Identification of Children's Resources in Collective Households

Geoffrey Dunbar, Arthur Lewbel, and Krishna Pendakur
Simon Fraser University, Boston College and Simon Fraser University

April 2010

Abstract

Children's resources matter, but they are hard to identify because consumption is typically measured at the household level. Modern collective household models permit some identification of household member resources, but these models typically either ignore children, or treat them as attributes of adults. We propose a collective household model in which children are people with their own utility functions (possibly assigned to them by parents). Extending the frameworks of Browning, Chiappori and Lewbel (2007) and Lewbel and Pendakur (2008), we show identification of children's resource shares within households, despite not being able to observe the consumption by individual household members of goods that are partly or wholly shared goods within the household. Specifically, by looking at how the budget shares for men's, women's and children's clothing and shoes vary across households in Malawi, with differing income levels and numbers of children, our structural model allows us to back out an estimate of the fraction of total household income that is consumed by each family member. Our identification does not require an assumption that persons with and without children share the same preferences. Further, we obtain identification using only Engel curves, and so do not require observed price variation in the data. Our models may be used to assess the impacts on children of policy interventions like micro lending that affect the level and distribution of income within households.

JEL codes: D13, D11, D12, C31, I32. Keywords: Collective Model, Cost of Children, Bargaining Power, Identification, Sharing rule, Demand Systems, Engel Curves. We would like to thank Martin Browning, Pierre-Andre Chiappori, and Federico Perali for helpful comments. 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

Most measures of economic well-being rely, to some degree, on individual consumption. Yet the measurement of individual consumption in data is often confounded because consumption is typically measured at the household, not the individual, level. 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 such as consumption expenditures. Given household data, useful measures of individual consumption expenditures are *resource shares*, defined as each member's share of total household consumption. If there is intra-household inequality, these resources shares will be unequal and per-capita measures are uninformative (or are at least misleading measures) of individual well-being.

Children differ from other household members in that they do not enter households by choice, they have little ability to leave, and generally bring little income or other resources to the household. Children may therefore be the most vulnerable of household members to intra-household inequality. It is thus imperative to measure children's resource shares in households in order to assess inequality and child poverty. This paper shows identification of children's resource shares in a collective household model, and offers simple methods to estimate them.

Children's resource shares in the collective household literature are not well understood. Children in collective household models are usually modeled as household attributes, or as consumption goods for parents, rather than as separate economic agents with individual utility/felicity functions. See, *e.g.,* Blundell, Chiappori and Meghir (2005). The implication is that children suddenly acquire utility functions once they reach adulthood. It may be a less extreme assumption to consider children as economic agents throughout their lives. Even if they are not fully expressing their own preferences when young, it is reasonable to assume that parents will try to allocate resources to maximize some measure of children's well being and hence utility. Our paper starts with the assumption that children are people with utility.

Dauphin et al (2008), and Cherchye, De Rock and Vermeulen (2008), test whether observed household demand functions are consistent with children having separate utility functions. They find some evidence that households behave as if children do have separate utility functions. Cherchye, De Rock and Vermeulen (2010) consider estimation, but their method generally only yields bounds on resource shares, rather than estimates of resource shares. Dauphin et al (2008) do not offer estimation or bounding.

There are a number of practical and technical difficulties in identifying and estimating household resource shares. First, many goods are shared and consumed jointly. Most collective models (including Dauphin et al 2008 and Cherchye, De Rock and Vermeulen 2008, 2010) either assume all goods are purely private (like food), or treat each good as being either purely private or purely public (like heat) within a household. But in reality many goods are partly shared, e.g., an automobile or cart may be used by a single household member part of the time, and by multiple members at other times. Second, even for goods that are privately consumed, we often only have data on the entire household's purchases of the good, and not its allocation to individual members. Our method of identification and estimation deals with both of these problems.

Based on the collective household model of Chiappori (1988, 1992), a series of papers starting from Bourguignon and Chiappori (1994), Browning, Bourguignon, Chiappori, and Lechene (1994), and Browning and Chiappori (1998) show identification of changes in resource shares as functions of some observables (called distribution factors). However, these papers (along with more recent variants such as Vermeulen (2002) and Lise and Seitz (2004)) do not identify the level of resource shares, and typically cannot be applied to model changes in children's resource shares because generally no observable distribution factors for children will exist. Later versions of some of these models can identify levels of (not just changes in) resource shares under some difficult to verify conditions (see, e.g., Chiappori and Ekelund 2008), but all the models in this class impose strong restrictions on how goods may be shared within households, specifically, they assume that all goods are either purely private or purely public within the household.

Browning, Chiappori and Lewbel (2007) (hereafter BCL) provide a model that nonparametrically identifies the levels of resource shares of all individual household members and which allows for very general forms of sharing of goods But, they show identification only when the demand functions of individuals can be separately observed, which is not the case for children since they are always in households that include adults. In practice, BCL observe the demand functions of individuals by observing data from single men and women living alone, and combine that with data the demands of couples living together, assuming limited differences between the utility functions of single and married men and between those of single and married women.

Our contribution is to extend the model of BCL to include children. Specifically, we show semiparametric identification of children's resource shares in the BCL model. We use two identification strategies. In the first, we assume that preferences are similar in certain limited ways across people (within household types), and use this similarity to help identify resource shares within households with a given number of children. In the second, we assume that preferences for a person are similar across household types, and compare the consumption choices of people across households with varying numbers of children. In comparison with BCL, we do not need to use information on childless households (either couples or singles). In that respect, our identification strategies impose milder conditions on preference stability across household types, since e.g. we would assume that fathers of two children have similar preference to fathers of three children, rather than assume that either are similar to single men. Related identification ideas go back at least to Lazear and Michael (1988, chapter 4).

Our identification uses private assignable goods. A good is defined to be *private* if it cannot be shared or consumed jointly by more than one person, and is defined to be *assignable* if it is consumed by one individual household member that is known to the researcher. Examples could include toys and diapers which are private goods assignable to children, or alcohol and tobacco which are private goods assignable to adults. Chiappori and Ekelund (2008) and Cherchye, De Rock and Vermeulen (2008) among others show how assignable goods can aid in the identification of resource shares. Our strategy follows this line in assuming the presence, and observability, of a small number of private assignable goods, and uses these to identify childrens' resource shares.

Previous papers have used private assignable goods to address children's resources without invoking a structural model of the household. For example, Lundberg, Pollak and Wales (1997) find that household budget shares on children's clothing are higher when households have (exogenously) higher female incomes, and conclude that children are therefore better off when female incomes are higher. This assumes a direct monotonic link between a child's clothing share and his or her economic well being. In contrast, we provide a structural model for calculating the child's economic well being, defined as the total amount of the household's resources consumed by the child, which is based in part on budget share equations for private assignable goods like clothing. Our structural model shows that the level of budget shares mixes both a price response (coming in part from the extent to which some goods are consumed jointly) and an income response. Our identification of children's resources accounts for these two types of responses.

We show that, exploiting semi-parametric restrictions similar to Pendakur (1999) and Lewbel and Pendakur (2008), we can identify resource shares using Engel curves. This greatly facilitates empirical application of the model as we do not require price data and do not model price effects. Thus, basing identification and estimation on Engel curves substantially reduces both model complexity and data requirements. These types of models could be used in principle to assess the impacts on children of policy interventions like school lunch programs or micro lending to women that can affect both the level and distribution of income within households.

We present empirical results for children's resource shares in Malawi. Malawi is one of the poorest countries in the world, with per-capita (2005 PPP) GDP of US$773 in 2008. Given the extreme poverty of most Malawian households, one may suspect that children are exceedingly vulnerable to intra-household inequality. We find that children command a reasonably large share of resources – roughly 20 percent for the first child – and that this share rises with the number of children – 5-10 percentage points per additional child. Moreover, fathers command a larger share of resources than mothers, and mothers seem to sacrifice more resources than fathers to their children. This latter finding is similar to Duflo (2003) who finds evidence that male household heads tend not to allocate additional resources to children while female household heads do. Our results are however not directly comparable to hers as we only consider two parent households with a male household head.

We find some evidence of gender discrimination within the household, similar to Rose (1999). We find that the father's resource shares rises as the proportion of children that are girls rises. Indeed, if all children are girls, then the father's resource share rises roughly five percentage points while the mother's share remains unchanged.

Yet, on the whole, our estimates suggest that while perhaps vulnerable, children in Malawi are not unvalued. Thus, if poverty is measured relative to an absolute level of resources, child poverty reduction should be achievable through development projects that increase household incomes.

2 Collective Households and Resource Shares

In the version of the BCL model we consider, 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. Within the household, each member faces this income constraint and a vector of Lindahl (1919) type shadow prices for goods. Each household member has a potentially different resource share, but all members face the same shadow price vector. The resource share of a person and shadow price vector of the household together define a shadow budget constraint faced by each individual within the household. Each household member then determines their own demand for each consumption good by maximizing their own utility function.

These shadow prices differ from market prices 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 by multiple household members. Goods that are not shared (ie, private goods) will have shadow prices equal to market prices. Each member faces the same shadow prices because the degree to which a good can be shared is an attribute of the good, rather than an attribute of the consumer.

The shadow budget constraint faced by individuals within households can be used to conduct consumer surplus exercises relating to individual well-being. One example of this is the construction of 'indifference scales', a tool BCL develop for comparing the welfare of individuals in a household to that of individuals living alone, analogous to an equivalence scale.

Resource shares for each individual may also be of interest even without knowledge of shadow prices. The resource share times the household expenditure level gives the extent of the individuals' budget constraint and is therefore a useful indicator of that individual's material well-being. For example, Lise and Seitz (2008) use these to construct national consumption inequality measures that account for inequality both within and across households.

In addition, because within-household shadow prices are the same for all household members, resource shares describe the relative consumption levels of each member. Consequently, they can be used to evaluate the relative welfare level of each household member, and are sometimes used as measures of the bargaining power of household members. BCL show a one to one relationship between resource shares and collective household model utility pareto weights, which are also used as measures of member bargaining power. Since we focus on the estimation of children's resource shares, we will not be interpreting the results in terms of bargaining power.

2.1 The Model

We begin by summarizing the BCL model, extended to include children. In general, we use superscripts to index goods, subscripts to index people and households. We consider three types t of individuals: m, f, and c, indicating male adult, female adult, and child. Our results readily extend to more types of individuals, such as younger and older children or boys and girls, but to simplify the presentation consider only households consisting of a mother, a father, and one or more children, so we can index households by the size measure $s = 1, 2, ...$ where s is the number of children in the family. Also to simplify notation, for now we suppress arguments corresponding to attributes like age, location, etc., that may affect preferences. We also suppress arguments corresponding to distribution factors, that is, variables like relative education levels that may help to determine bargaining power and hence resource shares devoted to each household member. All of our identification results may be conditioned on these types of variables, and when it comes to the empirical section, we will introduce them explicitly.

Households consume K types of goods. Let $p = \left(p^1, ..., p^K\right)'$ be the $K-$vectors of market prices and $z_s = \left(z_s^1, ..., z_s^K\right)'$ be the $K-$vectors of quantities of each good k purchased by a household of size s. Let $x_t = \left(x_t^1, ..., x_t^K\right)'$ be the $K-$vectors of quantities of each good k consumed by an individual of type t. Let x_{ts} denote this vector in a household with s children.

Let y denote total expenditure, which may be subscripted for households or individuals. Let $U_t\left(x_t\right)$ denote an ordinal measure of the utility that an individual of type t would attain if he or she consumed the bundle of goods x_t while living in the household. An individual's total utility may depend on the well being of other household members, on leisure and savings, and on being a member of a household, so $U_t\left(x_t\right)$ should be interpreted as just a subutility function over goods this period, which may be just one component of member t's total utility. For children, $U_c\left(x_c\right)$ might not represent their actual utility function over the bundle of goods x_c that the child consumes, but rather the utility function that parents believe the child has (or think he or she should have).

For their identification, BCL assume that for a person of type t, $U_t\left(x_t\right)$ also equals the utility function over goods of a single person of type t living alone. The Marshallian demand functions of a person t living alone, are then obtained by choosing x_t to maximize $U_t\left(x_t\right)$ under the linear budget constraint $p'x_t = y$. We do not impose this assumption, so for us $U_t\left(x_t\right)$ only describes the preferences over goods of individual t as a member of a family, which may be completely different from that person's preferences if he or she were living alone. In particular, it would not be sensible to define $U_c\left(x_c\right)$ as the utility function of a child living alone.

For simplicity, we assume that each child in a family has the same utility function $U_c\left(x_c\right)$. The underlying source of these preferences does not matter, e.g., this utility function could be imposed on them by parents. We may readily extend the model to include parameters that allow $U_c\left(x_c\right)$ to vary by, e.g., age and sex of the child, but these like other observed household and individual characteristics are omitted for the time being. However, up to the inclusion of such observable characteristics, we assume that the individual household member utility functions $U_f\left(x_f\right)$, $U_c\left(x_c\right)$, $U_m\left(x_m\right)$ are the same regardless of whether the household has one, two, or three children. So, e.g., in a household with given observed characteristics, mothers have the same preferences over privately consumed consumption goods regardless of how many children are in the household.

We assume that the total utility of person t is weakly seperable over the subutility functions for goods. So, e.g., a mother who gets utility from her husband's and child's well-being as well as her own would have a utility function of the separable form $U_f^*\left[U_f\left(x_f\right), U_c\left(x_c\right), U_m\left(x_m\right)\right]$ rather than being some more general function of x_f, x_m, and x_c.

Following BCL, assume that the household has economies of scale to consumption (that is, sharing and jointness or consumption) of a Gorman (1976) linear technologies type. The idea is that a bundle of purchased goods given by the K vector of purchased quantites z_s is converted by a linear K by K matrix A_s into a weakly larger (in magnitude of each element) bundle of 'private good equivalents' x, which is then divided among the household members, so $x = x_f + x_m + x_c$. Specifically, there is assumed to exist a K by K matrix A_s such that $x_f + x_m + x_c = x = A_s^{-1} z_s$. This "consumption technology" allows for much more general models of sharing and jointness of consumption than the usual collective model that categorizes goods only as purely private or purely public.

