, 377-91.

[21] Lewbel, Arthur (1991), ”The Rank of Demand Systems: Theory and Nonparametric Estimation,” Econometrica, 59, 711-730.

[22] Lewbel, Arthur (2001), "Demand Systems with and without Errors," American Economic Review, 91, 611-618

[23] 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.

[24] Pendakur, Krishna (1999) "Estimates and Tests of Base-Independent Equivalence Scales," Journal of Econometrics, 88, 1-40.

[25] Pendakur, Krishna (2002) "Taking Prices Seriously in the Measurement of Inequality," Journal of Public Economics, 86, 47-69.

[26] 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.

[27] Pollak, Robert A & Wales, Terence J, (1980) "Comparison of the Quadratic Expenditure System and Translog Demand Systems with Alternative Specifications of Demographic Effects," Econometrica, 48, 595-612.

[28] 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).

[29] 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 proofs, extensions, and additional mathematical properties of our models, which are relevant for evaluating these models and for other applications of our general methodology.

5.1 Proofs

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 (5) we have $C(\mathbf{p}, u) - u = \mathbf{p}' \left[\nabla_{\mathbf{p}} C(\mathbf{p}, u)\right]$. This is equivalent to $d(\mathbf{p}, u) = \mathbf{p}' \left[\nabla_{\mathbf{p}} d(\mathbf{p}, u)\right]$, which by the Euler Theorem shows that $d(\mathbf{p}, u)$ is linearly homogeneous in $\mathbf{p}$.

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 equalling the derivative of a linearly homogeneous function.

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 λ and κ, let $\lambda = 1/(p_2 - p_1)$ and $\kappa = -p_1/(p_2 - p_1)$ to obtain $d(0, 1, u) = \left[d(p_1, p_2, u)/\left(p_2 - p_1\right)\right] - \left[p_1/\left(p_2 - p_1\right)\right]$. The theorem holds with $m(u) = 1 - d(0, 1, u)$.

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 (7) is equivalent to $d(\mathbf{p}, u) - t(\mathbf{p}) - s(\mathbf{p})u = \mathbf{p}' \left[\nabla_{\mathbf{p}} d(\mathbf{p}, u)\right]$. Taking derivatives of this expression with respect to u gives

$$\begin{aligned}\nabla_u d(\mathbf{p},u) - s(\mathbf{p}) &= \mathbf{p}'\left[\nabla_{\mathbf{p}}\nabla_u d(\mathbf{p},u)\right]\\ \nabla_u^2 d(\mathbf{p},u) &= \mathbf{p}'\left[\nabla_{\mathbf{p}}\nabla_u^2 d(\mathbf{p},u)\right]\end{aligned}$$

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 $\nabla_u^2 d(\mathbf{p},u) = \mathbf{p}'\left[\nabla_{\mathbf{p}}\nabla_u^2 d(\mathbf{p},u)\right]$. 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 (7). Given exactness, Equation (7) implies $t(\mathbf{p})+s(\mathbf{p})u = u - C(\mathbf{p},u)+\mathbf{p}'\boldsymbol{\omega}(\mathbf{p},u)$. Ordinary homogeneity of cost functions and Marshallian demands requires that $\exp\left[u - C(\mathbf{p},u)+\mathbf{p}'\boldsymbol{\omega}(\mathbf{p},u)\right]$ 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 $s(\mathbf{p}+\mathbf{1}_J\kappa) = s(\mathbf{p})+\kappa$.

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

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

where the right side of the second equality follows from Stone index exactness for $\overline{C}(\mathbf{p},u)$ from Theorem 1. Comparing this equation to (7) proves the result.

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

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 $T(\mathbf{p})$ and $S(\mathbf{p})$ and having $y = u$ positive means that $T(\mathbf{p}) + S(\mathbf{p})u$ is concave, which suffices for cost function concavity. For monotonicity $\nabla_u C(\mathbf{p},u,\mathbf{z},\varepsilon) = 1 + \mathbf{p}'\nabla_y\mathbf{m}(y,\mathbf{z},\varepsilon) + S(\mathbf{p},\mathbf{z})$, which is positive given $1 + \inf\left[\mathbf{p}'\nabla_y\mathbf{m}(y,\mathbf{z},\varepsilon)\right] > -\inf\left[S(\mathbf{p},\mathbf{z})\right]$. If $S(\mathbf{0}_J,\mathbf{z}) \geq -1$ then by concavity of S, $S(\mathbf{p},\mathbf{z}) - \mathbf{p}'\left[\nabla_{\mathbf{p}}S(\mathbf{p},\mathbf{z})\right] \geq -1$ so the denominator of the expression for y in equation (11) is positive. Similarly, by concavity of T, the numerator of this expression for y is greater than $x - \mathbf{p}'\mathbf{w} - T(\mathbf{0}_J,\mathbf{z})$, so y is positive if $x > \sup[\mathbf{p}'\mathbf{w} + T(\mathbf{0}_J,\mathbf{z})]$. Now each element of $\mathbf{w}$ is bounded between zero and one, and boundedness of the supports of $\mathbf{p}$, and $T(\mathbf{0}_J, \mathbf{z})$ means that $\sup[\mathbf{p}'\mathbf{w} - T(\mathbf{0}_J, \mathbf{z})]$ is finite. Rescaling expenditures $\exp(x)$ to smaller units adds a constant to x, so $\inf(y) > 0$ if units are small enough to make $\inf(x) > \sup[\mathbf{p}'\mathbf{w} + T(\mathbf{0}_J, \mathbf{z})]$.

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

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

where y is given by equation (11). This uniquely defines ε 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 ε as preference parameters) and hence of a unique value of $\mathbf{w}$ associated with each possible value of $\mathbf{p}$, x, $\mathbf{z}$, ε, even though we do not have a closed form analytic expression for it.

5.2 Extensions

Instead of Stone index related constructions, we could more generally define real log total expenditures by

$$y = \frac{x - g(\mathbf{w}, \mathbf{p})}{1 + s(\mathbf{p})} \tag{21}$$

for any deflator function g. Having $y = u$ would then require

$$C(\mathbf{p}, u) = [1 + s(\mathbf{p})]u + g(\nabla_{\mathbf{p}} C(\mathbf{p}, u), \mathbf{p}) \tag{22}$$

and the resulting Pseudo-Marshallian demand functions would be

$$\mathbf{w} = \omega\left[\mathbf{p}, \frac{x - g(\mathbf{w}, \mathbf{p})}{1 + s(\mathbf{p})}\right]. \tag{23}$$

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

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

Theorem 5 provides a very weak sufficient condition for existence of pseudo-Marshallian demands $\mathbf{w} = \omega\left[\mathbf{p}, x - g(\mathbf{w}, \mathbf{p})\right]$ 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, then $C(\mathbf{p}, u) = u + t(\mathbf{p})$ for some function t and we may let $g(\mathbf{w}, \mathbf{p}) = -t(\mathbf{p})$. 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, $w_1 = \omega_1(\mathbf{p}, u)$, is invertible in u, so $u = \omega_1^{-1}(\mathbf{p}, w_1)$. Then a deflator g that makes $u = y$ is $g(\mathbf{w}, \mathbf{p}) = C(\mathbf{p}, \omega_1^{-1}(\mathbf{p}, w_1)) - \omega_1^{-1}(\mathbf{p}, w_1)$.

The function g satisfying equation (22) 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 Pseudo-Marshallian demands is not existence, but rather convenience for demand estimation, which is why this paper focused on examples where g has simple closed forms, like the log Stone deflator.

A generalization of the EASI class of cost functions (8) is

$$x = C(\mathbf{p}, u) = u + \mathbf{c}(\mathbf{p})'\mathbf{m}(u) + T(\mathbf{p}) + S(\mathbf{p})u \tag{24}$$

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

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

Solving this expression for $\mathbf{m}(u)$ and substituting the result into (24) gives

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

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

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

which we can write as

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

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

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

This is a generalization of EASI demands where y is an affine transform of $x - \overline{\mathbf{c}}(\mathbf{p})'\mathbf{w}$ instead of an affine transform of $x - \mathbf{p}'\mathbf{w}$. We could also introduce $\mathbf{z}$ into $\mathbf{c}(\mathbf{p})$. 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.

5.3 Global Regularity

Theorem 6: Assume a log cost function in the general EASI form of equation (9) with $\mathbf{1}'_J\mathbf{m}(u, \mathbf{z}, \varepsilon) = 1$, $\exp T(\mathbf{p}, \mathbf{z})$ and $\exp S(\mathbf{p}, \mathbf{z})$ homogeneous of degree zero in $\exp(\mathbf{p})$, and $S(\mathbf{0}_J, \mathbf{z}) > -1$. Then sufficient conditions for the cost function to be regular everywhere on the support of the data are concavity of $T(\mathbf{p}, \mathbf{z})$ and $S(\mathbf{p}, \mathbf{z})$ in $\mathbf{p}$, $\inf(y) > 0$, and $1 + \inf\left[\mathbf{p}'\nabla_y\mathbf{m}(y, \mathbf{z}, \varepsilon)\right] > \inf\left[S(\mathbf{p}, \mathbf{z})\right]$, where these infimums are over the support of the data. Also, if $T(\mathbf{p}, \mathbf{z})$ and $S(\mathbf{p}, \mathbf{z})$ are concave in $\mathbf{p}$, $\inf[S(\mathbf{0}_J, \mathbf{z})] > -1$, and the supports of x, $\mathbf{p}$, and $T(\mathbf{0}_J, \mathbf{z})$ are bounded, then $\inf(y) > 0$ if expenditures are measured in sufficiently small units.

By Theorem 6, the cost function can be globally (i.e., over the entire supports of the data) concave if $T(\mathbf{p}, \mathbf{z})$ and $S(\mathbf{p}, \mathbf{z})$ are bounded and concave and if prices and total expenditures have bounded support. In our empirical model based on the cost function (1), given boundedness of prices, $T(\mathbf{p}, \mathbf{z})$ and $S(\mathbf{p}, \mathbf{z})$ are bounded and concave if $\sum_{l=0}^{L} z_l\mathbf{A}_l$ and $\mathbf{B}$ are negative semidefinite.

For monotonicity in our empirical model, assume y is everywhere positive, let $\lambda_0 = \left[\inf(x) - \sup\left(\mathbf{p}'\mathbf{1}_J - \sum_{l=0}^{L} z_l\mathbf{p}'\mathbf{A}_l\mathbf{p}/2\right)\right]/[1 + \sup(-\mathbf{p}'\mathbf{B}\mathbf{p})]$ and let $\lambda_1 = \sup(x)$. It follows from equation (3) that $\lambda_0 \leq y \leq \lambda_1$. Let $\tau = -\inf\left[1 + \mathbf{p}'\mathbf{D}\mathbf{z} + \mathbf{p}'\mathbf{B}\mathbf{p}/2\right]$. 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. Then the sufficient monotonicity condition in Theorem 6 is satisfied if $-c_{-1}y^{-2} + \sum_{r=1}^{5} c_r r y^{r-1} > \tau$ 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 a lower bound on the slope of every Engel curve at 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 guarantee global regularity of demand systems can induce undesirable restrictions on flexibility. See, e.g., Diewert and Wales (1987) and Ryan and Wales (1998). We therefore follow the common practice of estimating without imposing the inequality constraints associated with regularity, and afterwards check them in the neighborhood of the data.

5.4 Coherency and Invertibility

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. In the demand system context (see, e.g., van Soest, Kapteyn, and Kooreman (1993), Brown and Matzkin (1998) or 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 ε 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 (9) that is regular. Assume $\widetilde{\mathbf{m}} = \mathbf{m}(u, \mathbf{z}, \varepsilon)$ is invertible in ε, so we may write $\varepsilon = \mathbf{m}^{-1}(u, \mathbf{z}, \widetilde{\mathbf{m}})$. Then the resulting budget share demand functions are coherent and invertible.

