

Estimation of Collective Household Models With Engel Curves

Arthur Lewbel and Krishna Pendakur
Boston College and Simon Fraser University

May 2008

Abstract

The structural consumer demand methods used to estimate the parameters of collective household models are typically either very restrictive and easy to implement or very general and difficult to estimate. In this paper, we provide a middle ground. We adapt the very general framework of Browning, Chiappori and Lewbel (2007) by adding a simple restriction that recasts the empirical model from a highly nonlinear demand system with price variation to a slightly nonlinear Engel curve system. Our restriction has an interpretation in terms of the behaviour of household scale economies and is testable. Our method identifies the levels of (not just changes in) household resource shares, and a variant of equivalence scales called indifference scales. We apply our methodology to Canadian expenditure data.

JEL codes: D12, D11, C30, I31, J12. Keywords: Consumer Demand, Collective Model, Sharing rule, Household Bargaining, Bargaining Power, Indifference Scales, Adult Equivalence Scales, Demand Systems, Barten Scales, Nonparametric Identification.

Corresponding Author: Arthur Lewbel, Department of Economics, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA, 02467, USA. (617)-552-3678, lewbel@bc.edu, <http://www2.bc.edu/~lewbel/>

1 Introduction

Dating back at least to Becker (1965, 1981), ‘collective household’ models are those in which the household is characterised as a collection of individuals, each of whom has a well defined objective function, and who interact to generate household level decisions. Here we focus on methods by which household level consumption data are used to recover information about individual household members. Examples include Chiappori (1988, 1992), Bourguignon and Chiappori (1994), Browning, Bourguignon, Chiappori, and Lechene (1994), Browning and Chiappori (1998), Vermeulen (2002), Browning, Chiappori and Lewbel (2004), Lise and Seitz (2004), and Cherchye, L., B. De Rock, and F. Vermeulen (2008). These models

are very useful because typical micro-data sources only have information on household level choices, but the objects of interest are based on the preferences of, and constraints faced by, the individuals who together make up the household.

Unfortunately, the structural models used to estimate these objects are either extremely restrictive and easy to estimate, or are very general but difficult to estimate. For example, in Chiappori (1994), households consume a purely public and purely private good and nothing else, yielding a model where the response of household member resource shares to changes in variables such as wages can be easily estimated. In contrast, Browning, Chiappori and Lewbel (2004), hereafter BCL, provide a model in which households consume a vector of goods ranging from purely private to very shareable, and show how to recover via demand system estimation both the resource shares and indifference scales (which are a variant of equivalence scales analogous to cost-of-living indices) for each household member. Unfortunately, this model is highly nonlinear in prices, expenditures and other characteristics, and is consequently difficult to estimate, both numerically and in terms of data requirements.

In this paper, we provide a middle ground. Our model is only a little more restrictive than BCL, but is easy to estimate. Specifically, we propose a version of BCL in which all the objects of interest can be obtained from estimates of Engel curves that are nearly linear in parameters. Basically, we offer a way to obtain identification without observing price variation, so that in our model the demand system reduces to a system of Engel curves. In this model, the nonlinearity is encompassed by a single parameter, and we can still recover the resource shares and indifference scales.

In the version of the BCL model we consider, the quantity of a household's demand for any consumption good is based on summing the demands of each household member for that good, and scaling the result by a Barten (1964) type scale. The Barten scale reflects the degree to which the good is shared among household members, that is, the degree of publicness or privateness of the good within the household. Each household member is allocated a *resource share*, that is, a share of the total resources (total expenditures) the household has to spend on consumption goods. Each household member determines their own demand for each consumption good by maximizing their own utility function, subject to a total budget constraint equal to the member's resource share of the total household budget, and facing a vector of Lindahl (1919) type shadow prices for goods. These shadow prices differ from market prices (by the Barten scales) because of economies of scale to consumption. In particular, shadow prices will be lower than market prices for goods that are shared or consumed jointly. Given each member's budget constraint within the household, BCL use ordinary consumer surplus methods to obtain *indifference scales* defined as the fraction of household expenditure that puts an individual living alone on the same indifference curve that she would attain living in the household. Resource shares and indifference scales are two objects of interest for each individual in a household, and BCL show that both these objects are identifiable from data on household budgets and market prices. We extend their results by providing identification without observing market prices in data without market price variation.

Unlike most empirical models in the collective household literature, BCL identify and estimate the level of resource shares, not just how those shares vary with so-called distribution factors. They do this by combining data on people living alone with data on couples. Lise and Seitz (2004) also identify the level of resource shares (though not indifference scales) by combining singles and couples data. They too require price data for identification, which they obtain by defining demand systems in terms of just two goods, consumption and leisure, with wage differences providing the required relative price variation. Another application of BCL's model requiring price variation is Cherchye, De Rock, and Vermeulen (2008).

Like BCL, our model does not consider marriage markets or matching. In particular, we do not model the connection between resource shares and the application of bargaining power or threat points in decisions to either marry or divorce, and so do not analyze the possibly endogenous process that determines who is single and who is married in our data. Our results, e.g., our estimates of the fraction of household

resources that are controlled by wives with varying amounts of income, are therefore conditional upon the outcome of marriage markets, and we do not explicitly model the constraints that marriage markets might impose on these resource shares. One reason for this is that an analysis of those constraints would require more explicit modeling of the household bargaining process than our methodology otherwise requires.

Our model obtains identification without price variation. Specifically, we show that combining BCL's general model with Barten (1964) scales for sharing goods and an Independence of Base assumption as in Blundell, Duncan and Pendakur (1998) yields the simple model

$$\mathbf{w}(x) = \mathbf{h} + \sum_j \eta_j \mathbf{w}_j (x - \ln I_j)$$

where $\mathbf{w}(x)$ is a household's vector of Engel curve budget shares given log total expenditures x , \mathbf{h} is a vector of constants, $\mathbf{w}_j(x)$ is the Engel curve budget shares of household member j , and the parameters of interest, η_j and I_j , are member j 's resource share and indifference scale, respectively. Estimation only requires parameterizing this model, adding demographic and error terms, and combining data on singles living alone to estimate $\mathbf{w}_j(x)$ along with collective household data to estimate $\mathbf{w}(x)$.

2 The Model

We begin by summarizing the BCL model of household demand equations. In general, we use lower-case to denote logged quantities, bold to denote vectors, superscripts to index goods and subscripts to index people and households. Let j denote individuals $j = 1, \dots, J$ and let $\mathbf{p} = [p^1, \dots, p^K]'$ be the K -vector of logged market-prices. Let x denote logged total expenditure, subscripted for households or individuals. Let $w_j^k(\mathbf{p}, x)$ denote person j 's budget share demand function for good k , that is, if person j were living alone he/she would spend the fraction $w_j^k(\mathbf{p}, x)$ of (unlogged) total expenditures e^x on the good k , for $k = 1, \dots, K$.

Assume that the household has economies of scale to consumption (that is, sharing and jointness or consumption) of a Barten (1964) type. Specifically, there exists a K vector of constants $\alpha = [\alpha^1, \dots, \alpha^K]'$, called log barten scales, such that the total log quantity of a good k that is consumed by the members of the household equals the log quantity of the good purchased by the household minus α^k . For example, suppose that a married couple ride together in a car (sharing the consumption of gasoline) half the time the car is in use. Then the total consumption of gasoline (as measured by summing the private equivalent consumption of each household member) is 3/2 times the purchased quantity of gasoline. Equivalently, if there had been no sharing of auto usage, so every member always drove alone, then the household would have had to purchase 50% more gasoline to have each member travel the same distance as before. In this example, we would have $\alpha^k = \ln(2/3)$ for k being gasoline. Thus α^k can be interpreted as the degree of "publicness" of good k within the household. A purely private good k would have $\alpha^k = \ln(1) = 0$, while a good that is shared has $\alpha^k < 0$, and the greater is the degree to which it is shared, the larger in magnitude is α^k .