For example, suppose that a married couple without children 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 couple would have had to purchase 50% more gasoline to have each member travel the same distance as before. In this example, we would have $x^k = (3/2)\, z^k$ for k being gasoline, so the k'th row of A would consist of 2/3 in the k'th column and zeros elsewhere.This 2/3 can be interpreted as the degree of "publicness" of good k within the household. A purely private good k would have 1 instead of $2/3$, and the greater is the degree to which it is shared, the further below one is this value. Nonzero off diagonal elements of A_s may arise when the extent to which one good is shared depends upon other goods, e.g., if leisure time is a consumption good, then the degree to which auto use is shared may depend on the time involved, and vice versa.

BCL assume the household is pareto efficient in its allocation of goods, and does not suffer from money illusion. This implies the existence of a monotonically increasing function $\widetilde{U}_s$ such that a household of type s buys the bundle of goods z_s given by

$$\max_{x_f, x_m, x_c, z_s} \widetilde{U}_s \left[U_f \left(x_f \right), U_m \left(x_m \right), U_c \left(x_c \right), p/y \right] \quad \text{such that} \quad z_s = A_s \left[x_f + x_m + x_c \right] \text{ and } y = z_s' p \tag{1}$$

Solving the household's maximization problem, equation (1) yields the bundles x_t of "private good equivalents" that each household member of type t consumes within the households. Pricing these vectors at within household shadow prices $A_s' p$ (which differ from market prices because of the joint consumption of goods within the household) yields the fraction of the household's total resources that are devoted to each household member.

Let η_{ts} denote the *resource share*, defined as fraction of the household's total expenditure consumed by a person of type t in a household with s children. This resource share has a one-to-one correspondence with the "pareto-weight", defined as the marginal response of $\widetilde{U}_s$ to U_t.

In this paper, we lean heavily on existence of private assignable goods for identification of resource shares. A *private good* for our purposes is defined as one where its corresponding diagonal element of A is equal to 1 and all off-diagonal elements in that row or column are equal to 0. This means that private goods are goods that do not have any economies of scale in consumption. For example, food is private to the extent that any unit consumed by one person cannot also be eaten by another (although there could be some economies of scale in reduced waste associated with preparation of larger quantities). A private good is *assignable* if it is consumed exclusively by one known household member. So, e.g., a sandwich would be assignable if we could observe who ate it. In our application we observe separate expenditures on men's, women's, and children's clothing, which we take to be private and assignable.

Suppose there exists a private assignable good for a person of type t. This good is not jointly consumed, and so appears only in the utility function U_t, not in the utility functions of any other type of household member. Let $W_{ts}(y, p)$ be the share of total expenditures y that is spent by a household with s children on the type t private good. For example $W_{cs}(y, p)$ could be the fraction of y that the household spends on toys or children's clothes. Also let $w_t(y, p)$ be the share of y that would be spent buying the type t private good by a (hypothetical) individual that maximized $U_t(x_t)$ subject to the budget constraint $p'x_t = y$. Unlike in BCL, these individual demand functions need not be observable.

While the demand functions for goods that are not private are more complicated (see the Appendix for derivations and details, especially equation 13), the household demand functions for private assignable goods, derived from equation (1), have the simple forms

$$\begin{aligned} W_{cs}(y, p) &= s\eta_{cs}(y, p)\, w_c\left(\eta_{cs}(y, p)\, y, A'_s p\right) \\ W_{ms}(y, p) &= \eta_{ms}(y, p)\, w_m\left(\eta_{ms}(y, p)\, y, A'_s p\right) \\ W_{fs}(y, p) &= \eta_{fs}(y, p)\, w_f\left(\eta_{fs}(y, p)\, y, A'_s p\right) \end{aligned} \tag{2}$$

This solution to BCL for the case of private assignables states that the household's budget share for a person's private assignable good is equal to her resource share multiplied by the budget share she would choose herself if facing her personal shadow budget constraint. Household demand functions W_{ts}, the left side of equation (2), are in principle observable by measuring the consumption patterns of households with various y facing various p regimes. Our goal is identification of features of the right side of equation (2), in particular η_{cs}, and moreover we wish to obtain identification using only data from a single price regime.

The way BCL obtain identification is that they assume $s = 0$ and they assume that the demand functions w_m, and w_f are observable via the demands of single men and women. This amounts to assuming that the demand functions for single men and women are identical to those of men and women living in collective households, with the only differences coming from the ability of couples to share goods. This corresponds to the assumption that, up to a monotonic transformation, the utility functions $U_f\left(x_f\right)$ and $U_m(x_m)$ apply to both single and married women and men (however, given this assumption BCL do not require the existence of private, assignable goods for identification). These assumptions yield nonparametric identification of η_{ts} and A_s, which together completely define the shadow budget constraint faced by each person, and are thus sufficient for consumer-surplus type calculations.

Two problems prevent us from using the BCL identification strategy in our setting with children. First, unlike what we can do with adults, we cannot observe the demand functions for children living alone. Moreover, the assumption that single and married individuals have the same underlying utility function is questionable, so we drop that assumption and replace it with the milder assumption that parents (and individual children) have utility functions over goods that do not depend on whether the number of children in the household is one, two, or three. Our formal assumption is actually even less restictive; only certain features of utility functions at low expenditure levels are assumed to not vary with the number of children. See Theorem 2 in the Appendix for details. Theorem 1 provides an alternative identification strategy that assumes some similar feature of demands across people within a household, rather than across the same people in different household sizes.

Note that the consumption technology (i.e., the degree to which goods are jointly consumed) and the resource shares do vary with the number of children. In terms of equation (2), the assumption is that (some features of) the functions w_t for $t = c, f, m$ do not depend on s, though the values they are evaluated at, $\eta_{ts}y$ and $A'_s p$, do vary with s.

A second problem with BCL is that identification of the household consumption technology A_s requires observable price variation and the measurement of price responses in household demand functions. The measurement of price responses in demand is typically difficult for at least two reasons: first, the rationality restrictions of Slutsky symmetry and homogeneity typically require that price effects enter demand functions in complicated nonlinear ways; and second, there is often not much observed price variation in our data, so estimated price responses can be very imprecise. Indeed, many real-world data sources which have information on household consumption of commodities have no information at all on the prices of those commodities.

We get around these two problems in two steps. First, we restrict the resource share functions η_{fs} to be independent of household expenditures y, at least at low expenditure levels (though they may depend arbitrarily on p). This restriction has real bite, but one can certainly write down parametric Bergson-Samuelson household welfare functions over parametric utility functions whose resulting resource shares satisfy this restriction (we present a class of such models in the Appendix). Similar to Lewbel and Pendakur (2009), this restriction allows us to recast the BCL model into an Engel-curve framework where price variation is not exploited for identification. Second, we invoke some restrictions on the shapes of individual Engel curves. These restrictions allow us to identify individual resource shares by comparing household demands for private assignables across people within households, or by comparing these demands across households for a given type of person.

3 Identification of Children's Resource Shares Using Engel Curves

In this section, we offer a brief nontechnical description of how we achieve identification of each person's resource share in the collective household, using only data on Engel curves for private assignable goods in households with children. Technical discussion and formal identification proofs are deferred to the Appendix.

An Engel curve is defined as the functional relationship between a budget share and total expenditure, holding prices constant. In a slight abuse of notation, we may write the BCL solutions for private assignables given by equation (2) in Engel curve form as

$$\begin{aligned} W_{cs}(y) &= s\eta_{cs} w_{cs}\left(\eta_{cs} y\right) \\ W_{ms}(y) &= \eta_{ms} w_{ms}\left(\eta_{ms} y\right) \\ W_{fs}(y) &= \eta_{fs} w_{fs}\left(\eta_{fs} y\right). \end{aligned} \tag{3}$$

Here, the *Engel curve* function w_{ts} gives the demand function for person t when facing the price vector $A_s' p$ for one particular value of p, so that, e.g., $w_{cs}\left(\eta_{cs} y\right) = w_c\left(\eta_{cs}(p) y, A_s' p\right)$ for that one value of p. The resource share η_{ts} does not depend on y by assumption, and its dependence on p is suppressed in the Engel curve $w_{cs}\left(\eta_{cs} y\right)$ because prices are held constant.

3.1 Identification by Shape Invariance of Private Assignable Goods

We propose two distinct methods for obtaining econometric identification of the resource shares η_{ts}. This subsection summarizes the first method, which is similar to the shape-invariance restriction of Pendakur (1999), except that we apply it only to the Engel curves for the private assignable goods and we apply it only at low expenditure levels. We restrict how preferences for the private assignable goods vary across people, so we consider the same good for all people. For example, the private assignable good could be clothing, so that the demand function $w_t(y, p)$ gives person t's (unobserved) budget-share function for clothing when facing the constraint defined by y, p. Then, we invoke restrictions on the preferences of individuals that cause the private assignable goods to satisfy shape-invariance for real expenditure levels below a threshold $y^*(p)$:

$$w_t(y, p) = d_t(p) + g\left(\frac{y}{G_t(p)}, p\right) \quad \text{for } y \leq y^*(p). \tag{4}$$

The phrase "shape-invariance" applies here because the budget share functions for all people have the same shape (given by the function g), except for the person-specific additive term $\delta_t\,(p)$ and the person-specific income deflator $G_t\,(p)$. We will use the acronym SI to refer to shape-invariance applied to the private assignable good at low expenditure levels.

Pendakur (1999) shows that if people have costs that differ only by (price-dependent) multiplicative equivalence scales, then budget share functions must satisfy shape-invariance for all goods and at all expenditure levels. A vast amount of empirical consumer demand analysis imposes this restriction on all goods at all expenditure levels. See, e.g., Blundell, Duncan, and Pendakur (1998), Blundell, Chen, and Kristensen (2007), and Lewbel (2010). In contrast, we only assume shape invariance for a single good and only at expenditure levels below a threshold $y^*(p)$.

Substituting this restriction into (3) we get, for $y \le y^*$,

$$\begin{aligned} W_{cs}\,(y) &= s\eta_{cs}\delta_{cs} + s\eta_{cs}\gamma_s\left(\frac{\eta_{cs}y}{\Gamma_{cs}}\right), \\ W_{ms}\,(y) &= \eta_{ms}\delta_{ms} + \eta_{ms}\gamma_s\left(\frac{\eta_{ms}y}{\Gamma_{ms}}\right), \\ W_{fs}\,(y) &= \eta_{fs}\delta_{fs} + \eta_{fs}\gamma_s\left(\frac{\eta_{fs}y}{\Gamma_{fs}}\right), \end{aligned}$$

where $\delta_{ts} = d_t(A_s'p)$, $\gamma_s(y) = g(y, A_s'p)$ and $\Gamma_{ts} = G_t(A_s'p)$. The key here is that g does not vary across people. All these functions are evaluated at the same shadow price vector $A_s'p$, and as a result the function γ_s does not vary across people either (it does not have a t subscript). We show in the Appendix that if the function g has sufficient nonlinearity, then the resource shares η_{ts} are identified from the Engel curve functions $W_{fs}\,(y)$ for any household size s.

A simple example (which we will use in our empirical work below) shows how this identification works. Suppose that each person has preferences over goods given by a PIGLOG (see the Appendix and Muellbauer 1979) indirect utility function, which has the form $V_t(p, y) = b_t(p)\left[\ln y - \ln a_t(p)\right]$. An example is the popular Almost Ideal demand system (Deaton and Muelbauer 1980). With PIGLOG preferences, a sufficient (but stronger than necessary) restriction for shape-invariance of the assignable goods is $b_t(p) = b(p)$.

By Roy's identity, corresponding budget share functions for each person's private assignable are then given by

$$w_t(y, p) = d_t\,(p) + \beta(p)\ln y,$$

where d_t is a function of $a_t(p)$ and $b(p)$, and $\beta(p)$ is minus the price elasticity of $b(p)$ with respect to the price of the private assignable good.

Plugging these budget share functions into (3) yields

$$\begin{aligned} W_{cs}\,(y) &= s\eta_{cs}\left(\delta_{cs} + \beta_s \ln \eta_{cs}\right) + s\eta_{cs}\beta_s \ln y, \\ W_{ms}\,(y) &= \eta_{ms}\left(\delta_{ms} + \beta_s \ln \eta_{ms}\right) + \eta_{ms}\beta_s \ln y, \\ W_{fs}\,(y) &= \eta_{fs}\left(\delta_{fs} + \beta_s \ln \eta_{fs}\right) + \eta_{fs}\beta_s \ln y, \end{aligned} \tag{5}$$

for any household size s, and where $\delta_{ts} = d_t\left(A_s'p\right)$ and $\beta_s = \beta(A_s'p)$. These three household Engel curves are linear in $\ln y$, with slopes that can be identified by linear regressions of the household budget shares W_{ts} on a constant and on $\ln y$. The slopes of these three Engel curves are proportional to the unknown resource shares η_{ts}, and the constant of proportionality is identified by the fact that resource shares must sum to one. Equivalently, we have four equations (three Engel curves and resource shares summing to one) in four unknowns (three resource shares and the preference parameter β_s. Consequently, resource shares are exactly identified from a single household's Engel curves for the private assignable good for each of its three members.

With more complex Engel curves for private assignable goods, identification is achieved by taking higher-order derivatives of the household Engel curves with respect to y or $\ln y$, but the spirit of the identification is the same. By assuming that individuals have budget share functions for their private goods that have the same shape across people for a given price vector, we are able to compare the shape of household Engel curves across people when they face the common within-household shadow price vector. Formal identification theorems are provided in the Appendix. The Appendix also provides more details regarding the construction of PIGLOG preference models and household models that are consistent with all of our assumptions, including, e.g., that resource shares be independent of y.

3.2 Identification by Independent of Base Scale Economies of Private Assignable Goods

Our second, alternative shape restriction for identifying resource shares invokes comparability across household types (equivalently, across differing shadow-price vectors) for a given person, rather than across persons for a given household type. In particular, here we assume that cross-price effects load onto an expenditure deflator for the shadow-price vectors associated with households with one, two, or three children.

Let $p = \left[p_m, p_f, p_c, \overline{p}, \widetilde{p}\right]$ where $\overline{p}$ is the subvector of p corresponding to purely private goods other than the assigned private goods, and $\widetilde{p}$ is the subvector of p corresponding to all the other goods. Note that $\overline{p}$ includes goods like food that are private but may not be assignable. Let L be the total number of private goods. The matrix A_s is block-diagonal, with an upper left block $\overline{A}_s = I_L$ and a lower-right block $\widetilde{A}_s$ which is unspecified. For private goods, the corresponding elements of $A_s p$ are $A_t p_t = p_t$ and $\overline{A}_s \overline{p} = \overline{p}$, i.e., by definition the shadow prices of private goods equal their market prices. The shadow price of non-private goods is $\widetilde{A}_s \widetilde{p}$. Thus, for private goods, the differences in a person's budget shares across household sizes is driven by two factors: their resource share, and their cross-price demand responses.