Our functional form for empirical work, the cost function (1), 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 (12), $\mathbf{m}^{-1}(u, \mathbf{z}, m) = \mathbf{m} - \left(\sum_{r=-1}^{5} \mathbf{b}_r u^r\right) - \mathbf{Cz} - \mathbf{Dz}u$. 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 $\mathbf{n}$.

A desirable property both for theory and estimation is to have the distribution of ε across individuals be independent of $\mathbf{p}, x, \mathbf{z}$. Theorem 7 does not impose this condition, and it may not hold in our cost function (1) because with budget shares bounded between zero and one, the support of $\varepsilon = \mathbf{w} - \omega(\mathbf{p}, y, \mathbf{z}, \mathbf{0}_J)$ may depend upon $\mathbf{p}, x, \mathbf{z}$. These errors could also be heteroskedastic if they have the property noted by Hildenbrand (1994) of variance increasing in x.

However, coherence and invertibility will also be satisfied if $\mathbf{m}(y, \mathbf{z}, \varepsilon) = \mathbf{Bn}(y, \mathbf{z}) + \mathbf{N}(y, \mathbf{z}, \varepsilon)$ where $\mathbf{N}$ is mean zero and invertible in ε, which would generate Pseudo-Marshallian budget errors given by $\mathbf{N}$. An open question is whether we can define a function $\mathbf{N}$ that results in independently distributed errors ε, but it may be possible to do so using methods from Matzkin (2005), e.g., sequentially defining each ε_j to equal the conditional distribution of $\mathbf{N}_j$ conditioning on covariates and on ε_1,...,ε_{j-1}.

5.5 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 horizontal translation 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

$$\omega(\mathbf{p}, u, \mathbf{z}, \varepsilon) = \nabla_{\mathbf{p}} f\left[\mathbf{p}, h(u, \mathbf{z}, \varepsilon)\right] + \nabla_{\mathbf{p}} g(\mathbf{p}, \mathbf{z}, \varepsilon).$$

We may invert C with respect to utility to give indirect utility

$$h(\cdot, u, \mathbf{z}, \varepsilon) = f^{-1}\left[\cdot, x - g(\mathbf{p}, \mathbf{z}, \varepsilon)\right],$$

and substituting this expression into Hicks demands yields

$$\mathbf{w} = \nabla_{\mathbf{p}} f\left[\mathbf{p}, x - g(\mathbf{p}, \mathbf{z}, \varepsilon)\right] + \nabla_{\mathbf{p}} g(\mathbf{p}, \mathbf{z}, \varepsilon).$$

This structure for demands may be very complex over x, but the characteristics $\mathbf{z}$ and ε 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 $h(u, \mathbf{z}, \varepsilon) = \overline{h}(u)$ holds then

$$C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = f\left[\mathbf{p}, \overline{h}(u)\right] + g(\mathbf{p}, \mathbf{z}, \varepsilon)$$

so

$$\ln E(\mathbf{p}, u, \mathbf{z}, \varepsilon) = g(\mathbf{p}, \mathbf{z}, \varepsilon) \tag{25}$$

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 (9) 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

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

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

Shape-invariance is easily imposed on our empirically estimated model given by (1). 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

$$C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{d}'\mathbf{z} + \mathbf{p}'\left[\sum_{r=-1}^{5} \mathbf{b}_r u^r + \mathbf{Cz} + \varepsilon\right] + \frac{1}{2}\sum_{l=0}^{L} z_l \mathbf{p}'\mathbf{A}_l\mathbf{p} + \frac{1}{2}\mathbf{p}'\mathbf{Bp}u, \quad (26)$$

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, log real-expenditure y includes the constant term $\mathbf{d}'\mathbf{z}$ and is given by

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

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.

5.6 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 (9), (10), and (11), or in our empirically estimated model defined by equations (1), (3), and (4), 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 ε terms, these are examples of the models described in the extensions section, specifically equation (24) with $\overline{\mathbf{c}}(\mathbf{p}) = \mathbf{c}(\mathbf{p}) = \mathbf{p} + \mathbf{k}$.

To check possible sensitivity of our empirical results to unit scaling, we tried to reestimate out 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 units.

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

$$C(\mathbf{p}, u, \mathbf{z}, \varepsilon) = u + \mathbf{p}'\mathbf{m}(u, \mathbf{z}, \varepsilon) + \mathbf{p}'\left[\mathbf{A}_1 + \mathbf{A}_2 h(u, \mathbf{z}, \varepsilon)\right]\mathbf{p}/2 \tag{27}$$

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

$$\begin{aligned} C(\mathbf{p}+\mathbf{k}, u, \mathbf{z}, \varepsilon) &= u + (\mathbf{p}+\mathbf{k})'\mathbf{m}(u, \mathbf{z}, \varepsilon) + (\mathbf{p}+\mathbf{k})'\left[\mathbf{A}_1 + \mathbf{A}_2 h(u, \mathbf{z}, \varepsilon)\right](\mathbf{p}+\mathbf{k})/2 \\ & \quad u^* + \mathbf{p}'\mathbf{m}^*(u^*, \mathbf{z}, \varepsilon) + \mathbf{p}'\left[\mathbf{A}_1 + \mathbf{A}_2 h^*(u^*, \mathbf{z}, \varepsilon)\right]\mathbf{p}/2 \end{aligned}$$

where

$$\begin{aligned} u^* &= u + \mathbf{k}'\mathbf{m}(u, \mathbf{z}, \varepsilon) + \mathbf{k}'\left[\mathbf{A}_1 + \mathbf{A}_2 h(u, \mathbf{z}, \varepsilon)\right]\mathbf{k}/2 \\ \mathbf{m}^*(u^*, \mathbf{z}, \varepsilon) &= \mathbf{m}(u, \mathbf{z}, \varepsilon) + \left[\mathbf{A}_1 + \mathbf{A}_2 h(u, \mathbf{z}, \varepsilon)\right]\mathbf{k} \\ h^*(u^*, \mathbf{z}, \varepsilon) &= h(u, \mathbf{z}, \varepsilon) \end{aligned}$$

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 (27), 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, 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.

5.7 Marshallian Elasticity Calculations

Cost functions in the class of equation (15) have y given equation (3), 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 υ gives

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

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

$$\begin{aligned}\nabla_\upsilon \mathbf{w}(\mathbf{p}, x, \mathbf{z}, \varepsilon) &= \left[I_J - \frac{\left[\nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)\right]\mathbf{p}'}{1 + \mathbf{p}'\mathbf{A}_2\mathbf{p}/2}\right]^{-1} \\ &\quad \left[\nabla_\upsilon \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon) + \left[\nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)\right]\left[\nabla_\upsilon\left(\frac{x - \mathbf{p}'\mathbf{w} - \frac{\mathbf{p}'\mathbf{A}_1\mathbf{p}}{2}}{1 + \mathbf{p}'\mathbf{A}_2\mathbf{p}/2}\right) - \frac{x}{1 + \frac{\mathbf{p}'\mathbf{A}_2\mathbf{p}}{2}}\right]\right]\end{aligned} \tag{28}$$

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

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

where $\nabla_y \omega(\mathbf{p}, y, \mathbf{z}, \varepsilon)$ is given by equation (17). Equation (28) could also be evaluated taking υ to be $\mathbf{p}$ to obtain Marshallian price elasticities, but it is simpler to recover them from the Hicksian $\mathbf{p}$ elasticities (16) and the above Marshallian x elasticities using the Slutsky matrix

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

Finally, again using equation (28), the Marshallian semielasticity with respect to $\mathbf{z}$ is

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

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 (17) and (18).

Some of the above elasticity expressions depend on ε either directly or through $\mathbf{w}$. Given estimated model parameters, ε 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 ε 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.

Table 1: Data Descriptives

Variable		Mean	Std Dev	Minimum	Maximum
expend	foodin	0.216	0.103	0.000	0.809
shares	foodout	0.061	0.063	0.000	0.643
	rent	0.334	0.127	0.000	0.949
	hhoper	0.082	0.049	0.000	0.636
	hhfurneq	0.049	0.057	0.000	0.646
	cloth	0.101	0.063	0.000	0.585
	privtrans	0.093	0.083	0.000	0.591
	pubtrans	0.030	0.039	0.000	0.452
log-prices	foodin	-0.179	0.539	-1.412	0.337
	foodout	-0.106	0.602	-1.458	0.534
	rent	-0.249	0.498	-1.316	0.366
	hhoper	-0.181	0.540	-1.397	0.319
	hhfurneq	-0.135	0.378	-0.939	0.198
	cloth	-0.006	0.386	-0.868	0.427
	privtrans	-0.234	0.595	-1.533	0.528
	pubtrans	-0.153	0.711	-1.580	0.687
	perscare	-0.126	0.463	-1.114	0.337
demo	couple	0.216	0.412	0.000	1.000
graphics	headage	-0.306	11.102	-15.000	24.000
	logsize	0.677	0.585	0.000	2.398
	singlepar	0.105	0.306	0.000	1.000
log-	x	9.379	0.643	6.683	11.613
expenditure	$\mathbf{w'p}$	-0.177	0.498	-1.412	0.465
	$x-\mathbf{w'p}$	9.556	0.456	7.691	11.434

Table 2: R^2 values for each equation

Equation	foodin	foodout	rent	hhoper	hhfurneq	cloth	privtrans	pubtrans
GMM-baseline	0.484	0.192	0.386	0.150	0.115	0.240	0.168	0.033
GMM-quadratic	0.484	0.191	0.378	0.150	0.115	0.238	0.159	0.033
SUR-baseline	0.489	0.188	0.396	0.152	0.114	0.250	0.170	0.034

Table 3: Model Tests

model	restriction	parameters	df	Wald	p-value
unrestricted	symmetry	$\mathbf{A}_l=\mathbf{A}_l'$; $\mathbf{B}=\mathbf{B}'$	168	2669	0
baseline		$\mathbf{A}_l=\mathbf{A}_l'$	140	457	0
		$\mathbf{B}=\mathbf{B}'$	28	100	0
symmetry-restricte	no py term	$\mathbf{B}=0$	36	453	0
baseline	no pz term	$\mathbf{A}_l=0$	144	838	0
	no zy term	$\mathbf{D}=0$	32	1127	0
	no y^5	$\mathbf{b}_5=0$	8	14	0.09
	no y^4,y^5	$\mathbf{b}_4=\mathbf{b}_5=0$	16	138	0
	no y^3,y^4,y^5	$\mathbf{b}_3=\mathbf{b}_4=\mathbf{b}_5=0$	24	249	0
	no y^{-1},y^3,y^4,y^5	$\mathbf{b}_{-1}=\mathbf{b}_3=\mathbf{b}_4=\mathbf{b}_5=0$	32	583	0

