Identification of Children's Resources in Collective Households

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

Oct. 2009 - Preliminary Draft

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), with resources within household that can be estimated from data. Extending the frameworks of Browning, Chiappori and Lewbel (2007) and Lewbel and Pendakur (2008), we show identification of children's resource shares within households both with and without price variation. We estimate the model using household micro-data from Canada and from Malawi. These types of 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.

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.

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.

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, but they do not provide estimates of the share of resources children consume.

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, allowing for very general forms of sharing of goods, but they only show identification 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, they observe individual's demand functions by observing data from single men and women living alone, and combine that with data on couples living together, assuming limited differences between the utility functions of single and married men and between those of single and married women.s versus those of

Our contribution is to extend the model of Browning, Chiappori and Lewbel (2007), which allow for very general types of consumption sharing, to include children. Specifically, we show nonparametric identification of children's resource shares in the BCL model. Unlike BCL, instead of combining data of couples with singles, we show identification by combining data on households with varying numbers of children. This imposes milder conditions regarding stability of preferences across household types than BCL, 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 (though assuming the latter similarity increases precision of the estimates if it holds).

A good is defined to be *assignable* if it is consumed exclusively by a particular known individual (or subset of individuals) within a household. Examples could include toys and diapers assignable to children, or alcohol and tobacco 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 identification strategy follows this line in assuming the presence of a small number of private assignable goods, and in fact only requires observation of the demand functions of some assignable goods for the identification of childrens resource shares.

Like the BCL model, our nonparametric identification is in a demand system setting that exploits both price and expenditure variation across households. Lewbel and Pendakur (2008) provide additional modeling restrictions that permit identification of resource shares in the BCL model in an Engel curve setting, without price variation. This greatly facilitates empirical application of the model, since it substantially reduces both model complexity and data requirements. We also extend the present paper's results to permit identification of childrens resource shares without price variation. These Engel curve results impose some functional form restrictions, but these functional forms are empirically tractable, and are commonly used for Engel curve models.

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 affect the level and distribution of income within households.

We present empirical results for children's resource shares in Canada and Malawi. In both countries, children command a fairly large share of resources, which rises with the number of children. In general, we find that children's resource shares come at the expense of the mother's resource share. In contrast, father's resource shares remains relatively constant regardless of the number of children. Canada and Malawi represent extreme ends of the development spectrum. Per-capita GDP in 2007 was $43,400 and $256 in US dollars, respectively. While our results are not definitive, they suggest that the share of household resources devoted to children is roughly comparable across both countries. Thus, development projects that raise household income may well improve the well-being of all household members, *ceteris paribus*.

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 then determines their own demand for each consumption good by maximizing their own utility function, subject to this shadow budget constraint defined by their share of household resources and the within-household shadow prices of goods.

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, are individually and exclusively consumed) will have shadow prices equal to market prices.

Together, resource shares and shadow prices define the shadow budget constraint faced by individuals within households, and therefore 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. We only consider households consisting of a mother, a father, and one or more children, so we index households by the size measure $s = 1, 2, ...$ where s is the number of children in the family. 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 wages that may help to determine bargaining power and hence resource shares devoted to each household member. We will reintroduce these explicitly later.

Household 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 $x_t = \left(x_s^1, ..., x_s^K\right)'$ be the $K-$vectors of quantities of each good k consumed by an individual of type t 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 one'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 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 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

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

3 Identification

In an Appendix we provide formal identification theorems and proofs regarding nonparametric and semi-parametric identification of our extension of the BCL model to children. Here we will summarize the main points required for our empirical work.

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 (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 fraction of the household's total resources consumed by a person of type t in a household with s children. Our main goal will be identification of η_{cs}. The total share of resources going to children in the household will be $s\eta_{cs}$, with each child getting a share η_{cs}.

Private goods are consumed exclusively by one person, and if that private good is assignable, then the researcher can observe which individual consumes it. Our identification strategies rely on some private assignable goods. 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 or assignable are more complicated (see the Appendix for details, especially equation 10), the household demand functions for private assignable goods, derived from equation (1), have the simple forms

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

Household demand functions W_{ts} (the left side of equation 2) are observable by measuring the consumption patterns of households in various y and p regimes. The goal is identification of features of the right side of equation (2), in particular η_{cs}.