Now we invoke the restriction that the private assignable goods have cross-price effects that load onto an income deflator, for real expenditure levels below a threshold $y^*(p)$:

$$w_t(y, p) = g_t\left(\frac{y}{G_t(\widetilde{p})}, p_t, \overline{p}\right) \text{ for } y \leq y^*(p). \tag{6}$$

This restriction is similar to the "Independent of Base Scale Economies" (IBSE) restriction invoked by Lewbel and Pendakur (2009), because the scale-economies associated with non-private goods load onto the expenditure deflator $G_t(\widetilde{p})$ which is independent of the expenditure level (i.e., independent of the base). Lewbel and Pendakur (2009) apply this restriction to all goods and all prices and at all expenditure levels. Here, we only apply a tiny part of that IBSE restriction: we apply it only to the cross-price effects of non-private goods on the private assignable good, and we apply it only at low expenditure levels.

Substituting this restriction into (3), we get

$$\begin{aligned} W_{cs}(y) &= s\eta_{cs}\gamma_c\left(\frac{s\eta_{cs}y}{\Gamma_{cs}}\right) \\ W_{ms}(y) &= \eta_{ms}\gamma_m\left(\frac{\eta_{ms}y}{\Gamma_{ms}}\right), \\ W_{fs}(y) &= \eta_{fs}\gamma_f\left(\frac{\eta_{fs}y}{\Gamma_{fs}}\right), \end{aligned} \tag{7}$$

where $\gamma_t(y) = g_t(y, p_t, \overline{p})$ and $\Gamma_{ts} = G_t\left(\widetilde{A}_s'\widetilde{p}\right)$. The key here is that the functions g_t, and therefore $\gamma_t(y)$, do not depend on household size s. We show in the Appendix that as long as $\gamma_t(0) \neq 0$ and that there is sufficient variation in resource shares across individuals and household sizes, then the resource shares η_{ts} are identified from the Engel curve functions $W_{fs}(y)$ for any three household sizes.

To illustrate, suppose again that each person has PIGLOG preferences over goods, so the indirect utility is given by $V_t(p, y) = b_t(p)\left[\ln y - \ln a_t(p)\right]$. IBSE for the cross-price effects of non-private goods is achieved via the restriction that $b_t(p) = \overline{b}_t(\overline{p}/p_t)$ and $a_t(p) = \overline{a}_t(\widetilde{p})$, so $\overline{b}_t$ is some function of private good prices and $\overline{a}_t$ is some function of the prices of other goods. By Roy's identity, corresponding budget share functions for each person's private assignable good are given by

$$w_t(y, p) = \beta_t(\overline{p}/p_t)\left[\ln y - \overline{a}_t(\widetilde{p})\right],$$

where $\beta_t(\overline{p}/p_t)$ is minus the own-price elasticity of $\overline{b}_t(\overline{p}/p_t)$. Plugging these budget share functions into (3) yields

$$\begin{aligned} W_{cs}(y) &= s\eta_{cs}\left(\delta_{cs} + \beta_c \ln \eta_{cs}\right) + s\eta_{cs}\beta_c \ln y, \\ W_{ms}(y) &= \eta_{ms}\left(\delta_{ms} + \beta_m \ln \eta_{ms}\right) + \eta_{ms}\beta_m \ln y, \\ W_{fs}(y) &= \eta_{fs}\left(\delta_{fs} + \beta_f \ln \eta_{fs}\right) + \eta_{fs}\beta_f \ln y, \end{aligned} \tag{8}$$

where $\delta_{ts} = -\beta_t(\overline{p}/p_t)\overline{a}_t(\widetilde{A}'_s\widetilde{p})$ and $\beta_t = \beta_t(\overline{p}/p_t)$ These Engel curves are linear in $\ln y$, with slopes that vary across household size s for any person t. The coefficient of $\ln y$ for person t in a household with s children, which can be identified by linearly regressing W_{ts} on a constant and on $\ln y$, is $\eta_{ts}\beta_t$. The ratio of $\ln y$ coefficients for a person of type t in two different households equals the ratio of that person's resource shares in the two households. Given three household sizes we have a total of twelve equations (three Engel curves for each of three households, plus three sets of resource shares summing to one) in twelve unknowns (three sets of three resource shares, plus three β_t parameters), so the order condition for identification is satisfied. The corresponding rank condition for identification is provided in the Appendix.

3.3 Combining restrictions

Our two restrictions, (4) and (6), can be used separately for identification, or combined to strengthen the identification. Either restriction is partly testable because one can test whether or not household demands fit into the structures given by equation (4) or equation (6). Semiparametric testing may follow the lead of Pendakur (1999) or Blundell, Chen and Christensen (2003). In this paper, we briefly explore parametric testing via overidentification with more than one private assignable good per person and overidentification from having more than three household sizes.

Either restriction is compatible with very large classes of indirect utility functions, described in the Appendix, though obviously the intersection of these restrictions is smaller. Using both restrictions together may allow for more efficient estimates, if both restrictions are true.

With PIGLOG preferences, SI holds if $V_t(p, y) = b(p)\left[\ln y - \ln a_t(p)\right]$ and IBSE holds if $V_t(p, y) = \overline{b}_t(\overline{p}/p_t)\left[\ln y - \ln \overline{a}_t(\widetilde{p})\right]$, so the combination of both holds if IBSE holds with $\overline{b}_t(\overline{p}/p_t) = \overline{b}(\overline{p}/p_t)$ and if the private assignable goods all have the same price, so $p_c = p_f = p_m$. Equal prices would hold if each member is buying the same type of private assignable good, like similar clothing. Corresponding budget share functions for each person's private assignable will then be given by

$$w_t(y, p) = d_t(p) + \beta \ln y,$$

and household demands for the private assignables are then

$$\begin{aligned} W_{cs}(y) &= s\eta_{cs}\left(\delta_{cs} + \beta \ln \eta_{cs}\right) + s\eta_{cs}\beta \ln y, \\ W_{ms}(y) &= \eta_{ms}\left(\delta_{ms} + \beta \ln \eta_{ms}\right) + \eta_{ms}\beta \ln y, \\ W_{fs}(y) &= \eta_{fs}\left(\delta_{fs} + \beta \ln \eta_{fs}\right) + \eta_{fs}\beta \ln y \end{aligned} \tag{9}$$

for all household sizes s and for all persons c, m, f. Essentially, here we take the household demands (5), which may have different slopes for each household size, and impose the IBSE restriction that the shapes are the same across different household sizes.

It is important to stress that invoking either or both of our identifying restrictions, we identify the levels of the resource shares themselves, not just how they vary with distribution factors, and we identify children's resource shares, not just those of adults. These features are not provided in the existing resource sharing rule identification results (as discussed in the introduction). Both are crucially important for our policy analysis, which is to measure the relative welfare of children in households of varying composition.

Another feature of our identification results is that the associated estimators can be easy to implement. We do not require any data on prices, and we do not require a breakdown of household total expenditures into many different goods (only for some private, assignable goods). When using the PIGLOG specification for individual utility functions (which includes as a case the Almost Ideal model), the equations to be estimated are linear in the variables. In the case with exactly three sizes of households, the reduced form parameters may be obtained via OLS estimation of these equations, with the structural parameters being given by nonlinear functions of the reduced form parameters. In the case with more than three types households, the model is still linear, but there are nonlinear restrictions on the parameters that, for efficiency, should be imposed upon estimation. In either case, estimation is far less onerous, both computationally and in terms of data requirements, than other empirical collective household models such as BCL, and is more in the spirit of the econometric shortcuts offered by Lewbel and Pendakur (2008).

4 Engel Curve Estimation

4.1 Malawian Expenditure Data

In this section, we estimate Engel curve systems in an environment without price variation using the identification results provided in Theorems 1 and 2, and summarized in the previous section. The data come from the two waves of the Malawi Integrated Household Survey, conducted in 1998-1999 (IHS1) and 2004-2005 (IHS2), respectively. The Surveys were designed by the National Statistics Office of the Government of Malawi with assistance from the International Food Policy Research Institute and the World Bank in order to better understand poverty at the household level in Malawi. Both surveys include roughly 11,000 households, drawn randomly from a stratified sample of roughly 500 strata. The sampling methods, while similar, differ between survey years because information from the 1998 Census was used to reweight the strata for the IHS2. The stratified sample is intended to provide poverty indices at the district level.

In both years, enumerators were sent to individual households to collect the data. Enumerators were monitored by Field Supervisors in order to ensure that the random samples were followed and also to ensure data quality. Cash bonuses, equivalent to roughly 30 per cent of average household income in Malawi, were used as an incentive system in the IHS2 for all levels of workers.[1] Roughly 5 per cent of the original random sample in both years was resampled because dwellings were unoccupied. Only 0.4 per cent of initial respondents refused to answer the survey in the IHS2, so endogenous selection of reporters is not likely to be a problem in these data.

In each Survey, households are asked questions from a number of modules relating to health, education, employment, fertility and, crucially for us, consumption. The consumption data are rich, particularly in the IHS2. Households are asked to recall their food consumption (one week recall) and their non-

[1]Unfortunately, the documentation for the IHS1 does not indicate whether the same incentive system was used for that Survey year.

food expenditure broken into four recall categories (one week, one month, three months and one year). Consumption amounts also include the value of home produced goods and services imputed at the value of those services consumed in the market.[2]

While rich, the actual consumption items in the data (particularly food) do vary across across households in the survey because the survey is conducted over a period of months which encompasses the wet and dry agricultural seasons. In addition, the data definitions and computations varied across the survey waves. We reconciled all IHS2 data back to IHS1 data definitions. For most expenditure items, this amounted to recoding. However, in the case of food expenditures, reconciliation required the use of auxilliary price data from the World Bank.[3]

The consumption data include (in the three month recall questionnaire) household expenditures on clothing and shoes for the household head, spouse(s), boys and girls. These are our assignable goods which we construct for each household from the detailed module data. For almost all the empirical work, we use a single private assignable good for each person equal to the sum of clothing and footwear expenditures for that person. As distribution and demographic factors, we use information from the remaining modules to construct measures of education, age, marital status, etc. We use the original survey questionnaire responses to recode and standardize the distribution factors to be consistent across surveys.

Our sample consists of 6327 households comprised of married couples with 1-4 children aged less than 15, These households (drawn from the database of approximately 20,000 households) satisfy the following additional sample restrictions: (1) polygamous marriages are excluded; (2) observations with any missing data on the age or education of members are excluded; (3) households with children aged 15 or over are excluded; and (4) households with any member over 65 are excluded. Our private assignable good is the sum of clothing and footwear expenditures. Table 3 gives summary statistics for our sample.

[2] In the IHS1, a diary method of one week recall of expenditure was also conducted although this data has since been purged as unreliable.

[3] There are differences in expenditure items across survey years which complicate the construction of comparable measures of total expenditure across survey years. In particular, the IHS1 does not include a large part of market food expenditure which, in the IHS2, is a significant component of food expenditures. As well, some significant non-food, durable, expenditures are not included. We therefore use an estimated value of total expenditure in IHS1, constructed by the World Bank for use with these data. They estimated total annual expenditure by household for the IHS2 data, using local data on prices, and, as part of their poverty alleviation research, constructed a total expenditure variable (scaled to 2004 prices) retrospectively for the IHS1 data. (The World Bank imputed total expenditure for the IHS1 using Povmap.) We constructed the equivalent total expenditure for the IHS2 data from the micro data using similar methods to those employed by the World Bank (but without replacing apparent outliers with imputed values). While this is obviously not ideal, the distribution of (real) total expenditure across surveys nevertheless appears similar. So, in our exercise, we use the total expenditure variable provided by the World Bank for IHS1 and our estimate of that variable for IHS2. For the assignable goods, we use reported clothing and footwear expenditures in IHS2, and rescale these assignable goods in the IHS1 to be comparable with 2004 nominals using Malawi's national overall inflation rate (goods-specific price increases are not available). This latter scaling is innocuous, since we include year dummies as covariates in our structural models (described below).

Table 1: Data Means, Malawian micro-data						
			couples with			
		1 child	2 children	3 children	4 children	
Number of Observations		2062	1914	1414	937	
clothing share	women	1.53	1.49	1.30	1.18	
(in per cent)	men	1.13	1.09	1.00	0.73	
	children	0.75	1.05	1.20	1.43	
footwear share	women	0.24	0.20	0.17	0.13	
(in per cent)	men	0.29	0.27	0.24	0.26	
	children	0.08	0.15	0.16	0.15	
log-total-expenditure		-0.236	-0.086	0.023	0.077	0.143

Because the Malawian data are very rich, we also include some demographic variables, which affect preferences and possibly resource shares, and some distribution factors, which affect only resource shares. Our theorems show identification for models without these variables, so one can apply the theorem conditionally on values of these additional variables to prove identification with them included. As in Browning and Chiappori (1998) the presence of distribution factors may help identification of resource shares, but, unlike Browning and Chiappori (1998) (and most other empirical collective household models), we do not require them for identification.

We include 7 demographic variables: region of residence (urban, non-urban North, non-urban Central and non-urban South, with urban as the left-out category); a dummy for the IHS2 wave; the average age of children less 5; the minimum age of children less 5; and the proportion of children who are girls. We include 2 distribution factors: the difference in age between husband and wife divided by 10; and the difference in years of education between husband and wife divided by 5. We allow all demographic factors to affect the preferences of every household member, and we allow resource shares to depend on all demographic factors and both distribution factors.

We estimate models corresponding to individuals with Almost Ideal indirect utility functions and their resulting log-linear Engel curves. Household budget share equations are given by

$$\begin{aligned} W_{cs}(y) &= s\eta_{cs}\left[\delta_{cs} + \left(\ln\eta_{cs} + \ln y\right)\beta_{cs}\right], \\ W_{ms}(y) &= \eta_{ms}\left[\delta_{ms} + \left(\ln\eta_{ms} + \ln y\right)\beta_{ms}\right], \\ W_{fs}(y) &= \eta_{fs}\left[\delta_{fs} + \left(\ln\eta_{fs} + \ln y\right)\beta_{fs}\right]. \end{aligned} \tag{10}$$

Implementation requires imposition of one or both of our identification restrictions. We impose $\beta_{ts} = \beta_s$ for all t to satisfy SI as in equation (5) or we impose $\beta_{ts} = \beta_t$ for all s to satisfy IBSE as in equation (8). Both conditions are satisfied if $\beta_{ts} = \beta$ for all t, s.