Table 4: Compensated Price Semi-Elasticities									
		Single Childless, median y=9.30				Couple, 2 Kids, median y=9.37			
		symmetry-restricted		unrestricted		symmetry-restricted		unrestricted	
		d**w**(**p**,u)/d**p**		d**w(.)**/d**p** - d**w(.)**/d**p**'		d**w(p,u)**/d**p**		d**w(.)**/d**p** - d**w(.)**/d**p**'	
Effect	Equation	Estimate	asy. t-stat	Estimate	asy. t-stat	Estimate	asy. t-stat	Estimate	asy. t-stat
pfoodin	foodin	-0.054	*-3.24*			-0.052	*-2.48*		
pfoodin	foodout	0.058	*5.12*	-0.072	*-2.48*	0.010	*0.79*	-0.091	*-2.60*
pfoodin	rent	-0.081	*-13.02*	0.008	*0.21*	-0.070	*-8.44*	0.052	*1.02*
pfoodin	hhoper	-0.002	*-0.22*	-0.024	*-0.82*	0.023	*1.54*	-0.041	*-0.88*
pfoodin	hhfurneq	0.044	*4.39*	-0.109	*-3.35*	0.077	*6.23*	-0.078	*-1.57*
pfoodin	cloth	0.023	*2.84*	-0.016	*-0.67*	0.018	*1.81*	0.092	*2.69*
pfoodin	privtrans	0.029	*4.79*	0.086	*2.85*	0.045	*5.89*	0.056	*1.60*
pfoodin	pubtrans	-0.042	*-8.55*	-0.046	*-2.68*	-0.081	*-12.86*	-0.155	*-6.95*
pfoodout	foodout	-0.039	*-3.15*			-0.037	*-3.31*		
pfoodout	rent	0.049	*8.83*	-0.097	*-3.38*	0.049	*9.97*	-0.237	*-6.47*
pfoodout	hhoper	0.015	*1.84*	-0.041	*-1.38*	-0.001	*-0.13*	-0.009	*-0.31*
pfoodout	hhfurneq	-0.008	*-0.93*	0.112	*3.72*	0.013	*1.39*	0.018	*0.62*
pfoodout	cloth	-0.052	*-7.59*	0.037	*1.72*	-0.028	*-3.96*	0.095	*4.01*
pfoodout	privtrans	0.007	*1.18*	-0.063	*-2.63*	-0.020	*-4.23*	0.008	*0.31*
pfoodout	pubtrans	-0.029	*-6.35*	-0.003	*-0.21*	0.026	*6.00*	0.038	*2.63*
prent	rent	0.070	*7.71*			0.045	*4.21*		
prent	hhoper	-0.001	*-0.17*	0.090	*2.40*	0.007	*1.34*	0.002	*0.04*
prent	hhfurneq	-0.027	*-5.97*	-0.070	*-1.84*	-0.039	*-7.30*	-0.083	*-1.64*
prent	cloth	-0.020	*-4.74*	0.051	*2.23*	-0.026	*-5.17*	0.006	*0.21*
prent	privtrans	-0.028	*-4.82*	-0.014	*-1.08*	0.002	*0.34*	-0.038	*-2.39*
prent	pubtrans	0.051	*15.78*	0.054	*4.25*	0.040	*10.74*	0.056	*3.47*
phhoper	hhoper	0.027	*2.34*			0.039	*2.44*		
phhoper	hhfurneq	-0.025	*-2.74*	0.000	*0.01*	-0.057	*-4.80*	0.029	*0.92*
phhoper	cloth	0.008	*1.23*	0.074	*3.39*	-0.002	*-0.24*	-0.043	*-1.49*
phhoper	privtrans	-0.027	*-6.87*	-0.053	*-1.74*	-0.036	*-6.86*	-0.087	*-2.58*
phhoper	pubtrans	0.028	*7.59*	0.015	*0.96*	0.025	*4.99*	0.098	*5.58*
phhfurneq	hhfurneq	-0.038	*-2.96*			-0.019	*-1.26*		
phhfurneq	cloth	0.031	*3.69*	-0.061	*-2.52*	0.019	*1.83*	-0.096	*-3.19*
phhfurneq	privtrans	-0.001	*-0.29*	-0.008	*-0.27*	0.004	*0.92*	0.008	*0.23*
phhfurneq	pubtrans	-0.012	*-2.82*	-0.031	*-1.80*	-0.032	*-6.77*	0.020	*1.09*
pcloth	cloth	0.032	*4.22*			0.071	*7.82*		
pcloth	privtrans	0.007	*1.65*	0.027	*1.42*	-0.022	*-4.63*	0.019	*0.90*
pcloth	pubtrans	0.001	*0.33*	-0.008	*-0.72*	-0.002	*-0.42*	0.006	*0.43*
pprivtrans	privtrans	0.016	*2.29*			0.014	*1.80*		
pprivtrans	pubtrans	-0.001	*-0.30*	0.001	*0.10*	0.023	*6.33*	0.013	*1.11*
ppubtrans	pubtrans	0.003	*0.92*			-0.007	*-2.08*		

Table 5: Implied Slutsky Terms, Symmetry-Restricted Estimates					
		single childless, median *y=9.30*		couples, 2 kids, median *y* =9.37	
Effect	Equation	Estimate	asy. t-stat	Estimate	asy. t-stat
pfoodin	foodin	-0.180	*-11.28*	-0.271	*-13.14*
pfoodin	foodout	0.072	*6.34*	0.022	*1.63*
pfoodin	rent	-0.021	*-3.22*	0.030	*3.39*
pfoodin	hhoper	0.008	*0.75*	0.049	*3.26*
pfoodin	hhfurneq	0.049	*4.83*	0.087	*6.87*
pfoodin	cloth	0.035	*4.33*	0.040	*4.01*
pfoodin	privtrans	0.044	*7.04*	0.072	*9.17*
pfoodin	pubtrans	-0.036	*-7.21*	-0.072	*-11.21*
pfoodout	foodout	-0.124	*-9.53*	-0.071	*-5.96*
pfoodout	rent	0.087	*15.10*	0.060	*11.68*
pfoodout	hhoper	0.021	*2.69*	0.002	*0.17*
pfoodout	hhfurneq	-0.005	*-0.58*	0.014	*1.50*
pfoodout	cloth	-0.044	*-6.54*	-0.026	*-3.66*
pfoodout	privtrans	0.016	*2.82*	-0.017	*-3.61*
pfoodout	pubtrans	-0.025	*-5.52*	0.027	*6.22*
prent	rent	-0.171	*-19.17*	-0.168	*-16.48*
prent	hhoper	0.027	*7.44*	0.032	*6.17*
prent	hhfurneq	-0.014	*-2.99*	-0.029	*-5.33*
prent	cloth	0.013	*3.02*	-0.005	*-0.96*
prent	privtrans	0.011	*1.83*	0.028	*4.30*
prent	pubtrans	0.068	*20.01*	0.048	*12.55*
phhoper	hhoper	-0.037	*-3.14*	-0.036	*-2.17*
phhoper	hhfurneq	-0.023	*-2.50*	-0.055	*-4.61*
phhoper	cloth	0.014	*2.14*	0.004	*0.40*
phhoper	privtrans	-0.020	*-5.17*	-0.030	*-5.56*
phhoper	pubtrans	0.031	*8.36*	0.028	*5.44*
phhfurneq	hhfurneq	-0.069	*-5.52*	-0.049	*-3.34*
phhfurneq	cloth	0.033	*4.03*	0.021	*2.05*
phhfurneq	privtrans	0.002	*0.39*	0.007	*1.43*
phhfurneq	pubtrans	-0.010	*-2.48*	-0.031	*-6.57*
pcloth	cloth	-0.044	*-5.26*	0.007	*0.66*
pcloth	privtrans	0.015	*3.53*	-0.016	*-3.45*
pcloth	pubtrans	0.005	*1.41*	0.000	*0.10*
pprivtrans	privtrans	-0.070	*-9.98*	-0.062	*-8.22*
pprivtrans	pubtrans	0.003	*0.92*	0.025	*6.93*
ppubtrans	pubtrans	-0.038	*-12.45*	-0.034	*-9.50*

Table 6: Slope of Compensated Semi-Elasticities with respect to *y*

Share	Price foodin	foodout	rent	hhoper	hhfurneq	cloth	privtrans	pubtrans
foodin	0.064	0.022	0.020	-0.057	-0.028	-0.021	-0.021	0.037
	2.24	*1.24*	*1.86*	*-2.89*	*-1.68*	*-1.56*	*-2.05*	*4.26*
foodout		-0.021	-0.043	0.040	-0.010	-0.008	0.037	-0.036
		-1.32	*-5.61*	*2.93*	*-0.71*	*-0.72*	*4.93*	*-5.39*
rent			0.151	0.003	-0.035	-0.052	-0.035	0.003
			10.07	*0.49*	*-4.38*	*-6.98*	*-4.44*	*0.51*
hhoper				0.020	0.013	0.005	0.006	-0.015
				0.94	*0.82*	*0.46*	*0.82*	*-2.23*
hhfurneq					-0.013	0.068	-0.010	0.009
					-0.60	*4.70*	*-1.37*	*1.28*
cloth						-0.003	0.003	0.010
						-0.24	*0.38*	*1.98*
privtrans							0.026	-0.006
							2.62	*-1.28*
pubtrans								0.001
								0.25

Figure 1: Estimated Food-In Shares: SUR & GMM; polynomial & quadratic; childless singles & couples with 2 kids

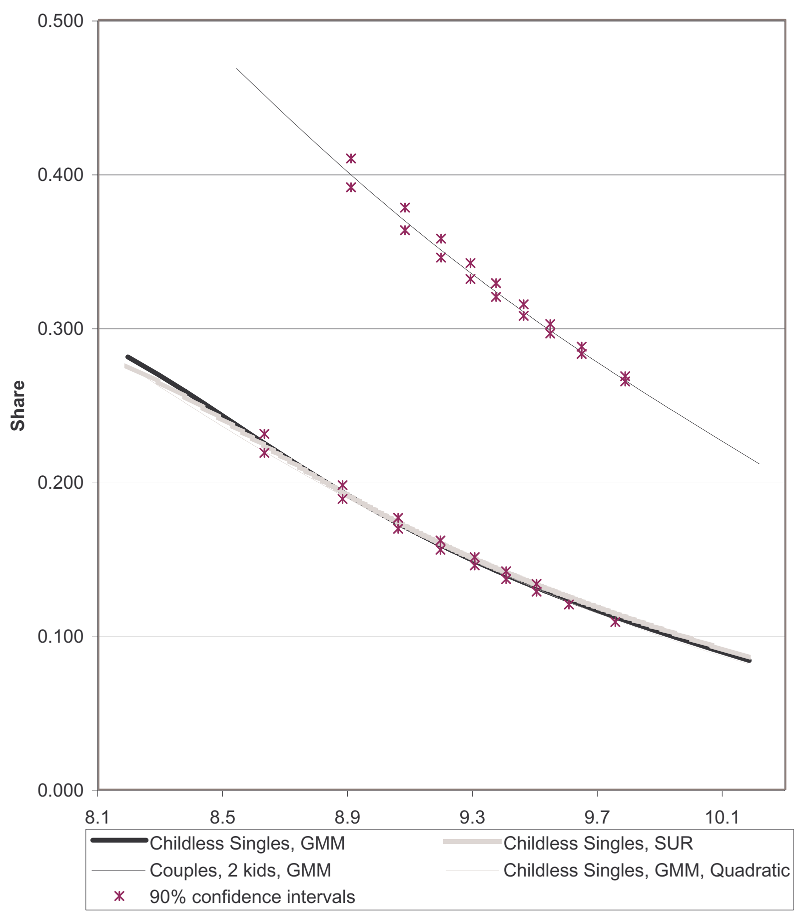


Figure 2: Estimated Food-Out Shares: SUR & GMM; polynomial & quadratic; childless singles & couples with 2 kids