Let $w^k(\mathbf{p}, x, \alpha)$ denote the budget share for good k of a household which is comprised of individuals $j = 1, \dots, J$, and has Barten economies of scale parameters α . If the household has e^x dollars to allocate toward purchasing goods, the household will spend the fraction $w^k(\mathbf{p}, x, \alpha)$ of e^x on the good k , for $k = 1, \dots, K$. Individuals living alone are assumed to have no economies of scale to consumption, and so have log Barten scale parameters equal to zero. For each good k and person j , let $w_j^k(\mathbf{p}, x)$ denote the budget share demand function of a household consisting just of person j living alone.

BCL prove that, subject to some technical conditions, if household purchase decisions are assumed to be Pareto efficient and if goods are shared by household members by the above Barten technology, then the

household having total log expenditures x and facing logged market prices \mathbf{p} will have purchased budget shares for each good k given by

$$w^k(\mathbf{p}, x, \alpha) = \sum_j \eta_j(\mathbf{p}, x, \alpha) w_j^k(\alpha + \mathbf{p}, x + \ln \eta_j(\mathbf{p}, x, \alpha)) \quad (1)$$

where $\eta_j(\mathbf{p}, x, \alpha)$ is the resource share of person j in the household and $\sum_j \eta_j(\mathbf{p}, x, \alpha) = 1$. The interpretation of this result is that efficiency is obtained by having each household member act as if they maximized their own utility function given a fraction η_j of the household's total expenditures e^x (which is why we call η_j a resource share) and facing log shadow prices $\alpha + \mathbf{p}$ which reflect the economies of scale from sharing.

We will consider households comprised of people who are observable both as single individuals and as members of collective households. This means that for each value of \mathbf{p} and x we can observe the budget shares w_j^k for people living singly, and the budget shares w^k for the household. BCL show that the remaining parameters of the model, the Barten scales α and the resource shares η_j , are identifiable from this data as long as individual's demand functions are sufficiently nonlinear. However, identification of these parameters requires price variation, and realistic models allowing for sufficient price flexibility and interactions require many parameters, which enter the household demand equations highly nonlinearly. These features make estimation of this model numerically difficult (e.g., BCL report finding ten different local minima for their model).

We propose testable restrictions on the preferences of individuals and Barten scales that can be used to identify resource shares and other parameters of interest without observing or modeling the effects of price variation, that is, using only Engel curve data. This greatly decreases data, specification, and computational burdens.

2.1 Indifference Scales

The two features of household demand functions that we are primarily interested in are resource shares and indifference curves. Most of the papers cited in the introduction consider identification of features of resource shares, which are often interpreted as measures of bargaining power within the household. Using data just on multiple person households, Browning and Chiappori (1998) show that resource shares cannot be nonparametrically identified. However, they also show that the derivatives of resource shares with respect to observable variables that affect these shares and do not affect individual's preferences (which they call "distribution factors") can be identified. BCL show that resource shares can be nonparametrically identified by combining data on multiple person households with data on single individuals living alone. Another application of this type of identification is Lise and Seitz (2004). These papers that obtain identification of resource shares require modeling and estimating how demand functions depend upon prices.

The previous section defined resource shares, and here we now define indifference scales. Let $V_j(\mathbf{p}, x)$ denote an ordinal indirect utility function describing the preferences of person j . This individual's demand functions $w_j^k(\mathbf{p}, x)$ for each good k are derived from $V_j(\mathbf{p}, x)$ using Roy's identity,

$$w_j^k(\mathbf{p}, x) = \frac{-\partial \ln V_j(\mathbf{p}, x) / \partial p^k}{\partial \ln V_j(\mathbf{p}, x) / \partial x}.$$

The indirect utility function $V_j(\mathbf{p}, x)$ implicitly defines person j 's indifference curves over bundles of goods.

Suppose a household has Barten (1964) scales α and person j in that household has resource share $\eta_j(\mathbf{p}, x, \alpha)$. BCL define an indifference scale $I_j(\mathbf{p}, x, \alpha)$ as the solution to

$$V_j(\mathbf{p}, \ln I_j(\mathbf{p}, x, \alpha) + x) = V_j(\alpha + \mathbf{p}, \ln \eta_j(\mathbf{p}, x, \alpha) + x) \quad (2)$$

As can be seen by this equation, if one took the total expenditures of the household, e^x , multiplied that by the indifference scale $I_j(\mathbf{p}, x, \alpha)$, and gave that amount of income to person j living alone (facing log market prices \mathbf{p}), then person j would be able to buy a bundle of goods that lies on the same indifference curve (to person j) as the bundle of goods that person j consumes as a member of the household, taking into account both his/her resource share $\eta_j(\mathbf{p}, x, \alpha)$, and the household economies of scale that come from sharing given by the Barton scales α .

Indifference scales differ from ordinary adult equivalence scales (see Lewbel 1997 for a survey) in that equivalence scales equate the utility of an individual to the utility of a household, and so suffer from fundamental identification problems associated with interpersonal comparisons of utility. Specifically, letting $V(\mathbf{p}, x)$ denote the indirect utility function of a household, an adult equivalence scale E_j is the solution to $V_j(\mathbf{p}, \ln E_j + x) = V(\mathbf{p}, x)$. Comparing this with the indifference scale that solves $V_j(\mathbf{p}, \ln I_j + x) = V_j(\alpha + \mathbf{p}, \ln \eta_j + x)$ shows that both equivalence scales E_j and indifference scales I_j are fractions of a household's income $\exp(x)$, but the indifference scale makes person j as well off living alone as he or she was inside the household, while the equivalence scale makes person j as well off as the household itself.

Unlike indifference scales, equivalence scales depend on how utility is cardinalized. To see this, let G be any strictly monotonically increasing function. Replacing a person j 's utility function V_j with the observationally equivalent utility function $\tilde{V}_j = G(V_j)$ in the definition of the equivalence scale, yields $\tilde{V}_j(\mathbf{p}, \ln E_j + x) = V(\mathbf{p}, x)$, which gives a different numerical value of E_j compared to the original definition using V_j . As a result, ordinary equivalence scales can never be identified from revealed preference data; they always require untestable assumptions regarding utility levels or functional forms for E_j . Of course, this analysis assumes one can even define a single utility function V for the household. Depending on how the household resolves its budget allocation decision, such a single measure of utility may not even exist.

In contrast, replacing person j 's utility function V_j with \tilde{V}_j in the definition of the indifference scale I_j gives $\tilde{V}_j(\mathbf{p}, \ln I_j + x) = \tilde{V}_j(\alpha + \mathbf{p}, \ln \eta_j + x)$, which yields the same numerical value for I_j as the definition using V_j . This shows that indifference scales depend only on the indifference curves of person j , not on how utility is cardinalized. As a result, unlike equivalence scales, indifference scales can at least potentially be identified just from revealed preference data. Intuitively, the reason for this difference is that equivalence scales compare the utility of two different entities, namely, an individual and a household. In contrast, indifference scales compare the same individual in two different settings, i.e., living alone facing market prices, versus living in a household, consuming his/her share of the household's resources and facing shadow prices.