4.2 Results

Let a be a vector of 4 dummy variables for the 4 household types (indexed by s), let d indicate the 2 distribution factors (relative education and relative age of husband and wife), let z_1 indicate the 3 area-of-residence dummy variables plus the dummy for the 2002 wave, and let z_2 indicate the 3 demographic variables describing the age and sex of children. By definition, distribution factors d are assumed to affect resource shares but not preferences, while the parameters can affect both preferences and resource shares. For each person t, the resource shares η_{ts} are specified as linear in a, d, z_1, and z_2, and the intercept preference parameters δ_{ts} are specified as linear in a, z_1 and z_2. The slope preference parameters β_{ts} are specified as linear in z_1, separately for each household type, or separately for each person (depending on the identifying restriction). Obviously, other arrangements for covariates are possible, and estimates from models with more and less covariates in each of η_{ts}, δ_{ts} and β_{ts} are available on request from the authors.

We implement the model by adding an error term to each equation of (10). These errors may covary across equations, so we estimate the model via nonlinear Seemingly Unrelated Regression. All estimates presented are for models using the sum of clothing and footwear expenditures for each person as the private assignable good. Asympotic standard errors are robust to heteroskedasticity of unknown form, and are given in *italics*.

We present estimates for η_{ts}. Estimates of other parameters are not presented, but they are available on request. Our identifying restrictions all concern β_{ts}. When imposing SI, we have β_{ts} linear in a constant and the 4 elements of z_1 for each household size s (20 parameters in total). When imposing IBSE, we have β_{ts} linear in a constant and the 4 elements of z_1 for each person t (15 parameters in total). Finally, when imposing both SI and IBSE, we have β_{ts} linear in a constant and the 4 elements of z_1 (5 parameters in total). We note that all estimated values of the constant terms in β_{ts} are statistically significantly different from zero (nonzero latent slopes are required for identification for all models).

Table 2 presents our estimated resource share parameters from the Malawian data. The leftmost block gives estimates using the SI restriction, the middle block gives estimates using the IBSE restriction, and the rightmost block imposes both restrictions. We report only coefficients relating to the resource shares η_{ts}, and only report coefficients relating to levels of resource shares and the effects of z_2 and d on resource shares. Parameters related to children's resource shares are computed off of the estimated values for adult resource share parameters.

Define a reference household as one in which $d = z_1 = z_2 = 0$, which is the case for urban households in the 1998/99 wave in which the man and woman have identical age and education, and the children are all boys aged 5 (so that the average and minimum are both 5). For a reference household, the resource share is given by the constant term in η_{ts}. In the Table, we report the level of the resource share for the man η_{ms}, woman η_{fs}, all children $s\eta_{cs}$, and each child η_{cs}.

Table 2: Estimates from Malawian Clothing Budget Shares

		SI		IBSE		SI&IBSE	
		Estimate	*Std Err*	Estimate	*Std Err*	Estimate	*Std Err*
one child	man	0.395	*0.036*	0.433	*0.150*	0.396	*0.035*
	woman	0.360	*0.035*	0.445	*0.172*	0.365	*0.035*
	children	0.245	*0.029*	0.122	*0.075*	0.239	*0.028*
	each child	0.245	*0.029*	0.122	*0.075*	0.239	*0.028*
two children	man	0.352	*0.035*	0.430	*0.138*	0.359	*0.036*
	woman	0.295	*0.031*	0.395	*0.159*	0.300	*0.032*
	children	0.353	*0.036*	0.176	*0.107*	0.341	*0.036*
	each child	0.176	*0.018*	0.088	*0.054*	0.170	*0.018*
three children	man	0.397	*0.042*	0.477	*0.140*	0.405	*0.041*
	woman	0.249	*0.031*	0.344	*0.156*	0.256	*0.031*
	children	0.353	*0.043*	0.180	*0.109*	0.340	*0.042*
	each child	0.118	*0.014*	0.060	*0.036*	0.113	*0.014*
four children	man	0.243	*0.045*	0.372	*0.140*	0.253	*0.047*
	woman	0.266	*0.035*	0.389	*0.166*	0.282	*0.036*
	children	0.491	*0.053*	0.239	*0.143*	0.465	*0.054*
	each child	0.123	*0.013*	0.060	*0.036*	0.116	*0.014*
min. age	man	-0.004	*0.004*	-0.004	*0.004*	-0.004	*0.004*
of children	woman	0.002	*0.003*	0.003	*0.003*	0.003	*0.003*
	children	0.002	*0.004*	0.001	*0.002*	0.001	*0.004*
avg. age	man	0.009	*0.003*	0.006	*0.004*	0.009	*0.003*
of children	woman	0.001	*0.002*	-0.001	*0.003*	0.001	*0.002*
	children	-0.010	*0.003*	-0.006	*0.004*	-0.010	*0.004*
proportion	man	0.055	*0.018*	0.044	*0.023*	0.058	*0.018*
girl children	woman	0.005	*0.014*	-0.005	*0.017*	0.006	*0.014*
	children	-0.059	*0.023*	-0.039	*0.027*	-0.064	*0.023*
man age-	man	-0.003	*0.002*	-0.003	*0.002*	-0.003	*0.001*
woman age	woman	0.003	*0.001*	0.004	*0.002*	0.003	*0.001*
	children	0.000	*0.001*	0.000	*0.000*	0.000	*0.001*
man educ-	man	-0.003	*0.005*	-0.003	*0.005*	-0.003	*0.005*
woman educ	woman	0.000	*0.004*	0.001	*0.004*	0.001	*0.004*
	children	0.003	*0.003*	0.002	*0.002*	0.003	*0.003*

Consider first the leftmost block which presents the estimates given the SI restriction. Looking at the coefficients giving the level of resource shares in reference households of different sizes, we see that, roughly speaking, if there are more children, the resources of adults decline so that children can have greater resources. A reference household with one child directs 24.5 per cent of its expenditure to children's consumption. With two children, this share rises to 35.3 per cent, and four children, to 49.1 per cent. The (weak) monotonicity of children's resources with respect to the number of children is consistent with our expectations. However, the resource share of *each* child declines with the number of children. But, this resource share is still about 12 per cent for each child, even if there are three or four children in the household.

Although the total resources of parents roughly decline with the number of children, this is not spread evenly across the male and female. Men absorb between 35 per cent and 40 per cent of household resources if there are 3 or less children. Given the standard errors, this is a relatively small amount of variation. In contrast, women see their resource shares drop by about 10 percentage points as the number of children goes from 1 to 3. Considering households with 4 children, we see a large drop in men's resource shares.

The coefficients corresponding to the minimum age of the children do not have individually statistically significant impacts, but those relating to their average age have a noticeable impact. If the average age of the children is 1 year older, the children together get resource share about 1 percentage point smaller, with these resources diverted to the man. The coefficients relating to the proportion of girl children also seem potentially important. In particular, if the children are all girls, the man's resource share is about 5 percentage points higher than if the children are all boys, and this increased resource share comes from a decrease in childrens' resources only (the woman's resource share is unaffected). Thus, unlike Deaton (1989, 1997) but similar to Rose (1999), we find statistically significant evidence of gender discrimination in consumption within the household. One difference between our finding and that of Rose (1999), is that we find that gender discrimination is the *status quo* and does not arise only in response to household income shocks.

The coefficients relating to differences the age and education of the man and the woman (ie., the distribution factors) do not suggest that these factors are very important. The estimated magnitudes are small and individually statistically insignificant. We emphasize that our findings here relate to possible bargaining over household resources conditional on being a household with children. Thus, it may be that household bargaining power (which may relate to the age and education differences) is determined or exercised at the point of household formation or over the determination of number of children. The insignificant results, particularly for education, may also arise because of the coarse nature of our education data.

Now, we briefly consider testing the SI restrictions. If there was more than one private assignable good for each person, indicated with a superscript k, we would have for AI individual preferences under the SI restriction:

$$\begin{aligned} W_{cs}^{k}(y) &= s\eta_{cs}\left(\delta_{cs}^{k}+\beta^{k}\ln s\eta_{cs}\right)+s\eta_{cs}\beta^{k}\ln y, \\ W_{ms}^{k}(y) &= \eta_{ms}\left(\delta_{ms}^{k}+\beta^{k}\ln \eta_{ms}\right)+\eta_{ms}\beta^{k}\ln y, \\ W_{fs}^{k}(y) &= \eta_{fs}\left(\delta_{fs}^{k}+\beta^{k}\ln \eta_{fs}\right)+\eta_{fs}\beta^{k}\ln y, \end{aligned} \tag{11}$$

for each assignable good k for each household size s. The important thing here is that although β^k varies across goods, the resource shares η_{ts} do not vary across goods. Thus, instead of having resource shares which are exactly identified by looking across people for a single private assignable good, the resource shares are overidentified in the presence of multiple private assignable goods. In particular, with 2 assignable goods and 3 people, we have 6 slopes revealed by the data for a given household size, and only 4 unobserved parameters (2 latent slopes β^k and 2 resource shares after taking into account that resource shares sum ton one).

A simple test of the SI restriction (under the PIGLOG model) is given by allowing η_{ts} to vary with the good, and testing whether or not this additional variation is necessary to explain behaviour. We implement this test by separating clothing and footwear expenditures into two private assignable goods, each of which has a budget-share equation for each person in the household. Then, we estimate the model under SI but with 2 extra parameters for each of the 4 household sizes, corresponding to differences in resource shares across the goods. The sample value of the test statistic for the exclusion of these 8 parameters is 13.3, and is distributed as a χ_8^2 with a p-value of 10.2 per cent under the null hypothesis that SI is true. So, we take this as evidence that SI is a tolerable restriction in this context.

The middle panel of Table 2 uses the IBSE restriction to identify resource shares. Here, we see that the estimated levels of resource shares are much less precisely estimated than in the SI case, with standard errors that are two to five times as large. However, the basic themes of estimated levels given SI are still evident: men's resources do not decline much with the number of children when there are $1, 2$ or 3 children; women's resources decline over the first 3 children; childrens' resources increase in with the number of children; and the resource share of each child (weakly) declines with the number of children. One difference stands out between the SI and IBSE estimates: given IBSE, the resource shares of each child are only about 6 per cent in households with three or four children.

Turning to the covariates, the estimates given IBSE are about as precise as those given SI, and the patterns are much the same. The only statistically significant covariate is the proportion of girl children, which is associated with higher resource shares for men and lower resource shares for children.

We can also test IBSE against a less restrictive alternative. Recall that given IBSE, we have:

$$W_{ts}(y) = \eta_{ts}\left(\delta_{ts} + \beta_t \ln \eta_{ts}\right) + \eta_{ts}\beta_t \ln y,$$

for men and women, and analogously for children. This model is exactly identified with 3 household sizes, and overidentified with more than 3 household sizes. In Table 2, we use 4 household sizes ($1-4$ children). In this case, there are 12 slopes revealed by the data and only 11 unobserved parameters—8 resource shares and 3 latent slope parameters (1 for each person). So, the resource shares given the IBSE restriction are overidentified with 1 degree of freedom. We test this overidentification restriction by testing the relevance of an extra parameter which allows the resource share of 1 person in 1 household type to deviate from the model. The sample value of the z-test on the hypothesis that this parameter is irrelevant is -1.39, so we take this as evidence that the IBSE restriction is tolerable.

We can assess the joint restriction that both SI and IBSE hold. A natural way to test this is to estimate (for a single private assignable good) under SI as in the leftmost columns, and to conduct a Wald test on the hypothesis that the β_s are same for all 4 household types. Since β_s has 20 parameters (5 for each type), this amounts to testing 15 restrictions. The sample value of this test statistic is 25.7, and it is distributed as a χ^2_{15}, yielding a p-value for the test of 4.4%. Alternatively, we can estimate under IBSE as in the middle columns, and conduct a Wald test on the hypothesis that the β_t are the same for all persons t. Since β_t has 15 parameters (5 for each person), this amounts to testing 10 restrictions. The sample value of this test statistic is 9.6, and it is distributed as a χ^2_{10}, yielding a p-value of 48%. Thus, we are more cautious on whether or not combining SI and IBSE is wise, but the joint restriction may not be ruled out by behaviour.

The rightmost columns of Table 2 present estimates given both SI and IBSE. All the same patterns emerge, and the estimates are slightly more precise. But the extra precision illuminates one additional feature. The distribution factor corresponding to the relative age of the parents may have an effect on the within-household distribution. In particular, if the man is older than the women, then resources are diverted from the man to the woman. If the man is 4 years older than the woman, then his resource share is about 1 percentage points smaller and hers is about 1 percentage point larger.

The estimates given both SI and IBSE are the most precise of the estimates presented because more identifying restrictions are imposed than with either SI or IBSE alone. These estimates allow for the sharpest testing of hypotheses about the behaviour of resource shares across household size. The patterns observed above, wherein for households with 3 or less children, men's resources are invariant to the number of children and women's resources decline with the number of children, evident in a statistical sense as well. The Wald test statistic for the hypothesis that men's resource shares are the same for households with $1, 2$ or 3 children is 5.1, and it is distributed as a χ^2_2, yielding a (two-sided) p-value of 15.6%. The Wald test statistic for the hypothesis that the woman's resource share is lower in households with 3 children than it is in households with 1 child is 7.5, and it is distributed as a χ^2_1, yielding a (one-sided) p-value of 0.8%. This pattern does not hold over the comparison between households with 3 and 4 children: both men and women sacrifice in terms of their resource shares with the addition of a fourth child.

The discussion above related to the levels of resource shares for persons in reference households, and to the marginal effects of various demographic and distribution factors. However, this does not tell us how resource shares change in aggregate across household sizes because the demographic and distribution factors covary with household size. To evaluate, for example, whether men or women make the larger sacrifice of consumption for their children, it is illustrative to consider their average resource shares in households of different sizes, averaging over all the values of demographic and distribution factors observed in the population.