The way BCL obtain identification is that they assume $s = 0$ and the demand functions w_m, and w_f (on the right side of equation 2) are are assumed to be 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, which is implied by the assumption that, up to a monotonic transformation, the utility functions $U_f\left(x_f\right)$ and $U_m\left(x_m\right)$ 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).

The BCL identification method cannot be immediately and applied to our application involving children, because unlike men or women we cannot observe 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. Note, however, 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 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 $A_t' p$, do vary with s.

The fact that resource shares η_{ts} appear both inside and outside the unobserved individual demand functions w_t complicates identification in this model. We get around this by imposing the restriction that resource shares η_{ts} are independent of prices p and household expenditure y (though this assumption is later relaxed to only requiring indepence of y in our Engel curve identification). This restriction may seem severe, but we offer two defenses: first, most empirical studies of resource shares implemented in the literature impose this restriction; and second, we only require that η_{ts} be independent of p and y after conditioning on other related features of individuals or households. So, e.g., η_{ts} can depend arbitrarily on the wages or income levels of household members, on household savings rates, and on other demographic and distribution factors (including the number of children s).

Theorem 1 in the Appendix presents our result for the identification of resource shares using assignable goods demands with price and expenditure variation. Even though the functions w_{ts} and parameters A_s and η_{ts} are not directly observed, we nevertheless show in Theorem 1 that sufficient information about the w_{ts} functions can be nonparametrically recovered from the household demand functions $W_{ts}(y,p)$ to identify the household member resource shares η_{ts}, particularly the child resource shares η_{cs}. The intuition for this result is that variation in s holding p and y fixed affects only η_{ts} and A_s, and variation in p and y identifies just enough features of w_{ts} to allow us to separate the effects of s on η_{ts} from the effects of s on A_s. The identification depends only on the private goods, so it is not necessary to estimate the more complicated demand functions of the other goods.

Given data limitations on observable relative price variation, and the complexity of specifying and estimating full demand systems, we also consider identification and estimation based on Engel curve data without price variation. Like the case with price variation, Engel curve identification relies on the presence of private assignable goods. Unlike the case with price variation, identification with Engel curves does not require that resource shares be independent of both prices and expenditure, but rather that they are independent of expenditure (though again they can depend on other related variables like wages and demographics, and only need to be conditionally independent of y after conditioning on these other observed variables). This is because one could in principle identify the resource at each price regime separately. However, because less information on demands are available for use without price variation, we impose some semiparametric functional form restrictions to identify resource shares just from Engel curve data.

Theorems 2 and 3 in the appendix show semiparametric identification of resource shares using just data on households of varying sizes and income levels in a single time period (i.e., a single price regime). In a slight abuse of notation, we may write these Engel curve based models as

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

where for identification we assume that the functions w_{ts} have one of two forms: either, based on Theorem 2,

$$w_{ts}\left(\eta_{ts} y\right) = \alpha_{ts} + \beta_t \ln\left(\eta_{ts} y\right), \tag{4}$$

for some constants α_{ts} and β_t, or, based on Theorem 3,

$$w_{ts}\left(\eta_{ts} y\right) = \alpha_t + M_{ts}\left(\eta_{ts} y\right) \tag{5}$$

for some constants α_t and functions $M_{ts}\left(\eta_{ts} y\right)$ where the dimension of the space spanned by the functions M_{ts} for $t = c, f$, and m is two.

Equation (4), is the class of Engel curves that are known as Working (1943) and Leser (1963) or PIGLOG (see Muellbauer 1975) Engel curves, and are equal to the Engel curves that would arise if individual utility functions $U_t(x_t)$ were those of Deaton and Muellbauer's (1980) Almost Ideal Demand (AID) model or of Jorgenson, Lau, and Stoker's (1982) exactly aggregable Translog demand system. This is a very popular class of Engel curves for Empirical work.