The above construction of indifference scales assumes that individual's indifference curves over goods themselves remain the same whether living as a single or as a member of a household. Changes in consumption behavior between living alone and living with another person are attributed to jointness of consumption and resource sharing rather than taste change. If tastes do change in an identified way, then identification is still possible, though one would then need to choose whether to construct an indifference scale based on preferences as a single or on preferences as a member of a couple, analogous to the difference between compensating variation and equivalent variation. The choice of which scale to use would depend upon context, and the results might be less useful for welfare analyses unless the differences between these two cases are relatively small.

BCL identify indifference scales from revealed preference data. Their identification of these scales depends upon identifying features of the household's and individuals' demand functions including dependence upon price variation.

2.2 IB Scale Economies

We now define the Independence of Base assumption that we will use to simplify the BCL model. For each person j , assume there exists a scalar-valued function $D_j(\alpha, \mathbf{p})$ such that the indifference curves of individual j and the Barten parameters α satisfy the condition

$$V_j(\alpha + \mathbf{p}, x) = V_j(\mathbf{p}, x - \ln D_j(\alpha, \mathbf{p})) . \quad (3)$$

Note that equation (3) is a joint restriction on the behavior of the individual and the household, because it involves the individual's utility function V_j and the household's scale economy parameters α . The function $D_j(\alpha, \mathbf{p})$ is a measure of the cost savings experienced by person j resulting from the scale economies of living in the household with others, because it equates utility for that person under market prices to utility with Barten-scaled market prices. Specifically, $D(\alpha, \mathbf{p})$ is the true cost-of-living index for person j associated with the change from log market prices \mathbf{p} to within-household log shadow-prices $\alpha + \mathbf{p}$. These scale economies are assumed to be independent of the base expenditure (and hence utility) level at which they are evaluated. Consequently, we refer to this restriction as imposing independent of base scale-economies (IBSE), which connects with a usage of that term in the equivalence scale literature.

The restriction above is similar to the independence-of-base (IB) restriction studied by Lewbel (1989), Blundell and Lewbel (1991), Blackorby and Donaldson (1993), Pendakur (1999) and Blundell, Duncan and Pendakur (1998). The functional form used by Jorgenson and Slesnick (1987) is an example, which imposes a restriction on Barten scales like the above. However, whereas this previous literature restricts how indirect utility responds to changes in demographic characteristics, equation (3) only restricts how an individual's indirect utility responds to changes in the price vector \mathbf{p} given by the economies of scale measures α . Thus, as discussed in Blundell and Lewbel (1991), the IB condition in the equivalence scale literature imposes both a testable restriction on how prices affect indirect utility and an untestable restriction on how demographics affect indirect utility. In contrast, because equation (3) only restricts price effects (and, e.g., does not entail comparing the utility of two demographically different individuals or households), it imposes only a testable restriction on how prices affect indirect utility.

Applying Roy's identity to equation (3) shows that individual j 's budget share functions $w_j^k(\alpha + \mathbf{p}, x)$ satisfy

$$w_j^k(\alpha + \mathbf{p}, x) = d_j^k(\mathbf{p}, \alpha) + w_j^k(\mathbf{p}, x - \ln D_j(\mathbf{p}, \alpha)) . \quad (4)$$

where

$$d_j^k(\mathbf{p}, \alpha) = \frac{\partial \ln D_j(\mathbf{p}, \alpha)}{\partial p_k} \quad (5)$$

is the elasticity of D_j with respect to the k 'th price. Equation (3) is exactly analogous to the shape-invariance restriction implied by IB equivalence scales (see, e.g., Pendakur 1999). In particular, the budget-share of person j as a function of log total expenditures x have the same shape when evaluated at (logged) market prices \mathbf{p} or at (logged) shadow-prices $\alpha + \mathbf{p}$. They differ only in that they are translated over log-expenditure x by $\ln D_j(\mathbf{p}, \alpha)$ and over each w_j^k by $d_j^k(\mathbf{p}, \alpha)$.

For an example of a general model of individual preferences that satisfies equation (3), consider individuals having indirect utility functions of the form

$$V_j(\mathbf{p}, x) = G \left(\frac{F_j^2(\mathbf{p})}{\ln x - \ln F_j^1(\mathbf{p})} + F_j^3(\mathbf{p}) \right) . \quad (6)$$

where the functions F_j^1 , F_j^2 , and F_j^3 , must satisfy homogeneity and other restrictions implied by having V be an indirect utility function, and G is any unknown strictly monotonically increasing transformation. Applying Roy's identity to equation (6) yields rank three demand systems with budget shares that are quadratic in log total expenditures. An example of equation (6) is the Quadratic Almost Ideal (QAI) model of Banks, Blundell, and Lewbel (1997). Assume that the the Barten parameters α and the functions F_j^2 and F_j^3 together satisfy

$$\begin{aligned} F_j^2(\alpha + \mathbf{p}) &= F_j^2(\mathbf{p}) \\ F_j^3(\alpha + \mathbf{p}) &= F_j^3(\mathbf{p}). \end{aligned}$$

For example, we could have $F_j^2(\alpha + \mathbf{p}) = \mathbf{c}'(\alpha + \mathbf{p})$ for some constant vector \mathbf{c} where the vector \mathbf{c} and Barton scales α are such that $\mathbf{c}'\alpha = 0$, and similarly for F_j^3 . The Barten scales in Jorgenson, Lau, and Stoker (1982) and Jorgenson and Slesnick (1987) satisfy this type of linear restriction. One can then immediately verify in this example that equation (3) holds with D_j given by

$$\ln D_j(\alpha, \mathbf{p}) = \ln F_j^1(\mathbf{p}) - \ln F_j^1(\alpha + \mathbf{p}).$$

3 Identification

Our goal is to identify resource shares and indifference scales from demand data without price variation. We now summarize our modeling assumptions.

ASSUMPTION A: Assume the general BCL household model, so equation (1) holds. For each household member j assume there exists a differentiable in \mathbf{p} function $D_j(\mathbf{p}, \alpha)$, with derivatives denoted by $d_j^k(\mathbf{p}, \alpha)$ in equation (5), such that equation (3) holds. Assume resource shares η_j do not depend on x , and so are given by $\eta_j(\mathbf{p}, \alpha) > 0$. By BCL, $\sum_j \eta_j(\mathbf{p}, \alpha) = 1$.

In addition to the restrictions already discussed, Assumption A adds the additional restriction that resource shares do not depend on x . Restricting resource shares in this way imposes constraints on the household social welfare or bargaining function, but BCL provide general results showing existence of household models (specifically, of so called Pareto weights) given restrictions on η_j . This is a potentially very strong restriction, but it can be mitigated by including measures of household wealth other than total expenditures in η_j . For example, Browning, Bourguignon, Chiappori, and Lechene (1994) model resource shares as depending on total household income but not total expenditures. This would be permitted in our model by including total income among the distribution factors (which we will later define as \mathbf{z}_h) in the model. Resource shares are weights on total expenditures for each member, and so must be nonnegative and sum to one.

A limitation implicit in the BCL model is that an individual's indifference curves over goods are assumed to be the same whether living as a single or in a couple. Differences between an individual's bundle of consumed goods as a single versus within a household are therefore assumed to be due to partially joint consumption, resource sharing, and to changes in total resources, but are not attributed to preference changes. BCL discuss how their model could be extended to handle preference changes by directly observing some private consumption within the household, or by parameterizing preference changes, or by allowing parameters like α to include both preference changes and economies of scale to consumption. In our model, the functional forms we estimate remain valid if preference changes take the form of either affecting the Barten scales α , or price dependent location shifts in the budget share

functions w_j^k . However, such changes would complicate the interpretation of our estimates, particularly of indifference scales.