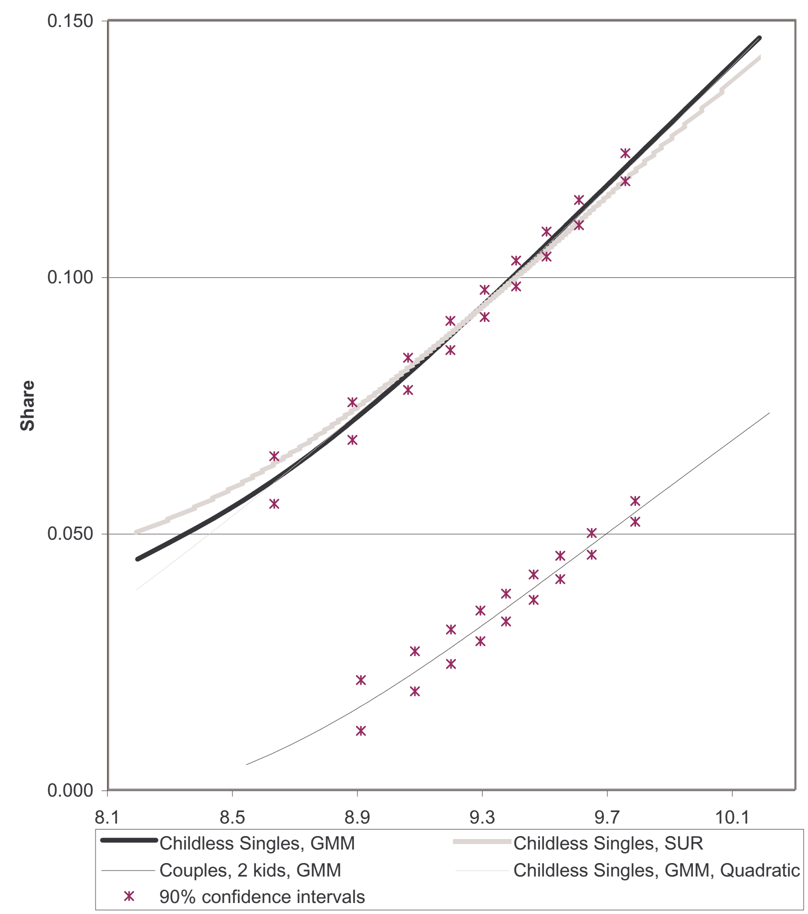


Figure 3: Estimated Rent Shares: SUR & GMM; polynomial & quadratic; childless singles & couples with 2 kids

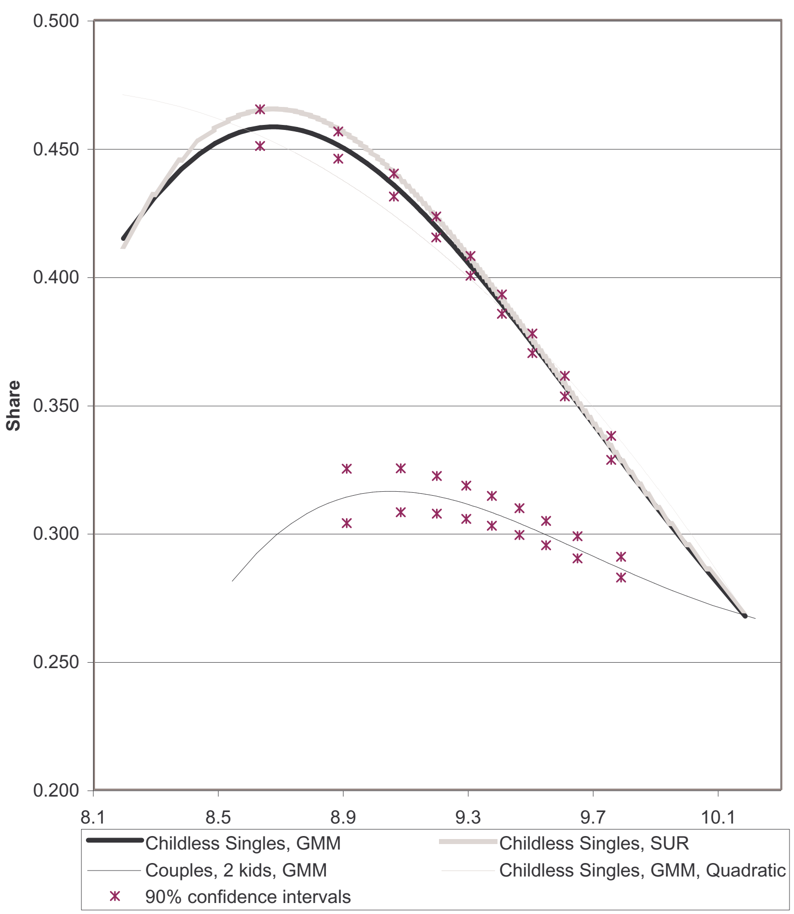
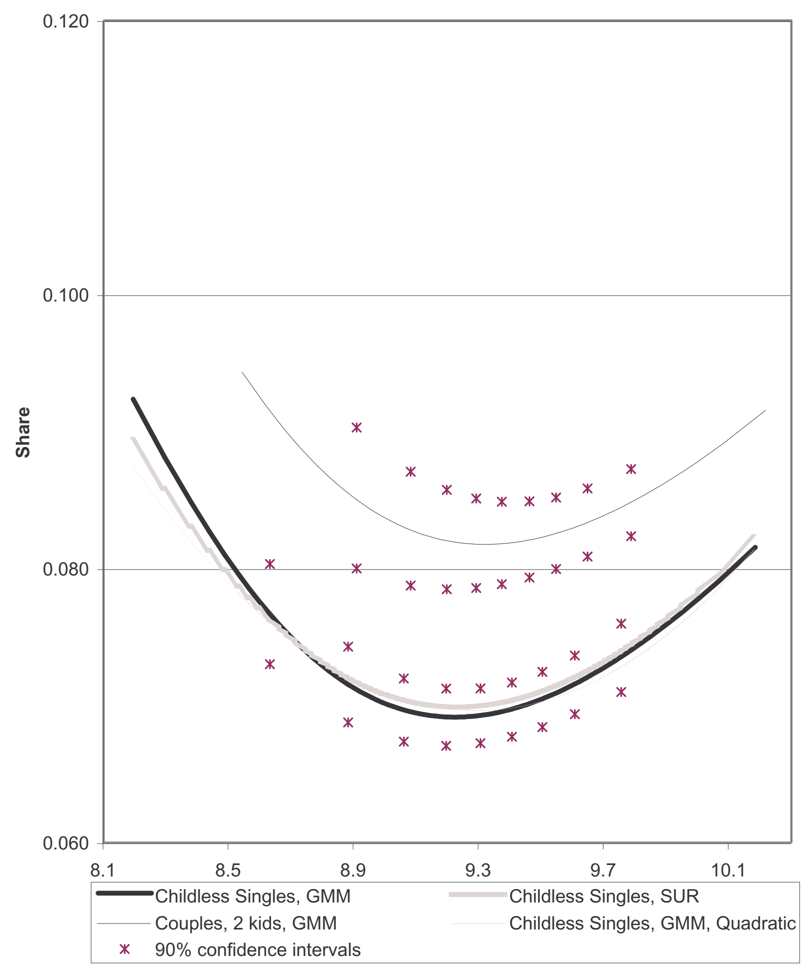


Figure 4: Estimated Household Operation Shares: SUR & GMM; polynomial & quadratic; childless singles & couples with 2 kids



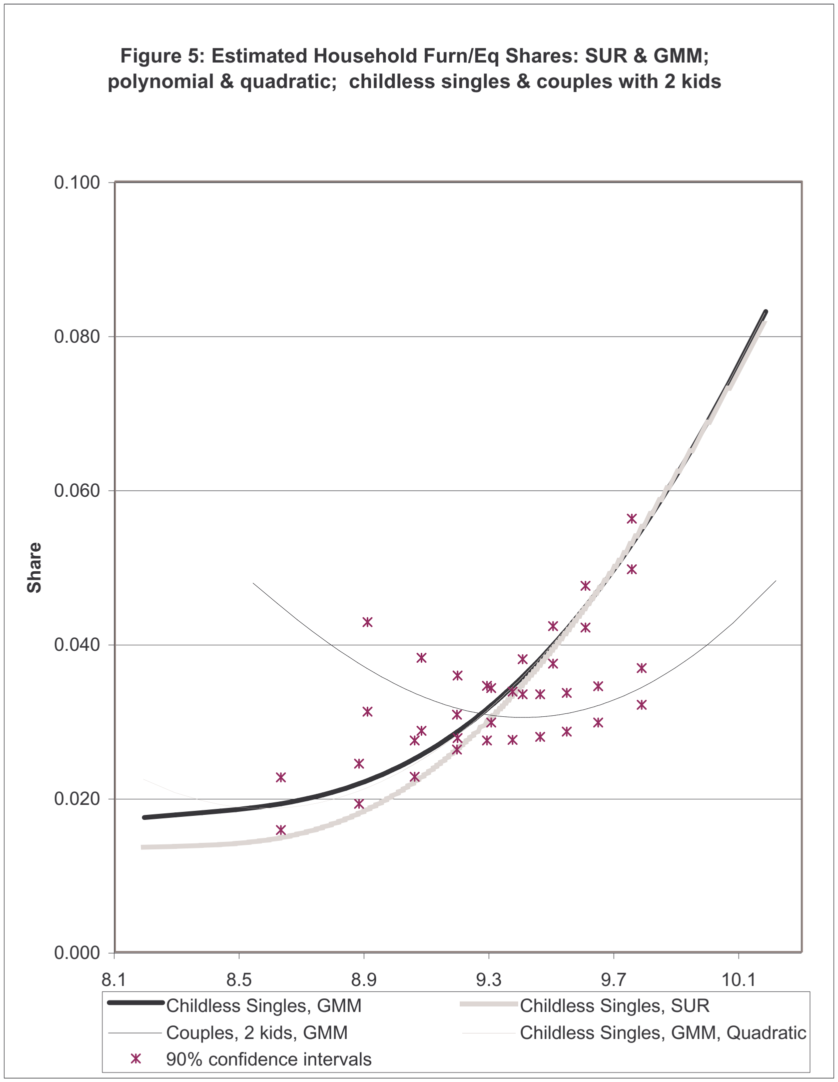


Figure 5: Estimated Household Furn/Eq Shares: SUR & GMM;
polynomial & quadratic; childless singles & couples with 2 kids
Share
0.100
0.080
0.060
0.040
0.020
0.000
8.1
8.5
8.9
9.3
9.7
10.1
Childless Singles, GMM
Childless Singles, SUR
Couples, 2 kids, GMM
Childless Singles, GMM, Quadratic
90% confidence intervals