The leftmost columns of Table 3 presents summary statistics on the estimated values of resource shares for people in households of different sizes. It is comforting to see that the minima and maxima of estimated resource shares do not go outside 0, 1 for any person in any household in the sample. Indeed, the standard deviations are quite small in most cases. Interestingly, the standard deviations of resource shares are larger for men than for women in all household sizes. Thus, the covariates are not very important in terms of their effects on the resource shares, though they do induce more variation for men than for women. Much more important than these factors are the household sizes themselves. This suggests that our ability to identify the *level* of resource shares, rather than just their response with respect to distribution factors, is important.

Table 3: Estimated Resource Shares and Poverty Rates

		Mean	Std Dev	Min	Max	Pov Rate *Unequal*	Pov Rate *Equal*
one child	man	0.426	0.036	0.317	0.533	0.791	0.889
	woman	0.381	0.019	0.339	0.486	0.843	
	children	0.193	0.044	0.057	0.263	0.892	
	each child	0.193	0.044	0.057	0.263		
two children	man	0.401	0.028	0.300	0.490	0.753	0.925
	woman	0.318	0.018	0.271	0.461	0.872	
	children	0.281	0.035	0.169	0.358	0.939	
	each child	0.140	0.017	0.085	0.179		
three children	man	0.456	0.025	0.325	0.542	0.588	0.954
	woman	0.275	0.018	0.225	0.454	0.893	
	children	0.270	0.027	0.179	0.335	0.989	
	each child	0.090	0.009	0.060	0.112		
four children	man	0.314	0.023	0.212	0.390	0.824	0.975
	woman	0.302	0.018	0.246	0.448	0.826	
	children	0.384	0.023	0.316	0.462	0.982	
	each child	0.096	0.006	0.079	0.115		
All Households	man	0.431	0.036	0.278	0.563	0.739	0.927
	woman	0.313	0.027	0.231	0.466	0.860	
	children	0.256	0.052	0.072	0.372	0.941	
	each child	0.139	0.056	0.058	0.271		
All Persons	all	0.194	0.129	0.057	0.542	0.902	0.935

Table 3 shows that children together absorb more resources when the are more children, but that this effect is muted when all the covariates are averaged over. As in Table 2 (which does not average), we see that for households with 1 to 3 children, men's resource shares are not very much smaller in households with more children, but women's resource shares are smaller in households with more children. In addition, men tend to absorb larger shares of household resources than the estimates in Table 2 would suggest. This is because they have favourable values of the covariates, on average.

The rightmost columns of Table 3 show the estimated poverty rates (at the household level) for households of different sizes. The rightmost column uses a per-capita poverty threshold of $US2 (2004 PPP adjusted) per day. The bottom block and bottom row give the estimated poverty rate for all households together and for all persons. Here, we see a poverty rate for households of 92.7% for our sample. For comparison, the World Bank reported poverty levels for 1999 and 2004 are 93.5% and 90.5%, respectively.

Our estimated poverty rates differ from the World Bank rates in two ways. First, the World Bank rates are based on per-capita expenditure, and thus use an equal resource share for each household member given by the reciprocal of the number of household members. This means that either all or none of the members of the household are poor. In contrast, in our approach, household members do not have equal shares. Second, the World Bank uses a poverty-threshold (for per-capita expenditure) that is the same for all household members. In our approach, different household members have different resource shares, but they may also have different needs. Consequently, we use the OECD estimate of the relative needs of children (60% of that of adults) to compute a poverty threshold for children.

There are at least three features to note in our poverty estimates. First, Table 3 shows that there are a lot more households with poor women than with poor men. For example, looking at the rows for All Households, we see that 73.9 per cent of households have a poor man, but 86.0 per cent of households have a poor woman. Second, for households with 1 to 3 children, the poverty rates of men seem to drop with household size, but the poverty rate for women and children rises with household size. Third, more households have poor children than have poor adults. In households with 3 or 4 children, nearly all children are poor. Taken together, these results suggest that intra-household inequality is an important determinant of poverty.

5 Conclusions

Child poverty is at the root of much inequality. Differences in human capital and physical health (among others attributes) have been traced to poor nutrition in the early years of life. Children are also among the least able in society to care for themselves. Despite the apparent importance of understanding the intra-household dimension of child inequality, very little research has focused on childrens' share of household resources. Most collective household models either ignore children, or treat them as attributes of adults.

We propose a collective household model in which children are people with their own utility functions (possibly assigned to them by parents). Children's resource shares within the household are identified given household level Engel curve data on private assignable goods. In particular, by looking at how the budget shares for men's, women's and children's clothing and shoes vary across households with differing income levels and numbers of children, our structural model allows us to back out an estimate of the fraction of total household income that is consumed by each family member.

Using household consumption data for Malawi, we find that children command a reasonably large share of household resources and that the share of resources devoted to children rises with the number of children. Mothers appear to contribute more resources than fathers to children, and we find some evidence of gender-bias in children's resource shares.

Our model and results are applicable to policy. Our model is applicable to situations where one has data on assignable goods and total expenditure for sufficient numbers of households. Policymakers can therefore identify child poverty and the intra-household inequality of children with relatively minimal data requirements. Our results suggest that increasing household income benefits all household members and so there exists a trade-off between the costs of targeting expenditure at one household member and the benefits of household-level assistance.

6 Appendices

6.1 Appendix 1: Theorems

Let $h_t^k(p, y)$ denote the Marshallian demand function for good k associated with the utility function $U_t(x_t)$, so an individual t that chooses x_t to maximize $U_t(x_t)$ under the usual linear budget constraint $p'x_t = y$ would choose $x_t^k = h_t^k(p, y)$ for every purchased good k. Let $h_t(p, y)$ be the vector of demand functions $h_t^k(p, y)$ for all goods k, so $x_t = h_t(p, y)$ and the indirect utility function associated with $U_t(x_t)$ is then defined as the function $V_t(p, y) = U_t(h_t(p, y))$.

For their identification, BCL assumed that for a person of type t, $U_t(x_t)$ was the same as the utility function of a single person of type t living alone, and so $h_t(y, p)$ would be that single person's observed demand functions over goods. We do not make this assumption.

Rewriting equation (1) we have

$$\max_{x_f, x_m, x_c, z_s} \widetilde{U}_s\left[U_f\left(x_f\right), U_m\left(x_m\right), U_c\left(x_c\right), p/y\right] \quad \text{such that} \quad z_s = A_s\left[x_f + x_m + s x_c\right] \text{ and } y = z_s' p \tag{12}$$

The demand functions for the household s arising from the household's maximization problem, equation (12), can be written as follows. Let A_s^k denote the row vector given by the k'th row of the matrix A_s.

Define $H_s^k(p, y)$ to be the demand function for each good k in a household with s children. Then an immediate extension of BCL (the extension being inclusion of the third utility function U_c) is that the household s demand functions are given by

$$z_s^k = H_s^k(p, y) = A_s^k\left[h_f\left(A_s' p, \eta_{fs} y\right) + h_m\left(A_s' p, \eta_{ms} y\right) + s h_c\left(A_s' p, \eta_{cs} y\right)\right] \tag{13}$$

where η_{ts} denotes the resource share of a person of type t in a household with s children. In general, resource shares η_{ts} will depend on the given prices p and total household expenditures y, however, we will assume that resource shares to do not vary with y, and so for now will denote them $\eta_{ts}(p)$. The resource shares $\eta_{ts}(p)$ may depend on observable household characteristics including distribution factors, which we suppress for now to simplify notation (recall we have also suppressed dependence of all the above functions on attributes such as age that may affect preferences).

Note in equation (13) that each child gets a share $\eta_{cs}(p)$, so the total share devoted to children is $s\eta_{cs}(p)$. By definition, resource shares must sum to one, so for any s

$$\eta_{fs}(p) + \eta_{ms}(p) + s\eta_{cs}(p) = 1 \tag{14}$$

Our first assumption is that the BCL model as described above holds, that is,

ASSUMPTION A1: Equations (12), (13), and (14) hold, with resource shares $\eta_{ts}(p)$ that do not depend upon y.

BCL show generic identification of their model by assuming the demand functions of single men, single women, and married couples (that is, the functions $h_m(r)$, $h_f(r)$, and $H_0(r)$) are observable, and assuming the utility functions $U_f\left(x_f\right)$ and $U_m(x_m)$ apply to both single and married women and men. Their results cannot be immediately extended to children and applied to our application, because unlike men or women we cannot observe demand functions for children living alone. We also do not want to impose the assumption that single and married adults have the same underlying utility functions $U_f\left(x_f\right)$ and $U_m(x_m)$.

The assumption that resource shares are independent of y is also made by Lewbel and Pendakur (2009). This assumption implies joint restrictions on the preferences of household members and on the household's bargaining or social welfare function $\widetilde{U}_s$ (see, proposition 2 of Browning, Chiappori, and Lewbel 2008). To illustrate the point, we later give an example of a model satisfying all of our assumptions which has resources shares independent of y, in which the household maximizes a Bergson-Samuelson social welfare function. Note that Assumption A1 permits resource shares to vary freely with other observables that are associated with total expenditures y, such as household income or the mother's and father's wages.

Definition: A good k is a private good if, for any household size s, the matrix A_s has a one in positition k,k and has all other elements in row k and column k equal to zero.

This is equivalent to the definition of a private, assignable good in models that possess only purely private and purely public goods. With our general linear consumption technology, this definition means that the sum of the quantities of good k consumed by each household member equals the household's total purchases of good k, so the good is not consumed jointly like a pure public good, or partly shared like the automobile use example.

Definition: A good k is an assignable good if it only appears in one of the utility functions U_f, U_m, or U_c, e.g. a child good is an assignable good that is only appears in U_c, and so is only consumed by children.

ASSUMPTION A2: Assume that the demand functions include a private, assignable child good, denoted as good c, and a private, assignable good for each parent, denoted as goods m and f.

Note that we do not require a separate assignable good for each child, so good c is consumed by all children. Our identification results will only require observing the demand functions for the three private, assignable goods listed in Assumption A2. Examples of child goods could be toys or children's clothes, while examples of adult goods could be alcohol, tobacco, or men's and women's clothing. Private, assignable goods are often used in this literature to obtain identification, or to increase estimation efficiency. See, e.g., Chiappori and Ekelund (2009).

It follows immediately from Assumptions A1 and A2 that, for the private, assignable goods $k = f, m, c$, equation (13) simplifies to

$$
\begin{aligned}
z_s^k &= H_s^k(p, y) = h_k\left(A_s' p, \eta_{ks}(p)\, y\right) \quad \text{for } k \in \{m, f\} && (15)\\
\text{and } z_s^c &= H_s^c(p, y) = s h_c\left(A_s' p, \eta_{cs}(p)\, y\right) && (16)
\end{aligned}
$$

We will now make some assumptions regarding individual's utility functions, that will translate into restrictions on the demand functions for assignable goods. We will show later that these assumptions are at least partly testable.

The first set of assumptions, leading to Theorem 1, will permit identification by imposing an element of similarity across different individual's demand functions for the assignable goods within a household of any given size. A second set of assumptions, leading to Theorem 2, will yield identification by permitting a comparison of the assignable good demand functions of each household member across households of different sizes.

Let $\widetilde{p}$ denote the vector of all prices except p_m, p_f, and p_c, so $\widetilde{p}$ consists of the prices of all goods except for the three private, assignable goods in Assumption A2. We may correspondingly define a square matrix $\widetilde{A}_s$ such that the set of prices $A_s' p$ is given by p_m, p_f, p_c, and $\widetilde{A}_s' \widetilde{p}$. Let $I(\cdot)$ be the indicator function that equals one when its argument $\cdot$ is true and zero otherwise.

ASSUMPTION A3: For $t \in \{m, f, c\}$ let

$$V_t(p, y) = I\left(y \leq y^*(p)\right) \psi_t\left[\upsilon\left(\frac{y}{G_t(p)}\right) + F_t(p), \widetilde{p}\right] + I\left(y > y^*(p)\right) \Psi_t(y, p) \tag{17}$$

for some functions y^*, Ψ_t, ψ_t, υ, F, and G_t where y^* is strictly positive, G_t is nonzero, differentiable, and homogeneous of degree one, υ is differentiable and strictly monotonically increasing, $F_t(p)$ is differentiable, homogeneous of degree zero, and satisfies $\partial F_t(p)/\partial p_t = \varphi(p) \neq 0$ for some function φ. Also, ψ_t and Ψ_t are differentiable and strictly monotonically increasing in their first arguments, and differentiable and homogeneous of degree zero in their remaining (vector valued) arguments.

As we show below, Assumption A3 only restricts people's demand functions for assignable goods at very low total expenditure levels. It places no restriction at all (except for standard regularity conditions) on the demand functions for all other goods, and place no restrictions on the assignable good demand functions anywhere other than at low total expenditure levels.

In Assumption A3, $y^*(p)$ is this low but positive threshold level of total expenditures. Households having total expenditures $y > y^*(p)$ have demand functions given by an arbitrary, unconstrained indirect utility function $\Psi_t(y, p)$. Assumption A3 only requires that $\Psi_t(y, p)$ have the standard homogeneity and differentiability properties of any regular indirect utility function. Assumption A3 therefore permits individuals to have any regular preferences at all over bundles of goods that cost more than some minimal level $y^*(p)$, and therefore the demand functions for all goods can have any smooth parametric or nonparametric functional form at total expenditure levels $y > y^*(p)$.

The key restriction in Assumption A3 is that the functions υ and φ do not vary across people. The function $\upsilon(y/g_t(p)) + F_t(p)$ with $\partial F_t(p)/\partial p_t = \varphi(p)$, if it were the entire indirect utility function, would, induce shape invariance on the Engel curves of the private, assignable goods. See Pendakur (1999), Blundell, Duncan, and Pendakur (1998), Blundell, Chen, and Kristensen (2007), and Lewbel (2010). However, the demand functions that arise from equation (17) are only constrained to satisfy same invariance shape at low expenditure levels, because this restriction is only imposed for $y \leq y^*(p)$. The result of this restriction will be that the Engel curves for assignable goods can have any shape, but they will all need to have the same shape at low total expenditure levels.

Also, even at low expenditure levels, shape invariance is only imposed on the demand functions of the private, assignable goods. The role of the function ψ_t and the lack of restriction on cross derivatives $\partial F_t(p)/\partial p_k$ for all $k \neq t$ is to remove constraints on the shapes of Engel curves of goods other than the private, assignable ones.