LEMMA 1: Let Assumption A hold. Then indifference scales do not depend on x , and are given by

$$I_j(\mathbf{p}, \alpha) = \frac{D_j(\mathbf{p}, \alpha)}{\eta_j(\mathbf{p}, \alpha)}. \quad (7)$$

Also household budget share demand functions are given by

$$w^k(\mathbf{p}, x, \alpha) = \sum_j \eta_j(\mathbf{p}, \alpha) \left[d_j^k(\mathbf{p}, \alpha) + w_j^k(\mathbf{p}, x - \ln I_j(\mathbf{p}, \alpha)) \right]. \quad (8)$$

PROOF: Equation (7) follows immediately from equations (3) and (2) with η_j independent of x . The second equation is then obtained by substituting equation (4) into (1) and applying equation (7).

Given Assumption A, which imposes the BCL collective household model, IB scale economies and expenditure-independent resource shares, equation (8) shows that household budget share equations are a simple function of individual budget share equations. The key distinction between this model and the unrestricted version of BCL is that household budget shares are a function of individual budget shares evaluated at the market, rather than the shadow, price vector (after translation). In particular, equation (8) shows that household budget shares are weighted averages of individual budget shares translated both in budget shares (vertically) and log-expenditure (horizontally). As in the equivalence scale literature (see, eg, Pendakur 1999), these translations are meaningful model parameters: translations in log-expenditure are individual indifference scales; translations in budget shares are weighted averages of scale economy price elasticities.

Now suppose we only observe data in one price regime, $\mathbf{p} = \mathbf{p}_0$. Both \mathbf{p}_0 and α are vectors of constants, so we can rewrite equation (8) in Engel curve form as

$$w^k(x) = h^k + \sum_j \eta_j w_j^k(x - \ln I_j). \quad (9)$$

for each good k , where $w^k(x)$ and $w_j^k(x)$ are respectively the household's and person j 's engel curves for good k (in the price regime \mathbf{p}_0 for a household with scale economy parameters α),

$$h^k = \sum_j \eta_j(\mathbf{p}_0, \alpha) d_j^k(\mathbf{p}_0, \alpha) = \sum_j \eta_j d_j^k \quad (10)$$

is a constant, and the resource shares η_j and indifference scales I_j are now constants for each household member j . We also have $\sum_j \eta_j = 1$ and $\eta_j > 0$ for each household member j .

We observe the expenditures of households and of individuals on each good k in this time period (price regime), so the Engel curve functions $w^k(x)$ and $w_j^k(x)$ are identified for each good k and person j . The question is then whether we can identify the resource shares η_j and indifference scales I_j from these Engel curves. Theorem 1 below shows that η_j and I_j are nonparametrically identified, as long as some of the goods have budget shares that are nonlinear and are sufficiently different across individuals. The following technical assumption formally describes sufficient nonlinearity restrictions.

ASSUMPTION B: Let $\Omega(s)$ denote the K by J matrix where the element in the k 'th row and j 'th column is $w_j^k(\ln s)$. Assume that the matrices $\partial\Omega(s)/\partial s$ and $\partial^2\Omega(s)/\partial s^2$, evaluated at $s = 0$, exist and have rank J .

Assumption B imposes restrictions on the behavior of Engel curves at the zero expenditure level. Given the assumed smoothness (differentiability) of Engel curves, this does not formally require observing budget shares at zero expenditure levels, but rather depends on the limits of Engel curve derivatives in the neighborhood of zero, i.e., at small expenditure levels. In practice, with actual data it may not be necessary to consider very low expenditure levels at all. Identification really just requires a certain degree of nonlinearity, which is most easily described at zero via Assumption B.

THEOREM 1: Let Assumptions A and B hold. Assume Engel curves $w^k(x)$ and $w_j^k(x)$ are identified for $j = 1, \dots, J$. Then all the parameters in equation (9) are also identified, that is, η_j and I_j are identified for $j = 1, \dots, J$ and h^k is identified for $k = 1, \dots, K$.

PROOF: Note that s is total expenditures, and $x = \ln s$ is log total expenditures. Also, by construction, each I_j is strictly positive. Define $\omega^k(s) = w^k(\ln s)$, $\omega_j^k(s/I_j) = w_j^k(\ln s - \ln I_j)$, $v_j^k(s) = \partial \omega_j^k(s) / \partial s$ and $\zeta_j^k(s) = \partial v_j^k(s) / \partial s$. By equation (9), $\omega^k(s) = h^k + \sum_{j=1}^J \eta_j \omega_j^k(s/I_j)$, so

$$\frac{\partial \omega^k(s)}{\partial s} = \sum_{j=1}^J \frac{-\eta_j}{I_j} v_j^k\left(\frac{s}{I_j}\right) \quad \text{and} \quad \frac{\partial^2 \omega^k(s)}{\partial s^2} = \sum_{j=1}^J \frac{\eta_j}{I_j^2} \zeta_j^k\left(\frac{s}{I_j}\right)$$

and therefore

$$\frac{\partial \omega^k(s)}{\partial s} \Big|_{s=0} = \sum_{j=1}^J \frac{-\eta_j}{I_j} v_j^k(0) \quad \text{and} \quad \frac{\partial^2 \omega^k(s)}{\partial s^2} \Big|_{s=0} = \sum_{j=1}^J \frac{\eta_j}{I_j^2} \zeta_j^k(0)$$

Letting $\omega(s)$ equal the vector of elements $\omega^k(s)$ for $k = 1, \dots, K$, these last two equations in matrix form are $\partial \omega(s) / \partial s \Big|_{s=0} = -[\partial \Omega(s) / \partial s \Big|_{s=0}] [\eta_1 / I_1, \dots, \eta_J / I_J]'$ and $\partial^2 \omega(s) / \partial s^2 \Big|_{s=0} = [\partial^2 \Omega(s) / \partial s^2 \Big|_{s=0}] [\eta_1 / I_1^2, \dots, \eta_J / I_J^2]'$. Since the derivatives of $\Omega(s)$ in these equations have full column rank, these equations can be solved for the vectors $[\eta_1 / I_1, \dots, \eta_J / I_J]$ and $[\eta_1 / I_1^2, \dots, \eta_J / I_J^2]$, which will then be expressed entirely in terms of derivatives of the functions $\omega^k(s)$ and $\omega_j^k(s)$ evaluated at $s = 0$, which are identified. Dividing each element in the first of these vectors by the corresponding element in the second yields the vector $[I_1, \dots, I_J]$, and multiplying each element in this vector by the corresponding element in $[\eta_1 / I_1, \dots, \eta_J / I_J]$ yields $[\eta_1, \dots, \eta_J]$. Finally, given I_j and η_j for all household members j , the constants h^k can be recovered from $h^k = w^k(x) - \sum_j \eta_j w_j^k(x - \ln I_j)$.