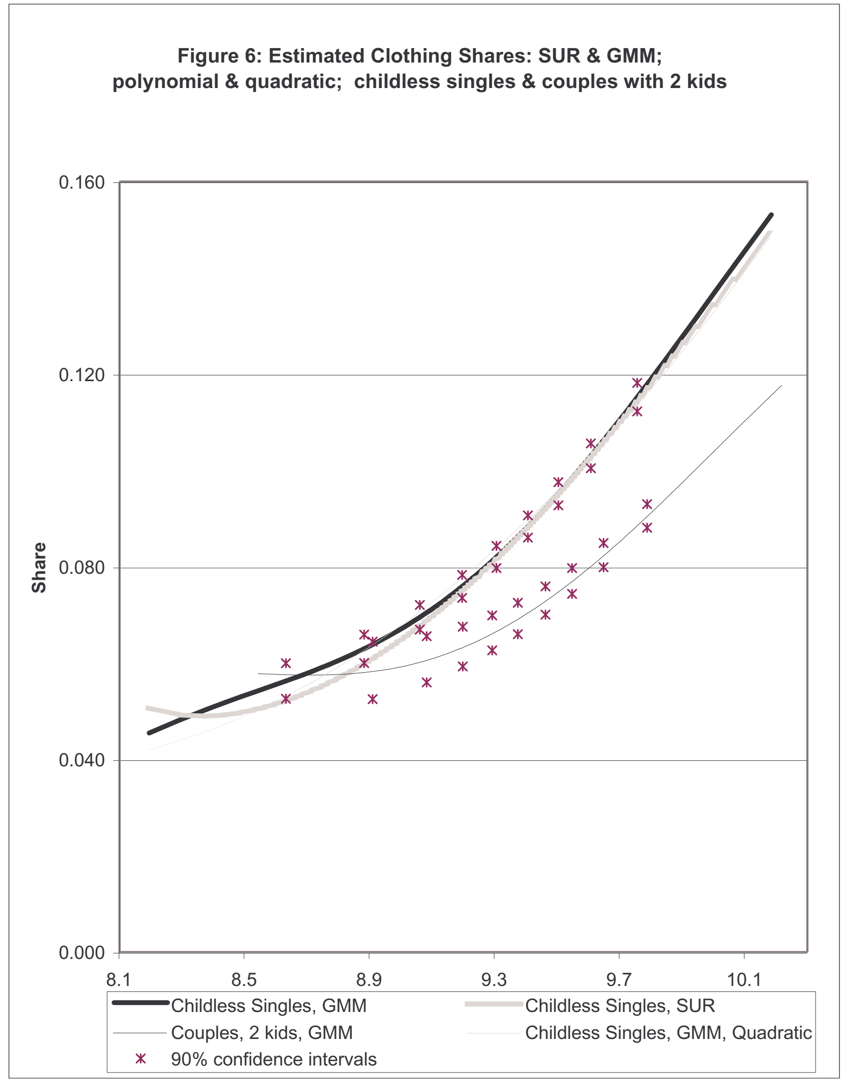


Figure 6: Estimated Clothing Shares: SUR & GMM;
polynomial & quadratic; childless singles & couples with 2 kids
Share
0.160
0.120
0.080
0.040
0.000
8.1
8.5
8.9
9.3
9.7
10.1
Childless Singles, GMM
Childless Singles, SUR
Couples, 2 kids, GMM
Childless Singles, GMM, Quadratic
90% confidence intervals

Figure 7: Estimated Private Trans Shares: SUR & GMM; polynomial & quadratic; childless singles & couples with 2 kids

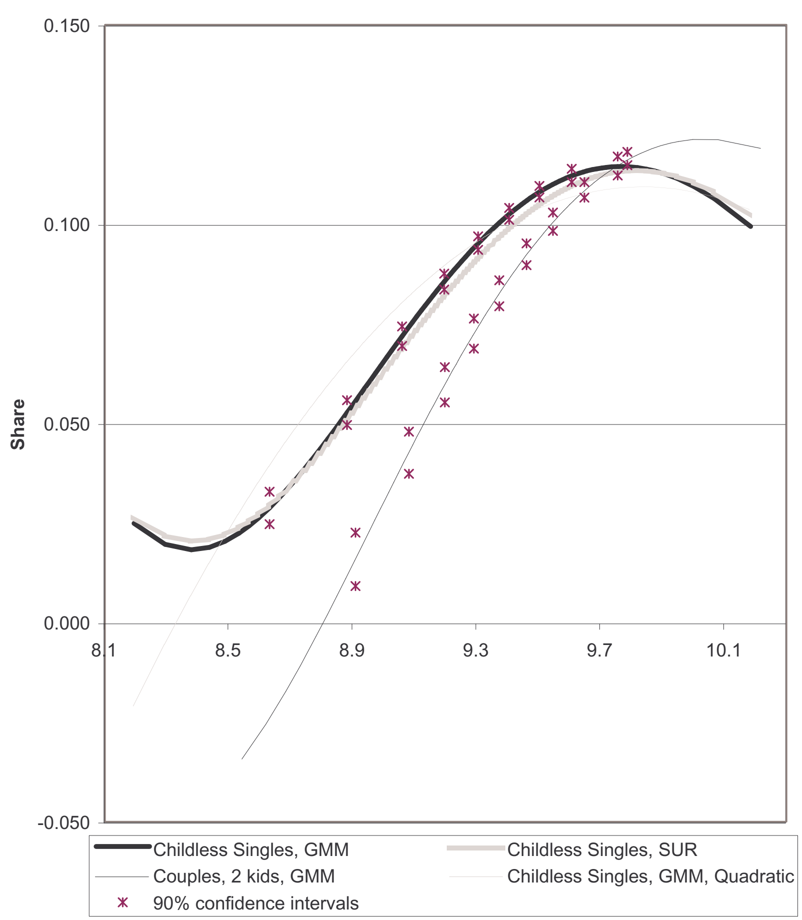


Figure 8: Estimated Public Trans Shares: SUR & GMM; polynomial & quadratic; childless singles & couples with 2 kids

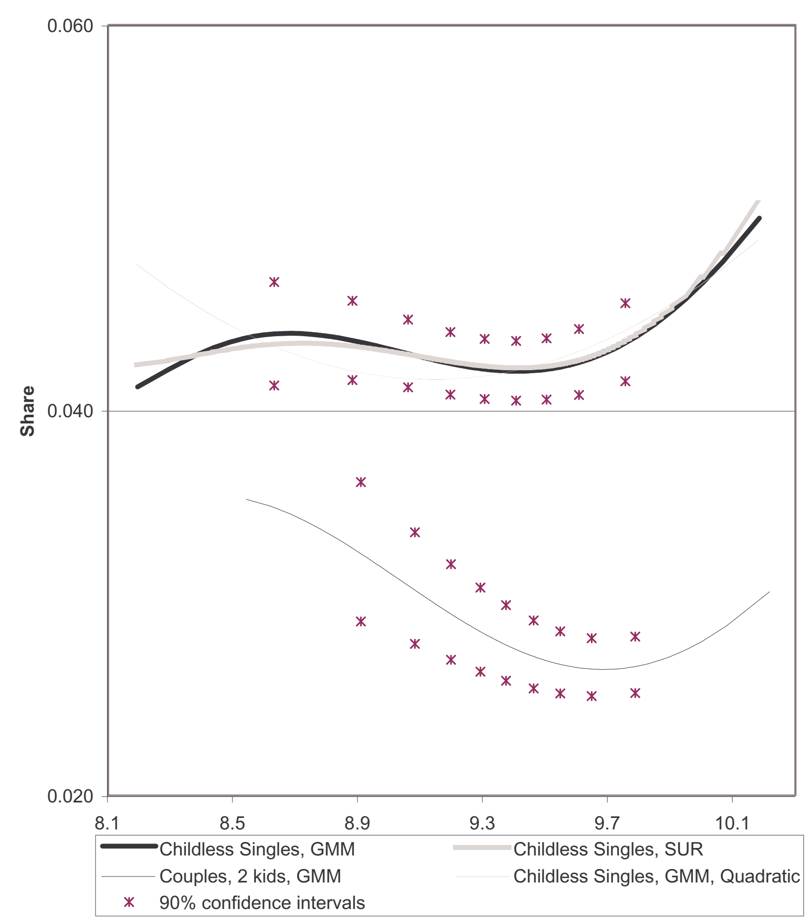


Figure 9: Cost-of-Living Experiment: 15% tax on rent

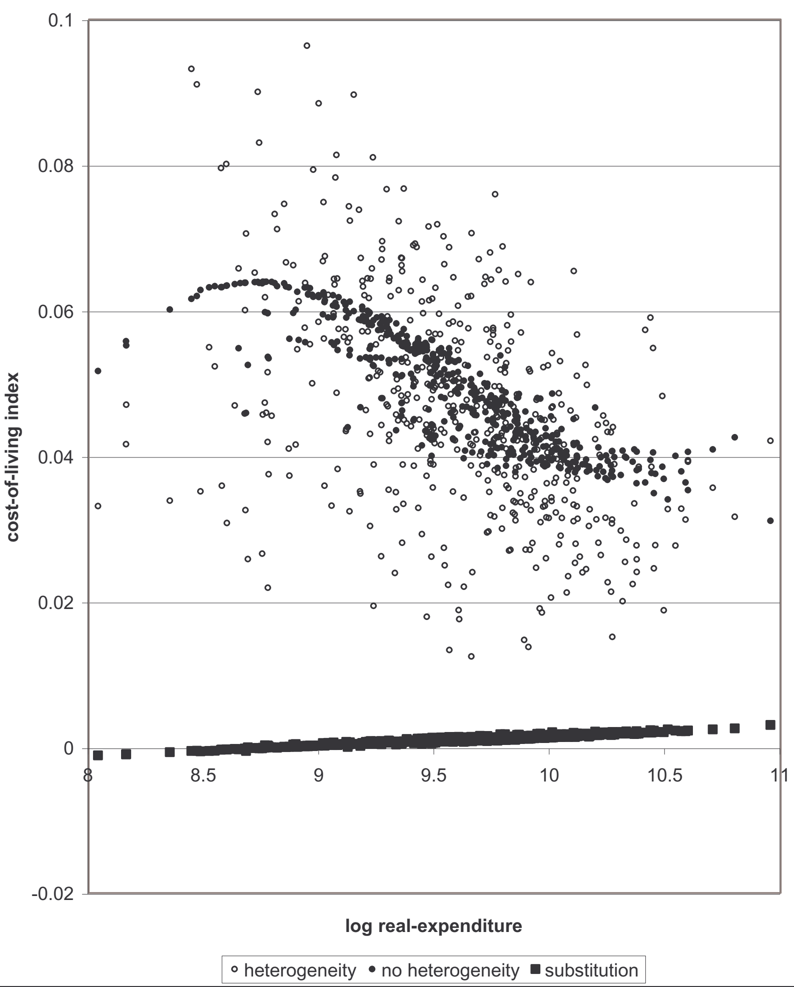


Table A1: Engel Curve Parameters, **b**

			GMM			SUR		
Parameter	Effect	Equation	Estimate	Standard Error	t-statistic	Estimate	Standard Error	t-statistic
b0_1	constant	foodin	0.463	0.308	1.505	0.375	0.176	2.136
b0_2	constant	foodout	0.116	0.119	0.968	0.009	0.006	1.417
b0_3	constant	rent	-0.320	0.309	-1.036	0.001	0.002	0.463
b0_4	constant	hhoper	0.142	0.188	0.754	0.012	0.006	2.199
b0_5	constant	hhfurneq	-0.025	0.102	-0.242	0.031	0.009	3.574
b0_6	constant	cloth	-0.209	0.124	-1.691	0.005	0.008	0.600
b0_7	constant	privtrans	0.788	0.149	5.292	-0.002	0.002	-1.298
b0_8	constant	pubtrans	-0.068	0.082	-0.834	0.012	0.009	1.415
b-1_ 1	y^{-1}	foodin	-0.040	0.063	-0.637	-0.027	0.032	-0.842
b-1_ 2	y^{-1}	foodout	-0.021	0.019	-1.076	-0.017	0.006	-2.711
b-1_ 3	y^{-1}	rent	0.118	0.065	1.818	0.002	0.012	0.187
b-1_ 4	y^{-1}	hhoper	-0.005	0.038	-0.125	0.018	0.006	2.773
b-1_ 5	y^{-1}	hhfurneq	0.006	0.015	0.403	-0.001	0.004	-0.362
b-1_ 6	y^{-1}	cloth	0.044	0.020	2.242	-0.437	0.156	-2.800
b-1_ 7	y^{-1}	privtrans	-0.116	0.031	-3.803	-0.005	0.005	-1.007
b-1_ 8	y^{-1}	pubtrans	0.023	0.014	1.687	-0.045	0.185	-0.242
b1_1	y	foodin	-0.136	0.555	-0.245	0.031	0.344	0.090
b1_2	y	foodout	-0.117	0.254	-0.459	-0.008	0.006	-1.389
b1_3	y	rent	1.303	0.555	2.347	0.001	0.020	0.064
b1_4	y	hhoper	-0.061	0.345	-0.178	-0.013	0.002	-5.341
b1_5	y	hhfurneq	0.099	0.228	0.434	-0.022	0.005	-4.558
b1_6	y	cloth	0.508	0.258	1.965	0.003	0.004	0.707
b1_7	y	privtrans	-1.616	0.270	-5.979	0.003	0.010	0.285
b1_8	y	pubtrans	0.177	0.167	1.066	0.046	0.094	0.485
b2_1	y^2	foodin	-0.116	0.480	-0.241	-0.264	0.320	-0.827
b2_2	y^2	foodout	0.100	0.247	0.406	-0.024	0.013	-1.941
b2_3	y^2	rent	-0.847	0.485	-1.747	-0.004	0.042	-0.099
b2_4	y^2	hhoper	-0.018	0.302	-0.061	-0.012	0.004	-3.113
b2_5	y^2	hhfurneq	-0.108	0.233	-0.461	0.002	0.009	0.199
b2_6	y^2	cloth	-0.467	0.248	-1.883	-1.321	0.353	-3.740
b2_7	y^2	privtrans	1.447	0.239	6.060	0.039	0.019	1.996
b2_8	y^2	pubtrans	-0.117	0.156	-0.754	0.000	0.000	6.196