The restriction that $\partial F_k(p)/\partial p_k$ be the same for k equal to m, f, and c limits either how $F(p)$ can depend on the prices of these goods, or on how the prices of these goods can covary. It follows from assignability that the indirect utility function for each person t will depend on p_t but not on the other two elements of the set $\{p_m, p_f, p_c\}$. Therefore, given assignability, it holds without loss of generality that $F_t(p) = \widetilde{F}_t(p_t, \widetilde{p})$ for some function $\widetilde{F}_t$ (a similar restriction must also hold for the function G_t). If the prices of the assignable goods are perfectly correlated over time, meaning they are Hicks aggregable, then $p_m = p_f = p_c$ (after appropriately rescaling units quantities are measured in if necessary) and it will follow automatically that $\partial F_k(p)/\partial p_k = \varphi(p)$ for the assignable goods k for any $F_k(p) = \widetilde{F}_k(p_k, \widetilde{p})$ function. Alternatively, if we have the functional form $F_t(p) = p_t \widetilde{\varphi}(\widetilde{p})$, then regardless of how the relative prices of the assignable goods vary, the constraint that $\partial F_k(p)/\partial p_k = \varphi(p)$ for k equal to m, f, and c will hold with $\varphi(p) = \widetilde{\varphi}(\widetilde{p})$

The role of the function ψ_t is to impose this low expenditure shape invariance only on the assignable goods, so the shapes of the Engel curves of all other goods are not restricted to be shape invariant anywhere. In short, although Assumption A4 looks complicated, it basically just says the budget share Engel curves of the household member's assignable goods all have same shape (differing only by translations) at low total expenditure levels, and are otherwise unrestricted.

To show this formally, apply Roy's identity to equation (17). The result is that, for person t and any good k, when $y > y^*(p)$, the demand function will be given by applying Roy's identity to $\Psi_t(y, p)$ giving $h_t(y, p) = -\left[\partial \Psi_t(y, p) / \partial p_k\right] / \left[\partial \Psi_t(y, p) / \partial y\right]$. However, when $y \leq y^*(p)$,applying Roy's identity to equation (17) gives

$$\begin{aligned} h_t(y, p) \quad = \quad & \frac{\psi_t'\left[v\left(\frac{y}{G_t(p)}\right) + F_t(p), \widetilde{p}\right]\left[v'\left(\frac{y}{G_t(p)}\right)\frac{y}{G_t(p)^2}\frac{\partial G_t(p)}{\partial p_k} - \frac{\partial F_t(p)}{\partial p_k}\right]}{\psi_t'\left[v\left(\frac{y}{G_t(p)}\right) + F_t(p), \widetilde{p}\right] v'\left(\frac{y}{G_t(p)}\right)\frac{1}{G_t(p)}} \\ & - \frac{\partial \psi_t\left[v\left(\frac{y}{G_t(p)}\right) + F_t(p), \widetilde{p}\right] / \partial p_k}{\psi_t'\left[v\left(\frac{y}{G_t(p)}\right) + F_t(p), \widetilde{p}\right] v'\left(\frac{y}{G_t(p)}\right)\frac{1}{G_t(p)}} \quad \text{for } y \leq y^*(p) \end{aligned}$$

Where ψ_t' and v' denote the derivatives of ψ_t and v with respect to their first elements.

For the assignable goods $k \in \{m, f, c\}$, the derivative $\partial \psi_t / \partial p_k$ is zero and $\partial F_k(p) / \partial p_k = \varphi(p)$, which makes the above demand function simplify to

$$h_k(y, p) = \frac{y}{G_k(p)} \frac{\partial G_k(p)}{\partial p_k} - \frac{\varphi(p)\frac{G_k(p)}{y}}{v'\left(\frac{y}{G_k(p)}\right)} y \quad \text{for } y \leq y^*(p) \tag{18}$$

which we can write more simply as

$$h_k(y, p) = \delta_k(p)\, y + g\left(\frac{y}{G_k(p)}, p\right) y \quad \text{for } y \leq y^*(p) \tag{19}$$

for functions δ_k and g. Substituting this into equation (15) gives household demand functions for the assignable goods

$$z_s^k = H_s^k(p, y) = \delta_k\left(A_s' p\right) \eta_{ks}(p)\, y + g\left(\frac{\eta_{ks}(p)\, y}{G_k\left(A_s' p\right)}, A_s' p\right) \eta_{ks}(p)\, y \quad \text{when } y \leq y^*(p), k \in \{m, f\}$$

and, for children

$$z_s^c = H_s^c(p, y) = \delta_c\left(A_s' p\right) s \eta_{cs}(p)\, y + g\left(\frac{\eta_{cs}(p)\, y}{G_c\left(A_s' p\right)}, A_s' p\right) s \eta_{cs}(p)\, y \quad \text{when } y \leq y^*(p).$$

Now consider Engel curves. For the given price regime p we can write the above equation more concisely as

$$\begin{aligned} z_s^k \quad &= \quad H_s^k(y) = \delta_{ks}\eta_{ks} y + g_s\left(\frac{\eta_{ks} y}{G_{ks}}\right)\eta_{ks} y \quad \text{for } y \leq y^*(p), k \in \{m, f\} \\ \text{and } z_s^c \quad &= \quad H_s^c(y) = \delta_{cs} s \eta_{cs} y + g_s\left(\frac{\eta_{cs} y}{G_{cs}}\right) s \eta_{cs} y \quad \text{for } y \leq y^*(p). \end{aligned}$$

ASSUMPTION A4: The function $g_s(y)$ is twice differentiable. Let $g_s'(y)$ and $g_s''(y)$ denote the first and second derivatives of $g_s(y)$. Either $\lim_{y \to 0} y^{\zeta} g_s''(y) / g_s'(y)$ is finite and nonzero for some constant $\zeta \neq 1$ or $g_s(y)$ is a polynomial in $\ln y$.

Polynomials in $\ln y$ can require $\zeta = 1$ to have $\lim_{y \to 0} y^{\zeta} g_s'' (y) / g_s' (y)$ be finite and nonzero, which is why Assumption A4 requires a separate statement to identify the polynomial case. The main implication of Assumption A4 is that identification requires some nonlinearity in the demand function, otherwise $g_s'' (y)$ would be zero.

For the formal proof it is easiest to have that nonlinearity be present in the neighborhood of zero as in Assumption A4, but in practice nonlinearity over other ranges of y values would generally suffice. Empirically, all points along the engel curves (or at least those below y^*) will generally contribute to the precision of estimation, not just data around zero.

A sufficient, but stronger than necessary, condition for the twice differentiability of g_s in Assumption A4 is that υ be three times differentiable.

THEOREM 1: Let Assumptions A1, A2, A3, and A4 hold. Assume the household's Engel curves of private, assignable goods $H_s^k (y)$ for $k \in \{m, f, c\}$, $y \leq y^* (p)$ are identified. Then resource shares η_{ks} for all household members $k \in \{m, f, c\}$ are identified.

Notes:

1. Theorem 1 says that just from estimates of the household's Engel curves (that is, demand functions in a single price regime) for assignable goods at low expenditure levels, we can identify the fraction of total household resources for all goods that are spent on each household member. Even though resource shares η_{ks} are the fractions of all the household's resources devoted to each household member, we only need to observe their expenditures on three assignable goods (one for each household member type) to identify these resource shares.

2. Many sharing rule identification results in the literature require the existence of "distribution factors," that is, observed variables that affect the allocation of resources within a household but do not affect the preferences and demand functions of individual household members. Theorem 1 does not require the presence of distribution factors. Many identification results also only identify how resource shares change in response to changes in distribution factors, but do not identify the levels of resource shares. Theorem 1 identifies the levels of resource shares, which are important for many policy related calculations such as poverty lines.

3. Theorem 1 assumes that all children in a family are treated equally, and so get equal resource shares. The theorem can be immediately extended to allow and identify, e.g., different shares for older versus younger children, or for boys versus girls, as long as expenditures on a separate assignable good can be observed for each type of child.

4. Theorem 1 applies to households with any number of children, including zero, and so could be used in place of the theorems in Browning, Chiappori, and Lewbel (2008) or Lewbel and Pendakur (2009) for identifying resource shares.

5. The assumptions in Theorem 1 imply that the household Engel curve functions for the assignable goods, $H_s^k (y)$, are shape invariant at low levels of total expenditures y. This can be empirically tested using, e.g., Pendakur (1999).

6. Shape invariance is is often assumed to hold for all goods and all total expenditures, not just assignable goods at low expenditures levels as we require (see, e.g., Blundell, Duncan, and Pendakur (1998), and Blundell, Chen, and Kristensen (2007)). If the assignable good Engel curves do satisfy the required shape invariance at all total expenditure levels, then everything above having to do with the cut off expenditure level $y^* (p)$ can be ignored. This will also help estimation precision, since in this case demand functions at all levels of y, not just those below some $y^* (p)$, will help identify the resource shares.

Now we consider alternative identifying assumptions, based on comparing demand functions across households of different sizes, instead of across individuals within a household. We maintain Assumptions

A1 and A2, but in place of Assumption A3 now assume the following:

ASSUMPTION B3: Define $\overline{p}$ to be the vector of prices of all goods that are private other than p_f, p_m, and p_c. Assume $\overline{p}$ is not empty, and for $t \in \{m, f, c\}$ assume

$$V_t(p, y) = I\left(y \leq y^*(p)\right) \psi_t\left[u_t\left(\frac{y}{G_t(\widetilde{p})}, \frac{\overline{p}}{p_t}\right), \widetilde{p}\right] + I\left(y > y^*(p)\right) \Psi_t(y, p) \tag{20}$$

for some functions y^*, u_t, ψ_t, F_t, and G_t where y^* is strictly positive, G_t is nonzero, differentiable, and homogeneous of degree one, F_t can be vector valued, is differentiable, and is homogeneous of degree zero, and ψ_t and u_t are differentiable and strictly monotonically increasing in their first arguments, and are differentiable and homogeneous of degree zero in their remaining (vector valued) arguments.

The goods in the price vector $\overline{p}$ are assumed to be private, and so have no economies of scale in household consumption, but they need not be assignable, so for example $\overline{p}$ might include food products that are consumed by all household members. Being private means that the elements of $A_s'p$ corresponding to $\overline{p}$ will just equal $\overline{p}$, so the term $\overline{p}/p_t$ will not change when p is replaced by $A_s'p$.

The difference between Assumption A3 and B3 is that the indirect utility function in B3 has the term $u_t\left[y/G_t(\widetilde{p}), \overline{p}/p_t\right]$ in place of $v(y/G_t(p)) + F_t(p)$. So A3 requires some similarity across individual's preferences, in that the function v is the same for all types of individuals t. In contrast, with B3 the u_t expression describing preferences can freely differ across types of individuals, so B3 allows men, women, and children to have completely different demand functions for their own private goods. However, B3 places more limits on how prices can appear inside u_t versus inside v and F_t, which will translate into strong restrictions on cross price effects in the demand functions of the private goods.

Other than replacing $v + F_t$ with u_t, Assumptions A3 and B3 are the same. In particular, the role of the function ψ_t in both cases is to allow the demand functions for all goods other than the private assignable goods to take on any shape, and the role of y^* and Ψ_t is to impose restrictions on preference only for low total expenditure households, leaving the demand functions at higher levels of y completely unconstrained.

To obtain demand functions corresponding to the indirect utility function in Assumption B3, apply Roy's identity to equation (20). As before, for person t and any good k, when $y > y^*(p)$, the demand function will be given by applying Roy's identity to $\Psi_t(y, p)$ giving $h_t(y, p) = -\left[\partial\Psi_t(y, p)/\partial p_k\right]/\left[\partial\Psi_t(y, p)/\partial y\right]$. However, when $y \leq y^*(p)$,applying Roy's identity to equation (20) gives

$$\begin{aligned} h_t(y, p) \quad = \quad & \frac{\psi_t'\left[u_t\left(\frac{y}{G_t(\widetilde{p})}, \frac{\overline{p}}{p_t}\right), \widetilde{p}\right]\left[u_t'\left(\frac{y}{G_t(\widetilde{p})}, \frac{\overline{p}}{p_t}\right)\frac{y}{G_t(\widetilde{p})^2}\frac{\partial G_t(\widetilde{p})}{\partial p_k} - \frac{\partial u_t\left(\frac{y}{G_t(\widetilde{p})}, \frac{\overline{p}}{p_t}\right)'}{\partial(\overline{p}/p_t)}\frac{\partial(\overline{p}/p_t)}{\partial p_k}\right]}{\psi_t'\left[u_t\left(\frac{y}{G_t(\widetilde{p})}, \frac{\overline{p}}{p_t}\right), \widetilde{p}\right]u_t'\left(\frac{y}{G_t(\widetilde{p})}, \frac{\overline{p}}{p_t}\right)\frac{1}{G_t(\widetilde{p})}} \\ & - \frac{\psi_{tk}\left[u_t\left(\frac{y}{G_t(\widetilde{p})}, \frac{\overline{p}}{p_t}\right), \widetilde{p}\right]}{\psi_t'\left[u_t\left(\frac{y}{G_t(\widetilde{p})}, \frac{\overline{p}}{p_t}\right), \widetilde{p}\right]u_t'\left(\frac{y}{G_t(\widetilde{p})}, \frac{\overline{p}}{p_t}\right)\frac{1}{G_t(\widetilde{p})}} \end{aligned}$$

Where ψ_t' and u_t' denote the derivatives of ψ_t and u_t' with respect to their first elements, ψ_{tk} denotes the partial derivative of ψ_t with respect to price p_k, and in a small abuse of notation $\partial u_t/\partial(\overline{p}/p_t)$ is the gradient vector of u_t with respect to the vector $\overline{p}/p_t$.

For the assignable goods $k \in \{m, f, c\}$ these simplify to

$$h_k(y, p) = \frac{\partial u_k\left(\frac{y}{G_k(\widetilde{p})}, \frac{\overline{p}}{p_k}\right)'}{\partial\left(\overline{p}/p_k\right)} \frac{\overline{p}}{p_k^2} \frac{G_k\left(\widetilde{p}\right)}{u_k'\left(\frac{y}{G_k(\widetilde{p})}, \frac{\overline{p}}{p_k}\right)} \quad \text{for } y \leq y^*(p) \tag{21}$$

which we can write simply as

$$h_k(y, p) = \widetilde{f}_k\left(\frac{y}{G_k\left(\widetilde{p}\right)}, p_k, \overline{p}\right) y \quad \text{for } y \leq y^*(p)$$

for functions $\widetilde{f}_k$. Recalling that p_k and $\overline{p}$ do not change when p is replaced with $A_s' p$, substituting this $h_k(y, p)$ expression into equation (15) gives household demand functions for the assignable goods

$$z_s^k = H_s^k(p, y) = \widetilde{f}_k\left(\frac{\eta_{ks}(p)\, y}{G_k\left(\widetilde{A_s'p}\right)}, p_k, \overline{p}\right) \eta_{ks}(p)\, y \quad \text{when } y \leq y^*(p), k \in \{m, f\}$$

and the same expression multiplied by s for $k = c$.