Theorem 1 essentially says that in a household with J people, the resource shares and indifference scales are identified if there are J goods having nonlinear Engel curves that differ both across people and across goods. For example, suppose that each individual j has budget share demand functions for good k given by $\omega_j^k(s) = \alpha_j^k + s\beta_j^k + s^2\gamma_j^k$ for some constants α_j^k , β_j^k , and γ_j^k where s is total expenditures (in levels, not in logged). Then $\partial \Omega(s) / \partial s \Big|_{s=0}$ is a matrix of elements β_j^k and $\partial^2 \Omega(s) / \partial s^2 \Big|_{s=0}$ is a matrix of elements γ_j^k , and by Theorem 1 all the parameters of the model are identified if these matrices have full column rank J , meaning that the vector $[\beta_1^k, \dots, \beta_J^k]$ for some good k is not a perfect linear combination of the corresponding β vectors for the other goods, and similarly for the vectors of γ_j^k elements. This example also illustrates the point that identification need not require observing demands in the neighborhood of zero, since the coefficients β_j^k and γ_j^k can be obtained just with data far from zero expenditure levels.

Note that Theorem 1 is sufficient but not necessary for identification. One might also obtain identification under weaker conditions by the presence of assignable goods (that is, goods that are only consumed by one household member), or by functional form restrictions.

4 Empirical Implementation

The households we consider are couples, so $J = 2$. We will index the members of the household by $j = f$ and $j = m$ (female and male). Since resource shares η_j sum to one we can just define a single share function $\eta = \eta_f$ with $1 - \eta = \eta_m$. Adding ε error terms, for $k = 1, \dots, K - 1$, the model for estimation can be written as the $3K - 3$ system of equations

$$\begin{aligned} w_f^k &= w_f^k(x) + \varepsilon_f^k \\ w_m^k &= w_m^k(x) + \varepsilon_m^k \\ w^k &= h^k + \eta w_f^k [x - \ln I_f] + (1 - \eta) w_m^k (x - \ln I_m) + \varepsilon^k \end{aligned}$$

where it should be understood that w_j^k is person j 's actual budget share and $w_j^k(x)$ is the Engel curve function for these shares. Equations for good K are unnecessary, since they follow from budgets shares summing to one. To exploit data from a survey of many people we must now specify how $w_j^k(x)$, h^k , I_j , and η vary by observable characteristics such as age.

We introduce a vector of demographic characteristics for each individual, \mathbf{z}_j , and a vector of distribution factors \mathbf{z}_h . Distribution factors are defined by Browning and Chiappori (1998) as variables that affect resource shares η_j but not the demand functions of individual household members. Thus, \mathbf{z}_j can enter the budget-share functions of the relevant singles, $w_j^k(x, \mathbf{z}_j)$ and the scale economies functions and elasticities of those singles, $D_j(\mathbf{z}_j)$ and $d_j^k(\mathbf{z}_j)$. Both \mathbf{z}_h and the individual characteristics \mathbf{z}_f and \mathbf{z}_m enter the resource share function $\eta_f(\mathbf{z}_f, \mathbf{z}_m, \mathbf{z}_h)$, and, since the indifference scale is equal to the scale economies divided by the resource share, these arguments also enter the indifference scale, $I_j(\mathbf{z}_f, \mathbf{z}_m, \mathbf{z}_h)$. Let $\mathbf{z} = (\mathbf{z}_f, \mathbf{z}_m, \mathbf{z}_h)$ denote the set of all the demographic characteristics, so the resource share and indifference scale functions are just $\eta_f(\mathbf{z})$ and $I_j(\mathbf{z})$. Adding demographic characteristics \mathbf{z} to the model gives us the estimating equations

$$w_f^k = w_f^k(x, \mathbf{z}_f) + \varepsilon_f^k \quad (11)$$

$$w_m^k = w_m^k(x, \mathbf{z}_m) + \varepsilon_m^k \quad (12)$$

$$w^k = h^k(\mathbf{z}) + \eta(\mathbf{z})w_f^k [x - \ln I_f(\mathbf{z}, \mathbf{z}_f)] + [1 - \eta(\mathbf{z})]w_m^k [x - \ln I_m(\mathbf{z}, \mathbf{z}_m)] + \varepsilon^k \quad (13)$$

where, based on equation (10),

$$h^k(\mathbf{z}) = \eta(\mathbf{z})d_f^k(\mathbf{z}_f) + [1 - \eta(\mathbf{z})]d_m^k(\mathbf{z}_m). \quad (14)$$

Theorem 1 shows that the parameters of interest, resource shares and indifference scales, are nonparametrically identified from Engel curve data. This means that, by nonparametrically estimating the Engel curve regression equations (11), (12), and $w^k = w^k(x, \mathbf{z}) + \varepsilon^k$ for each good k , we could nonparametrically recover the functions $\eta(\mathbf{z})$, $I_f(\mathbf{z})$ and $I_m(\mathbf{z})$, as well as $h^k(\mathbf{z})$ for each good k using equations (13) and (14). However, for empirical tractability we will flexibly parameterize the model instead of attempting nonparametric estimation. These parameterizations could be treated as polynomial nonparametric sieve estimators, though the order of the polynomials we specify is rather low for that interpretation.

Any parametric specification we propose must be consistent with equations (5), (7), (8), and (10), so the functions $I_m(\mathbf{z})$, $I_f(\mathbf{z})$, $\eta(\mathbf{z})$ and $h^k(\mathbf{z})$ must be related in the sense that $I_f(\mathbf{z})$ equals a $D_f(\mathbf{z}_f)$ function evaluated at \mathbf{p}_0 and α , divided by $\eta(\mathbf{z})$, and similarly for $I_m(\mathbf{z})$, while $h^k(\mathbf{z})$ equals a weighted average of the $d_f^k(\mathbf{z}_f)$ and $d_m^k(\mathbf{z}_m)$ functions with weights $\eta(\mathbf{z})$ and $1 - \eta(\mathbf{z})$. For parameterizing, it is convenient to rewrite equation (13) using (14) as

$$\begin{aligned} w^k &= \eta(\mathbf{z}) \left[w_f^k [x - \ln I_f(\mathbf{z}, \mathbf{z}_f)] + d_f^k(\mathbf{z}_f) \right] + \\ &\quad [1 - \eta(\mathbf{z})] \left[w_m^k [x - \ln I_m(\mathbf{z}, \mathbf{z}_m)] + d_m^k(\mathbf{z}_m) \right] + \varepsilon^k \end{aligned} \quad (15)$$

The functional forms we use for our empirical estimates are as follows. The budget share Engel curve models for single individuals are given by

$$w_f^k(x, \mathbf{z}_f) = a_f^{k0} + \mathbf{a}_f^{k'} \mathbf{z}_f + (x - \mathbf{e}'_f \mathbf{z}_f) b_f^k + (x - \mathbf{e}'_f \mathbf{z}_f)^2 c_f^k \quad (16)$$

for women and

$$w_m^k(x, \mathbf{z}_m) = a_m^{k0} + \mathbf{a}_m^{k'} \mathbf{z}_m + (x - \mathbf{e}'_m \mathbf{z}_m) b_m^k + (x - \mathbf{e}'_m \mathbf{z}_m)^2 c_m^k \quad (17)$$

for men, for each good $k = 1, \dots, K$. These budget shares are a rank three quadratic in log total expenditures, which many authors have found provide a good fit for Engel curves (see, e.g., Banks, Blundell, and Lewbel 1997), and are a quadratic in demographic characteristics restricted to satisfy shape invariance, which is also known to fit well (see, e.g., Blundell, Duncan and Pendakur 1998).

We take \mathbf{z}_f and \mathbf{z}_m to be the female's and male's age and highest level of schooling. We also normalise \mathbf{z}_f and \mathbf{z}_m so that they equals $\mathbf{0}$ for a reference set of characteristics, which in our example are possessed by a person aged 40 with High School as their highest level of school.