b3_1	y^3	foodin	0.095	0.214	0.444	0.165	0.152	1.088
b3_2	y^3	foodout	-0.026	0.121	-0.214	0.001	0.000	3.200
b3_3	y^3	rent	0.221	0.220	1.007	0.000	0.000	0.317
b3_4	y^3	hhoper	0.033	0.136	0.241	0.011	0.003	4.281
b3_5	y^3	hhfurneq	0.056	0.120	0.468	0.001	0.006	0.229
b3_6	y^3	cloth	0.221	0.121	1.827	-0.015	0.010	-1.462
b3_7	y^3	privtrans	-0.580	0.110	-5.290	0.000	0.000	1.459
b3_8	y^3	pubtrans	0.031	0.074	0.415	-0.005	0.002	-2.792
b4_1	y^4	foodin	-0.026	0.047	-0.552	-0.043	0.036	-1.199
b4_2	y^4	foodout	0.002	0.029	0.084	0.039	0.005	7.104
b4_3	y^4	rent	-0.023	0.050	-0.464	-0.011	0.010	-1.049
b4_4	y^4	hhoper	-0.011	0.030	-0.354	0.175	0.254	0.690
b4_5	y^4	hhfurneq	-0.012	0.030	-0.388	-0.024	0.012	-1.994
b4_6	y^4	cloth	-0.046	0.029	-1.605	0.006	0.005	1.168
b4_7	y^4	privtrans	0.106	0.025	4.245	0.009	0.004	2.341
b4_8	y^4	pubtrans	-0.001	0.017	-0.072	0.009	0.003	3.095
b5_1	y^5	foodin	0.002	0.004	0.615	0.004	0.003	1.241
b5_2	y^5	foodout	0.000	0.003	0.006	-0.002	0.008	-0.288
b5_3	y^5	rent	0.000	0.004	0.102	-0.040	0.017	-2.346
b5_4	y^5	hhoper	0.001	0.003	0.444	-0.023	0.012	-1.983
b5_5	y^5	hhfurneq	0.001	0.003	0.330	0.000	0.000	-1.237
b5_6	y^5	cloth	0.004	0.003	1.310	1.148	0.328	3.498
b5_7	y^5	privtrans	-0.007	0.002	-3.308	-0.008	0.006	-1.213
b5_8	y^5	pubtrans	0.000	0.002	-0.222	0.001	0.002	0.766

Table A2: Demographic Parameters, **D**

			GMM			SUR		
				Standard			Standard	
Parameter	Effect	Equation	Estimate	Error	t-statistic	Estimate	Error	t-statistic
C1_couple	couple	foodin	-0.002	0.010	-0.222	0.002	0.008	0.205
C2_couple	couple	foodout	-0.026	0.006	-4.548	-0.011	0.009	-1.196
C3_couple	couple	rent	0.026	0.012	2.101	-0.016	0.026	-0.616
C4_couple	couple	hhoper	0.045	0.005	8.800	-0.001	0.000	-3.275
C5_couple	couple	hhfurneq	-0.016	0.008	-2.181	0.001	0.000	1.847
C6_couple	couple	cloth	-0.022	0.007	-3.303	0.018	0.011	1.571
C7_couple	couple	privtrans	-0.002	0.008	-0.285	0.021	0.019	1.076
C8_couple	couple	pubtrans	-0.001	0.005	-0.147	0.005	0.004	1.155
C1_headage	headage	foodin	0.002	0.000	7.440	0.002	0.000	9.137
C2_headage	headage	foodout	-0.001	0.000	-6.482	-0.042	0.025	-1.675
C3_headage	headage	rent	-0.001	0.000	-2.099	0.003	0.005	0.567
C4_headage	headage	hhoper	0.001	0.000	6.795	-0.010	0.007	-1.496
C5_headage	headage	hhfurneq	0.001	0.000	4.330	0.074	0.037	2.029
C6_headage	headage	cloth	0.000	0.000	-1.425	-0.008	0.009	-0.921
C7_headage	headage	privtrans	-0.001	0.000	-3.986	-0.008	0.005	-1.757
C8_headage	headage	pubtrans	0.000	0.000	-3.103	0.000	0.000	-0.305
C1_logsize	logsize	foodin	0.214	0.009	22.551	0.199	0.008	26.390
C2_logsize	logsize	foodout	-0.027	0.006	-4.845	-0.001	0.000	-5.326
C3_logsize	logsize	rent	-0.205	0.012	-17.232	-0.146	0.236	-0.620
C4_logsize	logsize	hhoper	0.014	0.005	2.765	0.001	0.005	0.136
C5_logsize	logsize	hhfurneq	0.051	0.008	6.183	-0.005	0.003	-1.401
C6_logsize	logsize	cloth	0.022	0.007	3.035	-0.001	0.000	-3.328
C7_logsize	logsize	privtrans	-0.077	0.007	-10.295	0.086	0.172	0.501
C8_logsize	logsize	pubtrans	-0.001	0.004	-0.178	-0.001	0.004	-0.211
C1_singlepar	singlepar	foodin	-0.007	0.013	-0.526	0.012	0.011	1.078
C2_singlepar	singlepar	foodout	-0.006	0.006	-0.945	0.001	0.005	0.105
C3_singlepar	singlepar	rent	0.037	0.018	2.056	0.077	0.112	0.685
C4_singlepar	singlepar	hhoper	-0.035	0.009	-3.897	-0.002	0.010	-0.257
C5_singlepar	singlepar	hhfurneq	-0.010	0.008	-1.159	-0.098	0.033	-2.942
C6_singlepar	singlepar	cloth	0.030	0.008	3.762	-0.066	0.008	-8.818
C7_singlepar	singlepar	privtrans	0.009	0.009	0.944	-0.065	0.082	-0.790
C8_singlepar	singlepar	pubtrans	-0.021	0.005	-4.210	0.000	0.006	-0.039

Table A3: Price Parameters, $\mathbf{A}_0$

Parameter	Effect	Equation	GMM Estimate	GMM Standard Error	GMM t-statistic	SUR Estimate	SUR Standard Error	SUR t-statistic
A11	pfoodin	foodin	-0.169	0.056	-3.027	0.007	0.003	2.504
A12	pfoodin	foodout	0.019	0.033	0.557	0.011	0.023	0.454
A13	pfoodin	rent	-0.118	0.022	-5.392	-0.039	0.031	-1.247
A14	pfoodin	hhoper	0.100	0.037	2.691	0.052	0.012	4.168
A15	pfoodin	hhfurneq	0.096	0.031	3.091	-0.001	0.020	-0.069
A16	pfoodin	cloth	0.062	0.026	2.373	0.023	0.016	1.415
A17	pfoodin	privtrans	0.067	0.020	3.356	0.020	0.016	1.247
A18	pfoodin	pubtrans	-0.108	0.016	-6.617	-0.026	0.007	-3.999
A22	pfoodout	foodout	0.000	0.030	0.000	0.006	0.007	0.916
A23	pfoodout	rent	0.126	0.015	8.579	-0.013	0.023	-0.593
A24	pfoodout	hhoper	-0.058	0.025	-2.302	0.125	0.109	1.146
A25	pfoodout	hhfurneq	0.010	0.023	0.410	-0.009	0.026	-0.347
A26	pfoodout	cloth	-0.038	0.019	-2.019	0.056	0.009	6.237
A27	pfoodout	privtrans	-0.059	0.014	-4.228	-0.008	0.007	-1.223
A28	pfoodout	pubtrans	0.035	0.012	2.850	0.011	0.003	4.026
A33	prent	rent	-0.202	0.031	-6.625	0.054	0.041	1.337
A34	prent	hhoper	-0.006	0.012	-0.512	-0.052	0.213	-0.242
A35	prent	hhfurneq	0.037	0.014	2.634	0.002	0.012	0.135
A36	prent	cloth	0.075	0.014	5.436	-0.332	0.231	-1.434
A37	prent	privtrans	0.036	0.015	2.298	-0.009	0.012	-0.703
A38	prent	pubtrans	0.046	0.010	4.723	-0.006	0.014	-0.466
A44	phhoper	hhoper	-0.010	0.041	-0.242	-0.051	0.013	-4.005
A45	phhoper	hhfurneq	-0.049	0.029	-1.674	-0.055	0.013	-4.245
A46	phhoper	cloth	-0.002	0.021	-0.084	1.400	0.454	3.088
A47	phhoper	privtrans	-0.037	0.013	-2.822	0.007	0.013	0.556
A48	phhoper	pubtrans	0.056	0.013	4.420	-0.039	0.117	-0.337
A55	phhfurneq	hhfurneq	-0.015	0.038	-0.382	-0.014	0.007	-2.132
A56	phhfurneq	cloth	-0.093	0.025	-3.693	-0.026	0.008	-3.099
A57	phhfurneq	privtrans	0.017	0.013	1.325	-0.009	0.009	-0.917
A58	phhfurneq	pubtrans	-0.028	0.012	-2.255	0.035	0.021	1.691
A66	pcloth	cloth	0.038	0.025	1.533	0.016	0.013	1.271
A67	pcloth	privtrans	0.002	0.012	0.197	0.000	0.007	-0.061
A68	pcloth	pubtrans	-0.018	0.010	-1.852	-0.053	0.014	-3.697
A77	pprivtrans	privtrans	-0.030	0.019	-1.621	0.000	0.004	0.066
A78	pprivtrans	pubtrans	0.010	0.009	1.131	-0.025	0.007	-3.429
A88	ppubtrans	pubtrans	0.001	0.009	0.062	0.007	0.006	1.212