Now consider Engel curves. For the given price regime p we can write the above equation more concisely as

$$\begin{aligned} z_s^k &= H_s^k(y) = f_k\left(\frac{\eta_{ks} y}{G_{ks}}\right) \eta_{ks} y \quad \text{for } y \leq y^*(p), k \in \{m, f\} \\ \text{and } z_s^c &= H_s^c(y) = f_c\left(\frac{\eta_{cs} y}{G_{cs}}\right) s \eta_{cs} y \quad \text{for } y \leq y^*(p). \end{aligned}$$

Define the matrix Ω by

$$\Omega = \begin{pmatrix} \frac{\eta_{m1}}{\eta_{m3}} & 0 & -1 & 0 & 0 & 0 \\ 0 & \frac{\eta_{m1}}{\eta_{m2}} & -1 & 0 & 0 & 0 \\ 0 & \frac{\eta_{m1}}{\eta_{m2}} - \frac{\eta_{c1}}{\eta_{c2}} & 0 & 0 & \frac{\eta_{f1}}{\eta_{f2}} - \frac{\eta_{c1}}{\eta_{c2}} & 0 \\ 0 & 0 & 0 & \frac{\eta_{f1}}{\eta_{f3}} & 0 & -1 \\ 0 & 0 & 0 & 0 & \frac{\eta_{f1}}{\eta_{f2}} & -1 \\ \frac{\eta_{m1}}{\eta_{m3}} - \frac{\eta_{c1}}{\eta_{c3}} & 0 & 0 & \frac{\eta_{f1}}{\eta_{f3}} - \frac{\eta_{c1}}{\eta_{c3}} & 0 & 0 \end{pmatrix}.$$

ASSUMPTION B4: The matrix Ω is finite and nonsingular. $f_k(0) \neq 0$ for $k \in \{m, f, c\}$

Finiteness of Ω only requires that in households with two or three members, no member has a zero resource share. Violating Assumption B4 by having Ω singular would require a perfect coincidence relating the values of resource shares across households of different sizes. One of the few interpretable ways this could happen is if parents in households with two children each have the exact same resources shares as parents in households with three children. These statements, and the matrix Ω, have for simplicity been written using households consisting of s equal to 1, 2, and 3 children (with $s = 1$ shares as numerators), but in fact nonsingularity is only required to hold for any one set of three different household sizes.

The condition in Assumption B4 that $f_k(0) \neq 0$ will hold if the Engel curves for the private, assignable goods, written in budget share form, are continuous and bounded away from zero. This means that the budget shares will not be in a neighborhood of zero for very small total expenditure levels, and by continuity will not hit zero as y gets arbitrarily small. As with Theorem 1 and Assumption A4, the demand functions at all $y \leq y^*(p)$ help in identifying the model, but the technical conditions are easiest to prove in the neighborhood of zero.

THEOREM 2: Let Assumptions A1, A2, B3, and B4 hold for all household sizes s in some set S that has at least three elements. Assume the household's Engel curves of private, assignable goods $H_s^k(y)$ for $k \in \{m, f, c\}$, $y \leq y^*(p)$, $s \in S$ are identified. Then resource shares η_{ks} for all household members $k \in \{m, f, c\}$ and all $s \in S$ are identified.

Notes 1, 2, 3, and 4 listed after Theorem 1 also apply to Theorem 2.

It is possible to have models that satisfy the restrictions of both Theorems 1 and 2, by restricting the function $G_t(p)$ in Assumption A3 to only depend on $\widetilde{p}$ and restricting $F_t(p)$ in A3 to only depend on p_t and $\overline{p}$. Such models will be able to exploit comparisons of individuals both within and across households to strengthen the identification.

6.2 Appendix 2: An Example Model

In this example, we assume that at low total expenditure levels, individual's Engel curves for the assignable private goods m, f, and c, are linear in $\ln(y)$. This requires that the subutility function $\upsilon(Y/G_t(p)) + F_t(p)$ in equation (17) be in Muellbauer's (1976) Price Independent Generalized Logarithmic (PIGLOG) functional form. This form is usually written as $\ln(Y/G_t(p))/\widetilde{F}_t(p)$ for consumer t, for arbitrary (up to regularity) price functions G_t and $\widetilde{F}_t$. However, by ordinality of individual's utility functions, the same demand functions will be obtained using the monotonic transformation $\ln(\ln(Y/G_t(p))) + F_t(p)$, where $F_t(p) = -\ln \widetilde{F}_t(p)$. We therefore suppose that the Assumptions of Theorem 1 hold, with the function υ in equation (17) given by

$$\upsilon\left(\frac{y}{G_t(p)}\right) = \ln\left[\ln\left(\frac{y}{G_t(p)}\right)\right] \tag{22}$$

Then by equations (18) and (19), we can define a function $\widetilde{\delta}_k(p)$ such that

$$\begin{aligned} h_k(y,p) &= \frac{y}{G_k(p)}\frac{\partial G_k(p)}{\partial p_k} - \varphi(p)\frac{G_k(p)}{y}\left[\frac{y \ln y}{G_t(p)} - \frac{y \ln G_t(p)}{G_t(p)}\right] y \\ &= \widetilde{\delta}_k(p)\, y - \varphi(p) \ln y \quad \text{for } y \leq y^*(p). \end{aligned} \tag{23}$$

This then yields private assignable good Engel curves having the functional form

$$\begin{aligned} \frac{z_s^k}{y} &= \widetilde{\delta}_{ks}\eta_{ks} + \varphi_s \eta_{ks} \ln y \quad \text{for } y \leq y^*, k \in \{m, f\} \\ \text{and} \quad \frac{z_s^c}{y} &= \widetilde{\delta}_{cs} s \eta_{cs} + s\varphi_s \eta_{cs} \ln y \quad \text{for } y \leq y^*(p). \end{aligned} \tag{24}$$

with unknown constants $\widetilde{\delta}_{ks}$, φ_s, and η_{ks} for $k \in \{m, f, c\}$. It follows from Theorem 1 that η_{ks} are identified from these Engel curves, but in this case that is easily directly verified. One could simply project (i.e., regress) the observed private assignable good household budget shares z_s^k/y on a constant and on $\ln y$, just using household's having s children and low values of y, to identify the $\ln y$ coefficients $\rho_m = \varphi_s \eta_{ms}$, $\rho_f = \varphi_s \eta_{fs}$, and $\rho_c = \varphi_s \eta_{cs}$ (this last is the coefficient of $s \ln y$ for children) and then use $\eta_{ks} = \rho_{ks}/\left(\rho_{ms} + \rho_{fs} + s\rho_{cs}\right)$ for $k \in \{m, f, c\}$ to identify each η_{ks}.

In this example if $\varphi(p)$ only depends on the prices of private goods $\overline{p}$, then Assumption B3 will also be satisfied. In this case the assignable good Engel curves will be given by equation equation (24) with $\varphi_s = \varphi$, the same constant for all household sizes s. In this case, identification can be obtained by either Theorem 1 or Theorem 2, specifically, we can compare the coefficient of $\ln y$ both across individuals within a household and across households of different sizes to identify and hence estimate the resource shares η_{ts}.

6.3 Appendix 3: A Fully Specified Example Model

The information and derivation in the previous section is all that is required to apply our estimator empirically. However, to clarify how our assumptions work and interact, we will now provide an example of functional forms for the entire household model that incorporate the above piglog private goods, and in particular verify that resource shares can be independent of y.

First assume each household member t has utility given by Muellbauer's piglog model so, the function υ is given by equation (22), and let $\ln F_t(p) = \ln p_t - a' \ln \widetilde{p}$ for some constant vector a with elements a_k that sum to one. This is a simple example of a function that is homogeneous as required and is a special case of $F_t(p) = p_t \widetilde{\varphi}(\widetilde{p})$ as described in the text after Assumption A3. As noted there, if all the private assignable goods have the same price, then we could instead take F_t to be any suitably regular price function, instead of requiring $F_t(p) = p_t \widetilde{\varphi}(\widetilde{p})$.

For simplicity let $y^*(p)$ be larger than any household's actual y, so the functional forms of $y^*(p)$ and of $\Psi_t(y, p)$ are irrelevant and drop out of the model. This assumption makes private assignable good Engel curves be piglog, hence linear in $\ln y$, at all total expenditure levels, not just at low levels as the theorem requires. Also for simplicity let the function $\psi_t(\upsilon + F_t, \widetilde{p}) = \exp(\upsilon + F_t)$, which by not depending upon $\widetilde{p}$ makes individual Engel curves for all goods be the same as those of the private assignable goods, and exponentiating provides a convenient cardinalization for pareto weighting utility within the household. Finally, in a small abuse of notation let $G_t(p) = G_t(p_t, \widetilde{p})$, which makes explicit the assumption that the goods p_t are assignable, so e.g. the price p_m of the good that is assignable to the father does not appear in a child's utility function, and hence does not appear in $G_c(p_c, \widetilde{p})$.

The combination of all these assumptions means that the indirect utility functions for each household member t are given by

$$\ln V_t(p, y) = \ln \left[\ln \left(\frac{y}{G_t(p_t, \widetilde{p})} \right) \right] + p_t e^{-a' \ln \widetilde{p}} \tag{25}$$

Let the function $\widetilde{U}_s$, which describes how the household weighs together the utility functions of its members, be a general Bergson-Samuelson social welfare function

$$\widetilde{U}_s(U_f, U_m, U_c, p/y) = \omega_f(p)\left[U_f + \rho_f(p)\right] + \omega_m(p)\left[U_m + \rho_m(p)\right] + \left[U_c + \rho_c(p)\right]\omega_c(p) \tag{26}$$

Note that the positive Pareto weight functions $\omega_t(p)$ and the utility transfer or externality functions $\rho_f(p)$ must be homogenous of degree zero by our Assumptions, so e.g. $\omega_t(p) = \omega_t(p/y)$, but otherwise these functions are unrestricted.

Assume the matrix A_s, which defines the extent to which goods are consumed jointly rather than privately, is diagonal, and let A_{sk} denote the k'th element along the diagonal. In the terminology of Browning, Chiappori, and Lewbel (2008), this is a Barten type consumption technology, so each A_{sk} gives the degree of publicness vs privateness of the good k in a household with s children.

Substituting this structure for A_s and equation (26) into equation (12) gives a household with s children the maximization problem

$$\max_{x_f, x_m, x_c, z_s} \omega(p) + \omega_f(p) U_f(x_f) + \omega_m(p) U_m(x_m) + \omega_c(p) U_c(x_c)$$

$$\text{such that } z_s^k = A_{sk}\left[x_{fk} + x_{mk} + s x_{ck}\right] \text{ for each good } k, \text{ and } y = z_s' p$$

where $\omega(p) = \omega_f(p)\rho_f(p) + \omega_m(p)\rho_m(p) + \rho_c(p)\omega_c(p)$. This maximization can be decomposed into two steps as follows. Define resource shares η_{ts} for $t = m, f, c$ by $\eta_{ts} = x_t' A_s p / y = \sum_k A_{sk} p_k x_{tk} / y$, evaluated at the optimized level of expenditures x_t. In a lower step, conditional upon knowing η_{ts}, each household member can choose their optimal bundle x_t by maximizing $U_t(x_t)$ subject to the constraint $\sum_k A_{sk} p_k x_{tk} = \eta_{ts} y$. This is identical to standard utility maximization facing a linear budget constraint with prices $A_{sk} p_k$ and total expenditure level $\eta_{ts} y$. The resulting optimized utility level is then given by the individual's indirect utility function V_t evaluated at these shadow (Lindahl) prices, that is, $V_t\left(A_s' p, \eta_{ts} y\right)$.

Substituting these maximum attainable utility levels for each individual into the household's maximization problem then reduces the household's problem to determining optimal resource share levels by

$$\max_{\eta_{ms}, \eta_{fs}, \eta_{cs}} \omega(p) + \omega_f(p) V_f\left(A_s' p, \eta_{fs} y\right) + \omega_m(p) V_m\left(A_s' p, \eta_{ms} y\right) + \omega_c(p) V_c\left(A_s' p, \eta_{cs} y\right) \tag{27}$$

$$\text{such that } \eta_{ms} + \eta_{fs} + s\eta_{cs} = 1$$

Given our chosen functional form for utility, substituting equation (25), into equation (27) gives

$$\max_{\eta_{ms}, \eta_{fs}, \eta_{cs}} \omega(p) + \widetilde{\omega}_{fs}(p) \ln\left(\frac{\eta_{fs} y}{G_f\left(A_s' p\right)}\right) + \widetilde{\omega}_{ms}(p) \ln\left(\frac{\eta_{ms} y}{G_m\left(A_s' p\right)}\right)$$

$$+\widetilde{\omega}_{cs}(p) \ln\left(\frac{\eta_{cs} y}{G_c\left(A_s' p\right)}\right) \quad \text{such that } \eta_{ms} + \eta_{fs} + s\eta_{cs} = 1$$

where $\widetilde{\omega}_{ts}(p) = \omega_t(p) \exp\left(A_{st} p_t e^{-a'\left(\ln \widetilde{p} + \ln \widetilde{A}_s\right)}\right)$. Using a lagrange multiplier for the constraint that resource shares sum to one, the first order conditions for this maximum are

$$\frac{\widetilde{\omega}_{fs}(p)}{\eta_{fs}} = \frac{\widetilde{\omega}_{ms}(p)}{\eta_{ms}} = \frac{\widetilde{\omega}_{cs}(p)}{s\eta_{cs}}$$

which has the solution

$$\begin{aligned}
\eta_{ks}(p) &= \frac{\widetilde{\omega}_{ks}(p)}{\widetilde{\omega}_{fs}(p) + \widetilde{\omega}_{ms}(p) + \widetilde{\omega}_{cs}(p)} \quad \text{for } k \in \{m, f\} \\
\eta_{cs}(p) &= \frac{\widetilde{\omega}_{cs}(p)/s}{\widetilde{\omega}_{fs}(p) + \widetilde{\omega}_{ms}(p) + \widetilde{\omega}_{cs}(p)}
\end{aligned}$$

These explicit formulas for the resource shares in this example do not depend on y, as required by Assumption A1.