Substituting equations (16) and (17) into (15) then gives the model for couples of the form

$$w^k = \eta(\mathbf{z}) \left[a_f^{k0} + \mathbf{a}_f^{k'} \mathbf{z}_f + (x - \ln I_f(\mathbf{z}) - \mathbf{e}'_f \mathbf{z}_f) b_f^k + (x - \ln I_f(\mathbf{z}) - \mathbf{e}'_f \mathbf{z}_f)^2 c_f^k + d_f^k(\mathbf{z}_f) \right] \\ + [1 - \eta(\mathbf{z})] \left[a_m^{k0} + \mathbf{a}_m^{k'} \mathbf{z}_m + (x - \ln I_m(\mathbf{z}) - \mathbf{e}'_m \mathbf{z}_m) b_m^k + (x - \ln I_m(\mathbf{z}) - \mathbf{e}'_m \mathbf{z}_m)^2 c_m^k + d_m^k(\mathbf{z}_m) \right] + \varepsilon^k. \quad (18)$$

It remains to parameterize $\eta(\mathbf{z})$, $D_j(\mathbf{z}_j)$ and $d_j^k(\mathbf{z}_j)$. The indifference scale is given by $\ln I_j(\mathbf{z}) = \ln D_j(\mathbf{z}_j) - \ln \eta_j(\mathbf{z})$. We parameterize the resource share function $\eta(\mathbf{z})$ by

$$\eta(\mathbf{z}) = \mathbf{r}' \mathbf{z} = r_0 + \mathbf{r}'_h \mathbf{z}_h + \mathbf{r}'_m \mathbf{z}_m + \mathbf{r}'_f \mathbf{z}_f \quad (19)$$

Following Browning and Chiappori (1998), we take the distribution factor \mathbf{z}_h to be the relative wage of the husband versus the wife. Unlike many other models like Browning and Chiappori (1998), we do not require the presence of any distribution factors for identification, but we find this \mathbf{z}_h to be empirically significant.

Using the fact that \mathbf{p}_0 and α are invariant in our Engel curves, we parameterize $D_m(\mathbf{z}_m)$ and $D_f(\mathbf{z}_f)$ as

$$\ln D_m(\mathbf{z}_m) = d_{0m} + \mathbf{d}'_m \mathbf{z}_m, \quad \ln D_f(\mathbf{z}_f) = d_{0f} + \mathbf{d}'_f \mathbf{z}_f.$$

and the price elasticities of these IB scales are parameterised as

$$d_m^k(\mathbf{z}_m) = \delta_{0m}^k + \delta_m^{k'} \mathbf{z}_m, \quad d_f^k(\mathbf{z}_f) = \delta_{0f}^k + \delta_f^{k'} \mathbf{z}_f \quad (20)$$

for $k = 1, \dots, K$. These then imply the following functional forms for I_f and I_m :

$$\ln I_f(\mathbf{z}) = d_{0f} + \mathbf{d}'_f \mathbf{z}_f - \ln(\mathbf{r}' \mathbf{z}) \quad (21)$$

$$\ln I_m(\mathbf{z}) = d_{0m} + \mathbf{d}'_m \mathbf{z}_m - \ln(1 - \mathbf{r}' \mathbf{z}) \quad (22)$$

The model we actually estimate is then obtained by substituting equation (16) into (11) and (17) into (12) for singles, and substituting equations (19), (20), (21), and (22) into equation (18) for couples. We then have three equations for each $k = 1, \dots, K - 1$, giving a total system of $3K - 3$ equations. The equation for the K 'th good in each household type (m, f, and couple) does not need to be estimated, because its parameters are completely determined by budget shares summing to one for each household.

By Theorem 1, the above model is identified, as long as the 2 by K matrices consisting of rows $\begin{bmatrix} b_f^k, b_m^k \end{bmatrix}$ and $\begin{bmatrix} c_f^k, c_m^k \end{bmatrix}$ each have rank two and $\eta(\mathbf{z})$ is not equal to zero or one. However, precision of estimation is likely to be strengthened by the presence of an assignable good. This improvement is in the sense of reducing multicollinearity, or, more formally, in improving the conditioning of the covariance matrix of the estimated parameters, thereby mitigating likely numerical convergence issues and decreasing standard errors.

We call a good *assignable* if it is consumed exclusively by only one person in the collective household. To illustrate, suppose women are the exclusive consumers of women's clothing. Then, letting $k = f$ denote the women's clothing equation, we have $w_f^f = w_f^f(x, \mathbf{z}_f) + \varepsilon_f^f$ and $w_m^f = 0$, that is, the budget share of women's clothing in the men's demand function is zero. Also $d_m^f(\mathbf{z}_m) = 0$, that is, the effect of women's clothing prices on men's scale economies is zero. This then makes the couples equation for womens clothes simplify from equation (18) with $k = f$ to just

$$\begin{aligned} w^f &= \eta(\mathbf{z}) \left[w_f^f [x - \ln I_f(\mathbf{z}), \mathbf{z}_f] + d_f^f(\mathbf{z}_f) \right] + \varepsilon^f \\ &= \eta(\mathbf{z}) \left[a_f^{f0} + \mathbf{a}_f^f \mathbf{z}_f + \left(x - \ln I_f(\mathbf{z}) - \mathbf{e}_f' \mathbf{z}_f \right) b_f^f + \left(x - \ln I_f(\mathbf{z}) - \mathbf{e}_f' \mathbf{z}_f \right)^2 c_f^f + d_f^f(\mathbf{z}_f) \right] + \varepsilon^f \end{aligned}$$

An analogous simplification applies to the couple's share of any good that is only consumed by men. In our empirical application we consider $K = 12$ goods, and two of these goods, men's and women's clothing, are assumed to be assignable.

5 Data and Empirical Results

Data from the 1990 and 1992 Canadian Family Expenditure Surveys (Statistics Canada, 1990, 1992) are used to estimate the Engel curve system for $K = 12$ commodities: food at home, food out, gasoline, personal care, toys, recreation services, tobacco, alcohol, men's clothing, women's clothing, rent, and household furnishings and equipment. In these data, expenditures in each category are recalled by respondents over the previous 12 month period. In order to treat men's and women's clothing as assignable, we give a value of zero to men's clothing expenditures by single women and to women's clothing by single men. (Approximately 4% of each group reported nonzero expenditures in these categories.) We deflate expenditure in 1992 by the 2-year price change, and so treat the pooled data over these two years as facing the same relative price vector. Actual price changes over 1990-1992 for the commodity groups used in this paper were between 4.4% (for food-at-home) and 6.3% (for personal care) with the exception of clothing (one price index for men's and women's clothing rose by 10%).

Our selection criterion for households is as follows. Only rental-tenure households residing in cities with more than 30,000 people, where all members are full-year members and all members are aged 25-59 are used. We only select households comprised of either single men, single women, or married childless couples. Our sample consists of 419 single men, 450 single women, and 332 couples that meet these criteria.

The vectors \mathbf{z}_f and \mathbf{z}_m are each comprised of two demographic variables: age minus 40 and years of education minus 12. These vectors equal zero for men and women aged 40 with 12 years of education. We define distribution factors \mathbf{z}_h to be the demeaned woman's share of gross household income (the mean is 0.40) and an additional dummy for whether the woman's share of gross household income is less than 0.10. Often age differences and education differences between men and women are taken to be additional distribution factors, but these are already implicitly included in our model of resource shares, which we specify as depending on \mathbf{z} , and hence on \mathbf{z}_f and \mathbf{z}_m as well as \mathbf{z}_h .