Table A4: Interactions

	Parameter	Effect	(*Effect)	Equation	GMM Estimate	GMM Standard Error	GMM t-stat	SUR Estimate	SUR Standard Error	SUR t-stat
Demo graphics* Prices (**A**)	A11_couple	couple	pfoodin	foodin	-0.026	0.022	-1.139	0.004	0.004	1.042
	A12_couple	couple	pfoodin	foodout	-0.012	0.016	-0.761	0.003	0.002	1.237
	A13_couple	couple	pfoodin	rent	0.017	0.008	2.012	0.003	0.015	0.179
	A14_couple	couple	pfoodin	hhoper	0.035	0.015	2.421	0.004	0.006	0.687
	A15_couple	couple	pfoodin	hhfurneq	-0.013	0.014	-0.877	0.001	0.002	0.254
	A16_couple	couple	pfoodin	cloth	0.022	0.011	1.927	-0.009	0.008	-1.150
	A17_couple	couple	pfoodin	privtrans	-0.015	0.009	-1.772	-0.005	0.023	-0.232
	A18_couple	couple	pfoodin	pubtrans	-0.005	0.007	-0.732	-0.006	0.009	-0.685
	A22_couple	couple	pfoodout	foodout	0.021	0.016	1.322	-0.009	0.007	-1.403
	A23_couple	couple	pfoodout	rent	-0.005	0.006	-0.773	-0.001	0.011	-0.076
	A24_couple	couple	pfoodout	hhoper	-0.015	0.010	-1.425	0.045	0.005	8.471
	A25_couple	couple	pfoodout	hhfurneq	0.010	0.012	0.885	0.026	0.012	2.100
	A26_couple	couple	pfoodout	cloth	-0.011	0.009	-1.260	-0.003	0.004	-0.795
	A27_couple	couple	pfoodout	privtrans	0.003	0.007	0.445	-0.020	0.010	-2.015
	A28_couple	couple	pfoodout	pubtrans	0.006	0.006	1.036	-0.007	0.004	-1.803
	A33_couple	couple	prent	rent	0.014	0.011	1.334	-0.029	0.020	-1.486
	A34_couple	couple	prent	hhoper	0.000	0.005	0.086	-0.024	0.002	-10.85
	A35_couple	couple	prent	hhfurneq	-0.017	0.007	-2.524	0.002	0.006	0.378
	A36_couple	couple	prent	cloth	0.001	0.006	0.245	0.017	0.011	1.601
	A37_couple	couple	prent	privtrans	-0.013	0.007	-1.836	-0.020	0.017	-1.167
	A38_couple	couple	prent	pubtrans	0.002	0.004	0.437	0.039	0.013	2.948
	A44_couple	couple	phhoper	hhoper	-0.030	0.017	-1.790	0.010	0.006	1.551
	A45_couple	couple	phhoper	hhfurneq	-0.004	0.013	-0.294	-0.015	0.006	-2.455
	A46_couple	couple	phhoper	cloth	0.002	0.009	0.248	-0.004	0.005	-0.836
	A47_couple	couple	phhoper	privtrans	0.011	0.005	1.936	-0.019	0.018	-1.088
	A48_couple	couple	phhoper	pubtrans	0.006	0.005	1.106	0.115	0.246	0.467
	A55_couple	couple	phhfurneq	hhfurneq	-0.003	0.019	-0.164	0.013	0.007	1.968
	A56_couple	couple	phhfurneq	cloth	0.005	0.013	0.434	-0.025	0.007	-3.727
	A57_couple	couple	phhfurneq	privtrans	0.014	0.006	2.118	0.007	0.013	0.532
	A58_couple	couple	phhfurneq	pubtrans	-0.008	0.006	-1.315	0.001	0.001	0.992
	A66_couple	couple	pcloth	cloth	-0.016	0.012	-1.410	-0.003	0.006	-0.514
	A67_couple	couple	pcloth	privtrans	0.005	0.006	0.752	-0.005	0.010	-0.471
	A68_couple	couple	pcloth	pubtrans	0.000	0.005	0.001	-0.001	0.001	-1.818
	A77_couple	couple	pprivtrans	privtrans	-0.009	0.010	-0.882	-0.025	0.047	-0.532
	A78_couple	couple	pprivtrans	pubtrans	0.000	0.005	-0.056	0.000	0.000	-0.607
	A88_couple	couple	ppubtrans	pubtrans	0.001	0.004	0.183	0.011	0.009	1.298

A11_headage	headage	pfoodin	foodin	0.000	0.001	0.393	-0.056	0.127	-0.443
A12_headage	headage	pfoodin	foodout	-0.001	0.001	-0.947	-0.206	0.231	-0.892
A13_headage	headage	pfoodin	rent	0.000	0.000	-0.106	0.000	0.001	0.863
A14_headage	headage	pfoodin	hhoper	0.000	0.001	0.829	0.000	0.000	-1.391
A15_headage	headage	pfoodin	hhfurneq	0.000	0.001	-0.206	-0.004	0.198	-0.019
A16_headage	headage	pfoodin	cloth	0.000	0.000	-0.293	0.000	0.000	-1.308
A17_headage	headage	pfoodin	privtrans	0.000	0.000	-0.318	0.100	0.043	2.352
A18_headage	headage	pfoodin	pubtrans	0.000	0.000	-0.180	0.039	0.012	3.212
A22_headage	headage	pfoodout	foodout	0.001	0.001	0.934	0.174	0.249	0.698
A23_headage	headage	pfoodout	rent	0.000	0.000	-0.158	0.000	0.000	-0.405
A24_headage	headage	pfoodout	hhoper	0.000	0.000	-0.729	0.001	0.000	8.137
A25_headage	headage	pfoodout	hhfurneq	0.000	0.000	0.421	0.000	0.000	0.390
A26_headage	headage	pfoodout	cloth	0.000	0.000	1.002	0.001	0.000	3.133
A27_headage	headage	pfoodout	privtrans	0.000	0.000	0.912	-0.006	0.032	-0.181
A28_headage	headage	pfoodout	pubtrans	0.000	0.000	-1.669	0.114	0.135	0.841
A33_headage	headage	prent	rent	0.000	0.000	0.642	0.000	0.001	-0.381
A34_headage	headage	prent	hhoper	0.000	0.000	1.120	-0.001	0.000	-12.22
A35_headage	headage	prent	hhfurneq	0.000	0.000	1.489	0.000	0.000	1.893
A36_headage	headage	prent	cloth	-0.001	0.000	-3.427	0.000	0.000	-0.790
A37_headage	headage	prent	privtrans	0.000	0.000	-1.640	0.160	0.014	11.310
A38_headage	headage	prent	pubtrans	0.000	0.000	2.339	-0.121	0.265	-0.454
A44_headage	headage	phhoper	hhoper	-0.001	0.001	-0.790	0.000	0.000	1.267
A45_headage	headage	phhoper	hhfurneq	0.001	0.000	1.287	0.000	0.000	0.957
A46_headage	headage	phhoper	cloth	0.000	0.000	-0.490	0.000	0.000	3.230
A47_headage	headage	phhoper	privtrans	0.000	0.000	1.108	-0.055	0.025	-2.200
A48_headage	headage	phhoper	pubtrans	0.000	0.000	-1.033	-0.014	0.010	-1.389
A55_headage	headage	phhfurneq	hhfurneq	0.000	0.001	-0.643	-0.015	0.021	-0.697
A56_headage	headage	phhfurneq	cloth	0.000	0.000	0.056	0.049	0.013	3.630
A57_headage	headage	phhfurneq	privtrans	-0.001	0.000	-5.184	0.006	0.026	0.237
A58_headage	headage	phhfurneq	pubtrans	0.001	0.000	3.576	0.007	0.028	0.255
A66_headage	headage	pcloth	cloth	0.001	0.000	1.914	-0.001	0.000	-3.847
A67_headage	headage	pcloth	privtrans	0.000	0.000	0.906	-0.022	0.019	-1.138
A68_headage	headage	pcloth	pubtrans	0.000	0.000	-2.030	-0.018	0.037	-0.499
A77_headage	headage	pprivtrans	privtrans	0.001	0.000	2.127	-0.002	0.005	-0.366
A78_headage	headage	pprivtrans	pubtrans	0.000	0.000	0.873	0.023	0.024	0.959
A88_headage	headage	ppubtrans	pubtrans	0.000	0.000	0.182	-0.062	0.014	-4.498
A11_logsize	logsize	pfoodin	foodin	-0.002	0.019	-0.116	-0.001	0.000	-6.567
A12_logsize	logsize	pfoodin	foodout	-0.036	0.013	-2.696	0.120	0.110	1.089
A13_logsize	logsize	pfoodin	rent	0.007	0.007	0.909	-0.014	0.012	-1.151
A14_logsize	logsize	pfoodin	hhoper	0.021	0.014	1.553	0.006	0.005	1.155
A15_logsize	logsize	pfoodin	hhfurneq	0.025	0.012	2.006	0.016	0.094	0.170
A16_logsize	logsize	pfoodin	cloth	-0.003	0.010	-0.280	0.022	0.006	3.467
A17_logsize	logsize	pfoodin	privtrans	0.013	0.007	1.762	0.001	0.001	1.626
A18_logsize	logsize	pfoodin	pubtrans	-0.030	0.006	-5.025	-0.002	0.000	-8.092
A22_logsize	logsize	pfoodout	foodout	0.002	0.013	0.189	-0.020	0.022	-0.933
A23_logsize	logsize	pfoodout	rent	0.002	0.006	0.384	-0.007	0.009	-0.820
A24_logsize	logsize	pfoodout	hhoper	-0.013	0.010	-1.386	0.008	0.005	1.685
A25_logsize	logsize	pfoodout	hhfurneq	0.016	0.011	1.475	0.015	0.010	1.465
A26_logsize	logsize	pfoodout	cloth	0.017	0.008	2.195	-0.009	0.004	-2.369
A27_logsize	logsize	pfoodout	privtrans	-0.021	0.006	-3.784	0.000	0.000	0.826
A28_logsize	logsize	pfoodout	pubtrans	0.042	0.005	8.452	0.000	0.000	4.912
A33_logsize	logsize	prent	rent	-0.025	0.010	-2.614	0.020	0.016	1.232
A34_logsize	logsize	prent	hhoper	0.005	0.004	1.130	-0.001	0.002	-0.275
A35_logsize	logsize	prent	hhfurneq	-0.007	0.006	-1.170	0.000	0.005	0.028
A36_logsize	logsize	prent	cloth	-0.002	0.005	-0.418	-0.168	0.010	-17.44
A37_logsize	logsize	prent	privtrans	0.023	0.006	3.702	0.000	0.000	-0.424
A38_logsize	logsize	prent	pubtrans	-0.008	0.004	-2.337	-0.023	0.017	-1.329
A44_logsize	logsize	phhoper	hhoper	0.007	0.015	0.511	-0.012	0.005	-2.304
A45_logsize	logsize	phhoper	hhfurneq	-0.024	0.012	-2.060	-0.021	0.005	-3.939
A46_logsize	logsize	phhoper	cloth	-0.007	0.008	-0.888	0.062	0.004	14.086
A47_logsize	logsize	phhoper	privtrans	-0.007	0.005	-1.463	0.001	0.000	1.227
A48_logsize	logsize	phhoper	pubtrans	-0.001	0.005	-0.233	0.041	0.007	5.613
A55_logsize	logsize	phhfurneq	hhfurneq	0.015	0.016	0.917	0.010	0.016	0.617
A56_logsize	logsize	phhfurneq	cloth	-0.012	0.011	-1.116	0.000	0.006	-0.015
A57_logsize	logsize	phhfurneq	privtrans	0.005	0.005	0.856	0.000	0.000	-0.128
A58_logsize	logsize	phhfurneq	pubtrans	-0.015	0.005	-3.032	-0.001	0.003	-0.201
A66_logsize	logsize	pcloth	cloth	0.029	0.010	3.000	-0.003	0.005	-0.626
A67_logsize	logsize	pcloth	privtrans	-0.021	0.005	-4.167	0.000	0.000	1.011
A68_logsize	logsize	pcloth	pubtrans	-0.002	0.004	-0.625	-0.025	0.025	-1.001
A77_logsize	logsize	pprivtrans	privtrans	-0.003	0.008	-0.383	-0.933	0.421	-2.214
A78_logsize	logsize	pprivtrans	pubtrans	0.018	0.004	4.725	-0.012	0.017	-0.696
A88_logsize	logsize	ppubtrans	pubtrans	-0.007	0.003	-2.051	0.000	0.000	-1.865
A11_singlepar	singlepar	pfoodin	foodin	-0.008	0.036	-0.234	-0.024	0.002	-9.841
A12_singlepar	singlepar	pfoodin	foodout	0.000	0.019	-0.010	-0.030	0.026	-1.172
A13_singlepar	singlepar	pfoodin	rent	-0.026	0.013	-1.964	-0.043	0.021	-1.986
A14_singlepar	singlepar	pfoodin	hhoper	0.022	0.027	0.811	-0.003	0.009	-0.365
A15_singlepar	singlepar	pfoodin	hhfurneq	0.018	0.021	0.846	-0.005	0.022	-0.225
A16_singlepar	singlepar	pfoodin	cloth	0.002	0.017	0.124	0.017	0.011	1.513
A17_singlepar	singlepar	pfoodin	privtrans	-0.006	0.012	-0.474	0.014	0.013	1.086
A18_singlepar	singlepar	pfoodin	pubtrans	0.012	0.009	1.274	-0.028	0.006	-4.823
A22_singlepar	singlepar	pfoodout	foodout	0.009	0.016	0.601	0.045	0.013	3.398
A23_singlepar	singlepar	pfoodout	rent	-0.031	0.007	-4.588	0.034	0.016	2.192
A24_singlepar	singlepar	pfoodout	hhoper	0.002	0.015	0.148	-0.030	0.007	-4.196
A25_singlepar	singlepar	pfoodout	hhfurneq	-0.005	0.014	-0.384	0.013	0.018	0.716
A26_singlepar	singlepar	pfoodout	cloth	0.010	0.010	1.056	-0.012	0.006	-1.874
A27_singlepar	singlepar	pfoodout	privtrans	-0.009	0.006	-1.461	0.022	0.006	3.940
A28_singlepar	singlepar	pfoodout	pubtrans	0.008	0.006	1.395	-0.008	0.003	-2.963
A33_singlepar	singlepar	prent	rent	0.052	0.019	2.752	0.003	0.028	0.114
A34_singlepar	singlepar	prent	hhoper	-0.025	0.011	-2.204	0.029	0.003	8.816