Given these resource shares, the household's demand functions can now be obtained by having each household member choose their optimal bundle x_t by maximizing $U_t(x_t)$ subject to the constraint $\sum_k A_{sk} p_k x_{tk} = \eta_{ts} y$, which by standard utility duality theory is equivalent to applying Roys identity to the member's indirect utility function evaluated at prices $A_s' p$ and total expenditure level $\eta_{ts} y$, that is, $V_t\left(A_s' p, \eta_{ts} y\right)$, where the function $V_t(p, y)$ is given by equation (25).

Applying Roy's identity to equation (25) gives individual's demand functions

$$h_t^k(y, p) = \frac{y}{G_t(p_t, \widetilde{p})} \frac{\partial G_t(p_t, \widetilde{p})}{\partial p_k} - \frac{\partial\left(p_t e^{-a' \ln \widetilde{p}}\right)}{\partial p_k}\left[\ln y - \ln G_t(p_t, \widetilde{p})\right] y \tag{28}$$

for each good k and any individual t. Recalling that the sharing technology matrix A_s is diagonal, the household's quantity demand functions satisfy

$$z_s^k = A_{sk}\left[h_f^k\left(A_s' p, \eta_{fs}(p)\, y\right) + h_m^k\left(A_s' p, \eta_{ms}(p)\, y\right) + s h_c^k\left(A_s' p, \eta_{cs}(p)\, y\right)\right] \tag{29}$$

The demand functions of a household having s children, for each good k, are therefore obtained by substituting equation (28), and the above derived expression for $\eta_{ts}(p)$, for $t = f, m, c$, into equation (29).

Equation (28) can be written more simply as

$$h_t^k(y, p) = \widetilde{\delta}_{kt}(p)\, y - \varphi_t^k(p)\, y \ln y$$

which, when substituted into equation (29) gives household demand equations of the form

$$\begin{aligned}
\frac{z_s^k}{y} &= \left(\widetilde{\delta}_{kf}\left(A_s' p\right) + \widetilde{\delta}_{km}\left(A_s' p\right) + s\widetilde{\delta}_{kc}\left(A_s' p\right)\right) A_{sk} \\
&\quad - \left(\varphi_f^k\left(A_s' p\right) \ln \eta_{fs}(p) + \varphi_m^k\left(A_s' p\right) \ln \eta_{ms}(p) + s\varphi_m^k\left(A_s' p\right) \ln \eta_{cs}(p)\right) A_{sk} \\
&\quad - \left(\varphi_f^k\left(A_s' p\right) + \varphi_m^k\left(A_s' p\right) + s\varphi_m^k\left(A_s' p\right)\right) A_{sk} \ln y
\end{aligned}$$

For the private, assignable goods, this expression simplifies to the demand functions given earlier. Evaluating this equation in a single price regime shows that, in this model, the resulting Engel curves for all goods have the piglog form

$$\frac{z_s^k}{y} = \delta_{ks} + \varphi_s^k \eta_{ks} \ln y.$$

6.4 Appendix 4: Proofs

Proof of Theorem 1: We have already in the above text derived the household Engel curve functions for the assignable goods at low expenditure levels, that is, for $y \le y^*$, $H_s^k(y) = \delta_{ks}\eta_{ks} y + g_s\left(\frac{\eta_{ks} y}{G_{ks}}\right)\eta_{ks} y$ for $k \in \{m, f\}$, and the same equation multiplied by s for $k = c$. Define $\widetilde{h}_s^k(y) = \partial\left[H_s^k(y)/y\right]/\partial y$ and define $\lambda_s = \lim_{y\to 0}\left[y^\zeta g_s''(y)/g_s'(y)\right]^{\frac{1}{1-\zeta}}$, where by assumption $\zeta \ne 1$ (the alternative log polynomial case is considered below). Since the functions $H_s^k(y)$ are identified, we can identify $\kappa_{ks}(y)$ for $y \le y^*$, defined by

$$\begin{aligned}
\kappa_{ks}(y) &= \left(y^\zeta \frac{\partial \widetilde{h}_s^k(y)/\partial y}{\widetilde{h}_s^k(y)}\right)^{\frac{1}{1-\zeta}} \\
&= \left(\left(\frac{\eta_{ks}}{G_{ks}}\right)^{-\zeta}\left(\frac{\eta_{ks} y}{G_{ks}}\right)^{\zeta}\left[g_s''\left(\frac{\eta_{ks} y}{G_{ks}}\right)\frac{\eta_{ks}^3}{G_{ks}^2}\right] \Big/ \left[g_s'\left(\frac{\eta_{ks} y}{G_{ks}}\right)\frac{\eta_{ks}^2}{G_{ks}}\right]\right)^{\frac{1}{1-\zeta}} \\
&= \frac{\eta_{ks}}{G_{ks}}\left[\left(\frac{\eta_{ks} y}{G_{ks}}\right)^{\zeta} g_s''\left(\frac{\eta_{ks} y}{G_{ks}}\right) \Big/ g_s'\left(\frac{\eta_{ks} y}{G_{ks}}\right)\right]^{\frac{1}{1-\zeta}} = \frac{\eta_{ks}}{G_{ks}}\left(y_{ks}^\zeta \frac{g_s''(y_{ks})}{g_s'(y_{ks})}\right)^{\frac{1}{1-\zeta}}
\end{aligned}$$

and, in particular,

$$\kappa_{ks}(0) = \frac{\eta_{ks}}{G_{ks}}\lambda_s$$

so for any $y \leq y^*$ we can identify $\rho_{ks}(y)$ defined by

$$\rho_{ks}(y) = \frac{\widetilde{h}_s^k(y/\kappa_{ks}(0))}{\kappa_{ks}(0)} = g_s'\left(\frac{y}{\lambda_s}\right)\frac{\eta_{ks}}{\lambda_s}$$

and by equation (14), we can then identify the resource shares η_{ks} for each household member k by $\eta_{ks} = \rho_{ks}/\left(\rho_{ms} + \rho_{fs} + s\rho_{cs}\right)$.

Now consider the case where g_s is a polynomial of some degree λ in logarithms, so

$$g_s\left(\frac{\eta_{ks}y}{G_{ks}}\right) = \sum_{\ell=0}^{\lambda}\left(\ln\left(\frac{\eta_{ks}}{G_{ks}}\right) + \ln(y)\right)^{\ell} c_{s\ell}$$

for some constants $c_{s\ell}$, and therefore for any $y \leq y^*$ we can identify $\widetilde{\rho}_{ks}$ defined by

$$\widetilde{\rho}_{ks} = \frac{\partial^{\lambda}\left[h_s^k(y)/y\right]}{\partial(\ln y)^{\lambda}} = c_{s\lambda}\eta_{ks}$$

which identifies resource shares by $\eta_{ks} = \widetilde{\rho}_{ks}/\left(\widetilde{\rho}_{ms} + \widetilde{\rho}_{fs} + s\widetilde{\rho}_{cs}\right)$.

Proof of Theorem 2: In the text we derived the household Engel curve functions for the assignable goods at low expenditure levels, which are, for $y \leq y^*$, $H_s^k(y) = f_k\left(\frac{\eta_{ks}y}{G_{ks}}\right)\eta_{ks}y$ for $k \in \{m, f\}$, and the same equation multiplied by s for $k = c$. Let s and 1 be two elements of S. Since the functions $H_s^k(y)$ and $H_1^k(y)$ are identified, we can identify ς_{ks} defined by $\varsigma_{ks} = \lim_{y\to 0} H_1^k(y)/H_s^k(y)$, and

$$\varsigma_{ks} = \frac{f_k(0)\,\eta_{k1}y}{f_k(0)\,\eta_{ks}y} = \frac{\eta_{k1}}{\eta_{ks}} \text{ for } k \in \{m, f\}, \text{ and } \varsigma_{cs} = \frac{f_k(0)\,\eta_{c1}y}{f_k(0)\,s\eta_{cs}y} = \frac{\eta_{c1}}{s\eta_{cs}}$$

so

$$\begin{aligned}
\varsigma_{ms}\eta_{ms} + \varsigma_{fs}\eta_{fs} + \varsigma_{cs}s\eta_{cs} &= \eta_{m1} + \eta_{f1} + \eta_{c1} = 1 \\
\varsigma_{ms}\eta_{ms} + \varsigma_{fs}\eta_{fs} + \varsigma_{cs}\left(1 - \eta_{ms} - \eta_{fs}\right) &= 1 \\
\left(\varsigma_{fs} - \varsigma_{cs}\right)\eta_{fs} + \left(\varsigma_{ms} - \varsigma_{cs}\right)\eta_{ms} &= 1 - \varsigma_{cs}
\end{aligned}$$

These equations for $k \in \{m, f\}$ and for $s \in \{2, 3\}$ give the matrix equation

$$\begin{pmatrix}
\varsigma_{m3} & 0 & -1 & 0 & 0 & 0 \\
0 & \varsigma_{m2} & -1 & 0 & 0 & 0 \\
0 & \varsigma_{m2} - \varsigma_{c2} & 0 & 0 & \varsigma_{f2} - \varsigma_{c2} & 0 \\
0 & 0 & 0 & \varsigma_{f3} & 0 & -1 \\
0 & 0 & 0 & 0 & \varsigma_{f2} & -1 \\
\varsigma_{m3} - \varsigma_{c3} & 0 & 0 & \varsigma_{f3} - \varsigma_{c3} & 0 & 0
\end{pmatrix}
\begin{pmatrix}
\eta_{m3} \\ \eta_{m2} \\ \eta_{m1} \\ \eta_{f3} \\ \eta_{f2} \\ \eta_{f1}
\end{pmatrix}
=
\begin{pmatrix}
0 \\ 0 \\ 1 - \varsigma_{c2} \\ 0 \\ 0 \\ 1 - \varsigma_{c3}
\end{pmatrix}$$

The six by six matrix in this equation equals $\mathbf{\Omega}$ in the text using $\varsigma_{ks} = \eta_{k1}/\eta_{ks}$. Since $\mathbf{\Omega}$ is nonsingular, the above equation can be solved for η_{ms} and η_{fs} for $s \in \{1, 2, 3\}$, meaning that these resource shares are identified because they can be written entirely in terms of the identified parameters ς_{ks}. Children's resource shares are then identified for these household types by $\eta_{cs} = \left(1 - \eta_{ms} - \eta_{fs}\right)/s$, and resource shares for households of other types s are identified by $\eta_{ks} = \eta_{k1}/\varsigma_{ks}$ for any s.

7 References

Anderson, S. and Baland, J-M (2002), "The Economics of ROSCAs and Intrahousehold Resource Allocation," *Quarterly Journal of Economics*, August, 963-995.

Banerjee, A.V. and Dulfo, E. (2008), "What is middle class about the middle classes around the world," *Journal of Economic Perspectives*,

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

Becker, G. (1965), "A Theory of the Allocation of Time," *Economic Journal*, 75, 493-517.

Becker, G. (1981), "Altruism in the Family and Selfishness in the Market Place," *Economica*, 48, 1-15

Blundell, R., P.-A. Chiappori and C. Meghir (2005), "Collective Labor Supply with Children," *Journal of Political Economy*, 113, 1277-1306.

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

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.

Browning, M., and P.A. Chiappori (1998), "Efficient Intra-household Allocations: A General Characterization and Empirical Tests," *Econometrica*, 66, 1241-1278.

Browning, M., P.A. Chiappori and A. Lewbel (2008), "Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining power," mimeo.

Browning, M. and Thomas, Irene, 1999...

Cherchye, L., B. De Rock and F. Vermeulen (2008), "Opening the black box of intra-household decision-making: Theory and non-parametric empirical tests of general collective consumption models", *Journal of Political Economy*.

Cherchye, L., B. De Rock and F. Vermeulen (2010), "The revealed preference approach to collective consumption behavior: testing and sharing rule recovery", forthcoming in *Review of Economic Studies*

Chiappori, P.-A., (1988), "Nash-Bargained Households Decisions: A Comment," *International Economic Review*, 29, 791-796.

Chiappori, P.-A., (1992), "Collective Labor Supply and Welfare," *Journal of Political Economy*, 100, 437-467.

Chiappori, P.-A. and I. Ekelund (2009), "The Micro Economics of Efficient Group Behavior: Identification," *Econometrica*, 77, 763-99.

Chiappori, P.-A., R. Blundell and C. Meghir (2005), "Collective Labor Supply With Children," Journal of Political Economy, vol. 113, 1277-1306.

Dauphin, Anyck, Bernard Fortin, Guy Lacroix and Abdel-Rahmen el Lahga (2008), "Are Children Decision-Makers Within the Household", IZA Discussion Paper number 3728.

Deaton, Angus (1989), "Looking for Boy-Girl Discrimination in Household Expenditure Data," *World Bank Economic Review*, 3(1): 1-15.

Deaton, Angus (1997), *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Washington, DC: The World Bank.

Duflo, Esther (2003), "Grandmothers and Granddaughters: Old Age Pension and Intra-Household Allocation in South Africa," *World Bank Economic Review*, 17(1), 2003, 1-25.

Duflo, E. (2005), "Gender Equality in Development," mimeo.

Lazear, E. P. and R. T. Michael (1988), Allocation of Income Within the Household, University Chicago Press.

Lewbel, A. and K. Pendakur (2008), “Estimation of Collective Household Models With Engel Curves,” *Journal of Econometrics*, 147, 350-358.

Lindahl, E., (1919), “Just Taxation - A Positive Solution,” in R. Musgave and A. Peacock, eds., *Classics in the Theory of Public Finance*, London: MacMillan.

Lise, J. and S. Seitz, (2004). “Consumption Inequality and Intra-Household Allocations,” Working Papers, 1019, Queen’s University, Department of Economics.

Pendakur, K. (2002). “Taking Prices Seriously in the Measurement of Inequality,” *Journal of Public Economics*, 86, pp47-69..

Rose, Elaina (1999), “Consumption Smoothing and Excess Female Mortality in Rural India,” Review of Economics and Statistics, 81(1), 41-49.

Sen, Amartya (1990), “More Than 100 Million Women Are Missing,” *The New York Review of Books*, 37(20).

Vermeulen, F., (2002), “Collective Household Models: Principles and Main Results,” *Journal of Economic Surveys*, 16, 533-564.