Table 1: Descriptive Statistics

Budget Shares	Single Men		Single Women		Married Couples	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Food-in	0.160	0.087	0.164	0.091	0.193	0.083
Food-out	0.103	0.081	0.073	0.062	0.086	0.068
Gasoline	0.044	0.051	0.031	0.037	0.055	0.042
Personal Care	0.024	0.014	0.047	0.027	0.039	0.019
Toys	0.012	0.043	0.007	0.023	0.010	0.025
Recreation	0.070	0.064	0.057	0.054	0.062	0.049
Tobacco	0.047	0.073	0.039	0.065	0.055	0.069
Alcohol	0.056	0.070	0.024	0.034	0.041	0.049
Men's Clothing	0.060	0.053			0.037	0.027
Women's Clothing			0.087	0.066	0.054	0.040
Rent	0.389	0.133	0.427	0.131	0.321	0.099
Household Furn/Eq	0.037	0.050	0.044	0.055	0.056	0.056
log-expenditure	-0.069	0.455	-0.188	0.416	0.338	0.352
age (less 40)	2.723	9.376	4.496	10.121	1.121	10.125
education years (less 12)	0.141	1.368	0.124	1.314	0.198	1.294
woman's income share (less 0.4)					0.006	0.230
woman's share < 10%					0.139	0.346

Table 1 summarizes our data. Using assignability of clothing, a rough first estimate of resource shares can be obtained from the clothing shares in Table 1 as follows. If single men and women had budget shares independent of prices, expenditure and demographics (that is, Cobb-Douglas utility), then the resource share of men and women in married couples would just be given by the ratios of budget shares of the assignable goods for married couples versus singles. This gives crude estimates of men's resource shares equal to $0.037/0.060 = 0.62$ and women's resource shares $0.054/0.087 = 0.62$. These may be within twenty percent or so of reasonable values, but they violate the restriction of shares summing to one. We now consider more realistic estimates.

Table 2 gives estimated parameters and asymptotic standard errors for the collective household parameters that enter the D_f , D_m , and η functions. We ran four models: Model 1 contains all observations and all demographic variables \mathbf{z} ; Model 2 drops outliers defined as those single men and women and married couples lying in the top or bottom 5% of their respective total expenditure distributions; Model 3 drops education as a demographic variable; and Model 4 drops age as a demographic variable. Estimates for the roughly 150 coefficients comprising the singles' budget share equations for each model are available on request from the authors. We note that these equation systems are only slightly nonlinear, so that despite the large number of parameters to estimate, the models required only a few minutes to run, and we found no evidence of multiple local minima. This contrasts with BCL, who report finding "10 distinct local minima, each of which required a few days of iteration to converge."

Table 2: Parameter Estimates

	Model 1		Model 2		Model 3		Model 4	
	Est	Std Err	Est	Std Err	Est	Std Err	Est	Std Err
$d_{0,f}$	-0.351	0.406	-0.619	0.333	-0.233	0.277	-0.334	0.287
$d_{0,m}$	-0.245	0.377	-0.159	0.437	-0.301	0.269	-0.223	0.215
r_0	0.460	0.078	0.361	0.074	0.455	0.077	0.400	0.076
$r_{inc\ share}$	0.074	0.032	0.076	0.030	0.080	0.038	0.138	0.040
$r_{inc_share=0}$	0.012	0.015	0.005	0.018	0.022	0.017	0.007	0.023
$d_{age,f}$	-0.050	0.024	-0.064	0.017	-0.045	0.016		
$d_{age,m}$	-0.014	0.014	0.016	0.015	0.010	0.012		
$r_{age,f}$	0.000	0.001	0.004	0.002	0.003	0.002		
$r_{age,m}$	-0.001	0.001	-0.007	0.002	-0.008	0.002		
$d_{educ,f}$	0.125	0.149	-0.254	0.097				
$d_{educ,m}$	-0.210	0.107	-0.070	0.146				
$r_{educ,f}$	0.026	0.009	0.027	0.096				
$r_{educ,m}$	-0.018	0.013	-0.024	0.012				
include outliers?	Yes		No		No		No	
include education?	Yes		Yes		No		No	
include age?	Yes		Yes		Yes		No	
# observations	1201		1082		1082		1082	
# coefficients	177		177		131		85	
Log-Likelihood	22,600.2		20,526.5		20,459.8		20,368.8	

One should expect the scale-economy parameters comprising D_f and D_m to lie between 1/2 and 1, corresponding to the extremes of completely sharing consumption of all goods versus no sharing (completely private consumption). For persons aged 40 with 12 years of education, scale economies for women and men are given by $\exp(d_{0,f})$ and $\exp(d_{0,m})$, respectively. The point estimates from model 1 give scale economies of 0.70 for women and 0.78 for men of this type in married-couple households. These point estimates thus imply that for women aged 40 with 12 years of education, the shadow-prices faced within the married couple are associated with a cost-of-living index that is 70% of the costs faced by single women who face market prices. Men face slightly smaller scale economies within the married couple household. Thus these estimates are near the middle of the expected range between a half and one.

Although reasonable in magnitude, these scale economies are very imprecisely estimated. These scale economy parameter estimates are not statistically significant, and there is a fair amount of instability in the parameter estimates across the models. The basic problem is that the scale economy parameters $d_{0,j}$ shift the married-couple's Engel curves horizontally in $\ln x$, and are consequently identified only by the relationship between the quadratic and linear effects in $\ln x$, which are imprecisely estimated due to multicollinearity. The same problem arises in models that estimate equivalence scales from Engel curves. Model 4 provides the most precise scale economy estimates, because it drops outliers and some covariates, but even this model's standard errors are uncomfortably large.

Demographics affect the demands of single people, and so should also affect the scale economies faced by people when they live in married couple households. The estimates suggest that older women in married couples have larger scale economies (more negative $\ln D_f$) than younger women in married couples, while for men age has very little effect on scale economies. The education effects are quite unstable across the specifications and, like the levels of scale economies, are very imprecisely estimated.

The estimates of resource shares η are much more precise than those measuring scale economies, and that they are more stable across specifications. The parameter r_0 gives the resource share η of a woman

aged 40 with 12 years of education who earns 40% of the gross income in a married-couple household. The estimated value of r_0 ranges from 0.36 to 0.46, and has a standard error near 0.08. Our use of assignable goods (men's and women's clothing) strengthens the identification of resource shares, and thereby aids the precision of our estimates of r_0 coefficients. This ability to identify and estimate the level of the resource share, rather than just its response to changes in distribution factors, is a novel feature of BCL and of our model which is not shared by the earlier literature on collective household model estimation.

We also estimate the effect of demographic variables on the resource share. We find that the difference in age between the spouses has little effect on resource shares, but the husband's age may be negatively related to the wife's resource share. In Models 2 and 3, the estimated value of $r_{age,m}$ is near -0.008, which implies that a 45 year old couple has a women's resource share that is 12 percentage points smaller than a 30 year old couple. For education, the pattern is reversed. The education level of the man does not much affect the resource shares, but the difference in education level between the man and women is positively related to the wife's share. That is, women have larger resource shares in married couples when they have less education than their husband. If we interpret this as due to a distribution factor, then it is somewhat puzzling, because one might expect that if women have more human capital, then their ability to extract within-household resources would be larger. However, as we noted above, one cannot distinguish a pure education (demographic) effect from a relative education (distribution factor) effect.