	A35_singlepar	singlepar	prent	hhfurneq	0.017	0.009	1.868	-0.025	0.009	-2.902
	A36_singlepar	singlepar	prent	cloth	0.015	0.008	1.832	0.002	0.014	0.115
	A37_singlepar	singlepar	prent	privtrans	0.020	0.010	1.978	-0.010	0.010	-0.988
	A38_singlepar	singlepar	prent	pubtrans	-0.018	0.005	-3.506	-0.063	0.007	-8.513
	A44_singlepar	singlepar	phhoper	hhoper	-0.020	0.032	-0.610	0.009	0.009	1.045
	A45_singlepar	singlepar	phhoper	hhfurneq	-0.009	0.021	-0.420	0.016	0.009	1.752
	A46_singlepar	singlepar	phhoper	cloth	0.007	0.016	0.412	0.016	0.007	2.348
	A47_singlepar	singlepar	phhoper	privtrans	0.010	0.012	0.837	0.021	0.011	1.928
	A48_singlepar	singlepar	phhoper	pubtrans	-0.002	0.009	-0.182	-0.039	0.007	-5.992
	A55_singlepar	singlepar	phhfurneq	hhfurneq	0.021	0.024	0.882	0.009	0.012	0.727
	A56_singlepar	singlepar	phhfurneq	cloth	-0.027	0.015	-1.747	0.013	0.007	1.844
	A57_singlepar	singlepar	phhfurneq	privtrans	-0.027	0.008	-3.272	0.005	0.008	0.684
	A58_singlepar	singlepar	phhfurneq	pubtrans	0.027	0.008	3.517	-0.016	0.025	-0.661
	A66_singlepar	singlepar	pcloth	cloth	-0.009	0.014	-0.621	0.019	0.009	2.095
	A67_singlepar	singlepar	pcloth	privtrans	0.003	0.007	0.403	-0.025	0.006	-4.431
	A68_singlepar	singlepar	pcloth	pubtrans	-0.006	0.006	-1.092	-0.037	0.051	-0.732
	A77_singlepar	singlepar	pprivtrans	privtrans	0.027	0.012	2.297	0.247	0.200	1.235
	A78_singlepar	singlepar	pprivtrans	pubtrans	-0.014	0.005	-2.889	-0.024	0.034	-0.713
	A88_singlepar	singlepar	ppubtrans	pubtrans	0.001	0.005	0.303	0.042	0.005	8.476
Prices * Real Expend	B_11	y	pfoodin	foodin	0.064	0.028	2.244	-0.003	0.027	-0.101
	B_12	y	pfoodin	foodout	0.022	0.018	1.243	0.043	0.018	2.358
	B_13	y	pfoodin	rent	0.020	0.011	1.862	0.056	0.010	5.806
	B_14	y	pfoodin	hhoper	-0.057	0.020	-2.885	-0.031	0.019	-1.643
	B_15	y	pfoodin	hhfurneq	-0.028	0.017	-1.680	-0.045	0.017	-2.739
	B_16	y	pfoodin	cloth	-0.021	0.014	-1.560	-0.013	0.013	-1.002
	B_17	y	pfoodin	privtrans	-0.021	0.010	-2.046	-0.030	0.010	-3.177
	B_18	y	pfoodin	pubtrans	0.037	0.009	4.261	0.049	0.008	6.048
	B_22	y	pfoodout	foodout	-0.021	0.016	-1.319	-0.010	0.013	-0.766
	B_23	y	pfoodout	rent	-0.043	0.008	-5.612	-0.174	0.018	-9.445
	B_24	y	pfoodout	hhoper	0.040	0.014	2.929	0.006	0.015	0.400
	B_25	y	pfoodout	hhfurneq	-0.010	0.014	-0.709	0.060	0.036	1.659
	B_26	y	pfoodout	cloth	-0.008	0.010	-0.719	-0.011	0.015	-0.715
	B_27	y	pfoodout	privtrans	0.037	0.007	4.931	0.025	0.012	2.158
	B_28	y	pfoodout	pubtrans	-0.036	0.007	-5.393	0.080	0.018	4.418
	B_33	y	prent	rent	0.151	0.015	10.067	0.013	0.009	1.437
	B_34	y	prent	hhoper	0.003	0.007	0.488	0.021	0.025	0.831
	B_35	y	prent	hhfurneq	-0.035	0.008	-4.383	0.000	0.001	0.037
	B_36	y	prent	cloth	-0.052	0.008	-6.975	0.120	0.031	3.830
	B_37	y	prent	privtrans	-0.035	0.008	-4.436	0.032	0.024	1.318
	B_38	y	prent	pubtrans	0.003	0.005	0.510	0.000	0.000	-0.270
	B_44	y	phhoper	hhoper	0.020	0.022	0.944	0.000	0.001	0.231
	B_45	y	phhoper	hhfurneq	0.013	0.016	0.822	0.026	0.013	2.078
	B_46	y	phhoper	cloth	0.005	0.012	0.460	0.000	0.000	-0.300
	B_47	y	phhoper	privtrans	0.006	0.007	0.816	0.000	0.000	0.200
	B_48	y	phhoper	pubtrans	-0.015	0.007	-2.231	-0.039	0.006	-6.225

	B_55	y	phhfurneq	hhfurneq	-0.013	0.022	-0.604	0.027	0.018	1.536
	B_56	y	phhfurneq	cloth	0.068	0.015	4.696	0.006	0.010	0.655
	B_57	y	phhfurneq	privtrans	-0.010	0.007	-1.365	0.021	0.007	2.854
	B_58	y	phhfurneq	pubtrans	0.009	0.007	1.284	0.005	0.011	0.517
	B_66	y	pcloth	cloth	-0.003	0.014	-0.241	0.031	0.022	1.381
	B_67	y	pcloth	privtrans	0.003	0.007	0.375	-0.009	0.013	-0.672
	B_68	y	pcloth	pubtrans	0.010	0.005	1.978	0.001	0.007	0.134
	B_77	y	pprivtrans	privtrans	0.026	0.010	2.621	-0.016	0.017	-0.944
	B_78	y	pprivtrans	pubtrans	-0.006	0.005	-1.284	-0.129	0.015	-8.482
	B_88	y	ppubtrans	pubtrans	0.001	0.005	0.248	-0.015	0.009	-1.667
Demo graphics* Real Expend	D1_couple	y	couple	foodin	-0.004	0.004	-0.949	-0.006	0.004	-1.594
	D2_couple	y	couple	foodout	0.012	0.003	4.467	0.029	0.009	3.278
	D3_couple	y	couple	rent	-0.007	0.005	-1.394	0.045	0.029	1.577
	D4_couple	y	couple	hhoper	-0.024	0.002	-10.92	0.000	0.000	0.788
	D5_couple	y	couple	hhfurneq	0.006	0.003	1.718	0.000	0.000	1.285
	D6_couple	y	couple	cloth	0.011	0.003	3.481	-0.023	0.005	-4.276
	D7_couple	y	couple	privtrans	0.005	0.004	1.530	0.008	0.014	0.590
	D8_couple	y	couple	pubtrans	0.001	0.002	0.576	-0.002	0.004	-0.501
	D1_headage	y	headage	foodin	0.000	0.000	-1.974	0.000	0.000	-2.706
	D2_headage	y	headage	foodout	0.000	0.000	3.309	0.012	0.013	0.970
	D3_headage	y	headage	rent	0.001	0.000	3.823	0.041	0.024	1.739
	D4_headage	y	headage	hhoper	-0.001	0.000	-10.25	0.027	0.008	3.267
	D5_headage	y	headage	hhfurneq	-0.001	0.000	-6.919	0.004	0.015	0.261
	D6_headage	y	headage	cloth	0.000	0.000	-1.118	0.661	0.180	3.668
	D7_headage	y	headage	privtrans	0.000	0.000	2.830	0.002	0.017	0.137
	D8_headage	y	headage	pubtrans	0.000	0.000	4.612	-0.001	0.000	-4.728
	D1_logsize	y	logsize	foodin	-0.044	0.004	-10.66	-0.039	0.003	-11.79
	D2_logsize	y	logsize	foodout	-0.010	0.003	-3.719	0.001	0.000	2.343
	D3_logsize	y	logsize	rent	0.077	0.005	14.410	-0.001	0.001	-0.663
	D4_logsize	y	logsize	hhoper	-0.003	0.002	-1.172	-0.024	0.006	-3.799
	D5_logsize	y	logsize	hhfurneq	-0.029	0.004	-7.294	-0.010	0.011	-0.902
	D6_logsize	y	logsize	cloth	-0.018	0.003	-5.327	0.000	0.000	1.986
	D7_logsize	y	logsize	privtrans	0.034	0.003	10.097	0.000	0.000	1.070
	D8_logsize	y	logsize	pubtrans	-0.005	0.002	-2.676	-0.001	0.004	-0.341
	D1_singlepar	y	singlepar	foodin	-0.004	0.006	-0.661	-0.011	0.005	-2.170
	D2_singlepar	y	singlepar	foodout	-0.009	0.003	-3.117	-0.013	0.005	-2.693
	D3_singlepar	y	singlepar	rent	0.005	0.008	0.613	0.018	0.016	1.112
	D4_singlepar	y	singlepar	hhoper	0.032	0.005	7.108	-0.082	0.129	-0.633
	D5_singlepar	y	singlepar	hhfurneq	0.005	0.004	1.259	0.079	0.022	3.580
	D6_singlepar	y	singlepar	cloth	-0.015	0.004	-3.738	0.029	0.003	8.569
	D7_singlepar	y	singlepar	privtrans	-0.019	0.004	-4.264	0.016	0.008	1.981
	D8_singlepar	y	singlepar	pubtrans	0.007	0.002	2.808	-0.024	0.006	-4.103