The gross income share of women is highly statistically significant as a distribution factor in all models, but the magnitude of its effect on resource shares is modest. The parameter estimates range from 0.074 to 0.138, which implies that a woman who earns half of household income has a resource share that is 4 to 7 percentage points larger than a woman with no personal income. The dummy indicating that a woman has less than 10% of gross household income is insignificant in all 4 specifications.

A big difference between our empirical results and the full demand-system estimates of corresponding parameters in BCL is that our estimates suggest that women have resource shares less than 1/2. In contrast, BCL report women's resource shares in excess of 0.60. However, our results are more in line with Lise and Seitz (2004) though their estimates are based on just consumption and leisure, and with Cherchye, De Rock, and Vermeulen (2008), who apply the BCL model to Dutch data and report an average women's resource share of .49.

Indifference scales, which give the scale to household income that puts a single person on the same indifference curve they would attain in a married couple, are equal to the scale economy measure divided by the resource share in our model. Thus, given the estimates from Model 1, a woman aged 40 with 12 years of education who earns 40% of gross household income has an estimated indifference scale of $0.70/0.46 = 1.53$. This implies that such a woman, if she were living alone, would need approximately two-thirds ($1/1.53$) of the couple's income to reach the same indifference curve, and hence the same standard of living, that she attains as a member of the couple.

6 Conclusions

Collective household models are those in which the household is characterised as a collection of individuals, each of whom has a well-defined objective function, and who interact to generate household-level decisions. Typical micro-data sources only have information on household-level choices, but many of the objects of interest in these models are features of and constraints faced by the individuals who together make up the household. Existing structural models used to estimate these objects are either easy to estimate and extremely restrictive, or are unrestrictive but are difficult to estimate and have demanding data requirements. In this paper, we provide a middle ground, where the model is mildly restrictive yet easy to estimate, using data that does not require price variation.

We propose and estimate a version of Browning, Chiappori and Lewbel (2004) in which all the objects of interest are estimable via nearly linear Engel curve estimation. These objects include resource shares and indifference scales, both of which are quite useful in a range of policy and welfare applications. In contrast to most collective household models we identify the levels of resource shares, not just how they change in response to distribution factors like wives' versus husbands' relative wage levels. We obtain reasonable estimates of these objects in an empirical application of our model.

Our model relies on relatively strong assumptions regarding possible changes in individuals' preferences after marriage, by assuming that differences between the purchasing patterns of singles versus couples are largely due to jointness of consumption of goods and the separate allocation of household resources to husbands and wives. A useful direction for further work would be to explore ways to incorporate more preference changes after marriage. This might be done through some combination of exploiting current overidentification in the model and obtaining additional data. Such data could include direct observations of the allocation and separate consumption of some goods by individuals within the household. This type of data might also allow us to extend the model to handle children in the household.

Another useful extension would be to tie our results to the marriage market literature. For example, our main assumption regarding the process by which resources are allocated between spouses is Pareto efficiency. A more explicit model of the bargaining process could lead to useful additional modeling restrictions, e.g., one might assume that a woman could threaten to divorce (or never marry) if her resource share in a household puts her on a much lower indifference curve than the one she could attain living alone with her own wages. Our indifference scales are precisely the information that would be required to perform this calculation.

References

- [1] Banks, J., R. Blundell, and A. Lewbel (1997), "Quadratic Engel curves and Consumer Demand," *Review of Economics and Statistics*, 79, 527–539.
- [2] Barten, A. P. (1964), "Family Composition, Prices, and Expenditure Patterns," in *Econometric Analysis for National Economic Planning: 16th Symposium of the Colston Society*, P. Hart, G. Mills, and J. K. Whitaker, eds., 277-292, London: Butterworth.
- [3] Becker, G. (1965), "A Theory of the Allocation of Time," *Economic Journal*, 75, 493-517.
- [4] Becker, G. (1981), "Altruism in the Family and Selfishness in the Market Place," *Economica*, 48, 1-15
- [5] Blundell, R., A. Duncan and K. Pendakur (1998), "Semiparametric Estimation of Consumer Demand", *Journal of Applied Econometrics* 13, 435-461.
- [6] Blackorby, C., and D. Donaldson (1993), "Adult-Equivalence Scales and the Economic Implementation of Interpersonal Comparisons of Well-Being," *Social Choice and Welfare*, 10, 335-61.
- [7] Blundell, R. and A. Lewbel (1991), "The Information Content of Equivalence Scales," *Journal of Econometrics*, 50, 49-68.
- [8] Bourguignon, F. and P.-A. Chiappori, (1994) "The Collective Approach to Household Behavior," in *The Measurement of Household Behavior*, R. Blundell, I. Preston, and I. Walker, Eds., 70-85, Cambridge: Cambridge University Press

- [9] Browning, M., F. Bourguignon, P.A. Chiappori, and V. Lechene (1994), "Incomes and Outcomes: a Structural Model of Intrahousehold Allocation," *Journal of Political Economy*, 102, 1067-1096.
- [10] Browning, M., and P.A. Chiappori, (1998), "Efficient Intra-household Allocations: A General Characterization and Empirical Tests," *Econometrica*, 66, 1241-1278.
- [11] Browning, M., P.A. Chiappori and A. Lewbel, (2004) "Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power," *Boston College Working Papers in Economics* 588.
- [12] Cherchye, L., B. De Rock, and F. Vermeulen (2008), "Economic Well-Being and Poverty Among the Elderly: An Analysis Based on a Collective Consumption," *Tilburg University CentER Discussion Paper No. 2008-15*.
- [13] Chiappori, P.-A., (1988), "Nash-Bargained Households Decisions: A Comment," *International Economic Review*, 29, 791-796.
- [14] Chiappori, P.-A., (1992), "Collective Labor Supply and Welfare," *Journal of Political Economy*, 100, 437-467.
- [15] Jorgenson, D. W., L. J. Lau, and T. M. Stoker (1982), "The Transcendental Logarithmic Model of Aggregate Consumer Behavior," in: *Advances in Econometrics*, R. Basman and G. Rhodes, Eds., Greenwich: JAI Press.
- [16] Jorgenson, D. W., and D. T. Slesnick, (1987), "Aggregate Consumer Behavior and Household Equivalence Scales," *Journal of Business and Economic Statistics* 5, 219-232.
- [17] Lewbel, A. (1989), "Household Equivalence Scales and Welfare Comparisons," *Journal of Public Economics*, 39, 377-391.
- [18] Lewbel, A. (1997), "Consumer Demand Systems and Household Equivalence Scales," *Handbook of Applied Econometrics, Volume II: Microeconomics*, M. H. Pesaran and P. Schmidt, eds., Oxford: Blackwell Publishers Ltd.
- [19] Lindahl, E., (1919), "Just Taxation - A Positive Solution," in R. Musgrave and A. Peacock, eds., *Classics in the Theory of Public Finance*, London: MacMillan.
- [20] Lise, J. and S. Seitz, (2004), "Consumption Inequality and Intra-Household Allocations" *Queen's IER Working Paper No. WP1019*.
- [21] Pendakur, K. (1999) "Estimates and Tests of Base-Independent Equivalence Scales," *Journal of Econometrics*, 88, 1-40.
- [22] Vermeulen, F., (2002), "Collective Household Models: Principles and Main Results," *Journal of Economic Surveys*, 16, 533-564.