Equation (5) is a general class of rank two models (see Lewbel 1991). Note equation (4) is also rank two. Unlike equation (4) which has demand functions linear in $\ln y$, equation (5) allows for general Engel curve shapes, such as polynomials where

$$M_{ts}\left(\eta_{ts} y\right) = \sum_{j=1}^{J} \beta_{stj}\left(\eta_{ts} y\right)^j$$

and for each s the $3 \times (J+1)$ matrix of coefficients consisting of α_t and β_{stj} for $t = c, f$, and m and $j = 1, ..., J$ has rank two. However, relative to Theorem 2, Theorem 3 imposes some strong restrictions on the utility functions and sharing parameters A_s, though these restrictions would only be observable and testable given data with price variation. See the Appendix for details.

It is important to stress that, unlike most existing resource sharing rule identification results (as discussed in the introduction) 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. Both are crucially important for our policy analysis, which is to measure the relative welfare of children in households of varying composition and income levels.

A second feature of our identification results is that the associated estimators are easy to implement estimators. In particular, if the researcher uses the form (4), then 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 in parameters, but there are nonlinear restrictions on the parameters that, for efficiency, should be imposed upon estimation. In either case, estimation is much less onerous than most 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 Demand System Estimation

In this section, we estimate complete consumer demand systems using Canadian household-level price and expenditure microdata. Here, we estimate models under the identification conditions of Theorem 1, where we observe price variation. We can then use the results to assess the restrictions that allow for identification from just Engel curve data as in Theorems 2 and 3. In the next section, we will then use Malawian household-level Engel curve data to investigate children's resource shares in a developing country context.

The Canadian data used here come from the following public use sources: (1) the Family Expenditure Surveys 1969, 1974, 1978, 1982, 1984, 1986, 1990, 1992 and 1996; and (2) Browning and Thomas (1999), with extensions to rental prices from Pendakur (2002). These data have been used in many applications, and Pendakur (2002) provides a detailed description of them. Price and expenditure data are used from 9 years in 5 regions (Atlantic, Quebec, Ontario, Prairies and British Columbia) yielding 45 distinct price vectors. Prices are normalised so that the price vector facing residents of Ontario in 1986 is (1, ..., 1).

The empirical analysis uses annual expenditure in the following expenditure categories: food-in, food-out, clothing and other nondurable expenditures (comprised of household operation, household furnishing & equipment, transportation operation, recreation and personal care). Clothing is broken into three assignable goods: men's, women's and children's clothing. Other nondurable expenditure is the left-out equation, yielding five expenditure share equations to be estimated. These expenditure categories, which exclude large durables (such as shelter), account for about 55% of the current consumption of the households in the sample. The log of total expenditure is normalised so that its mean across the entire sample is zero.

Clothing is taken to be the assignable private good in this set of demand system estimates. The assumption that clothing is assignable is reasonable in this context, since respondents are asked to identify the household member for whom clothing purchases are intended.

Our sample for estimation consists of 4193 observations of families with at least one, and not more than 4, children. Due to data topcoding, we only use households with 1-3 children in the 1992 data, and households with 1-2 children in the 1996 data. In addition to estimation with a sample of households with children, we provide some estimates with the addition of 3581 childless couples, and with the addition of 5778 single men and women. We use only persons aged 25 to 64 living in cities with at least 30,000 residents. City residents are used to minimise the effects of possible home and farm production. Table 1 gives data means for our estimation sample.

Table 1: Data Means, Canadian micro-data		singles	childless couples	couples with 1 child	2 children	3 children	4 children
Number of Observations		5778	3581	1995	1517	531	150
budget-shares	food-in	0.265	0.256	0.305	0.337	0.379	0.427
	food-out	0.117	0.096	0.065	0.058	0.048	0.048
	men's clothing	0.039	0.053	0.044	0.039	0.037	0.036
	women's clothing	0.073	0.077	0.054	0.045	0.042	0.036
	children's clothing	0.000	0.000	0.022	0.039	0.051	0.061
	nondurable cons.	0.506	0.518	0.508	0.482	0.442	0.402
log-total-expenditure		-0.184	0.308	0.213	0.293	0.152	-0.015

We assume that clothing demand equations for each individual, $j = c, m, f$, have the following Quadratic Almost Ideal (QAI, see Banks, Blundell and Lewbel 1997 for details) form:

$$w_{js}(y) = a_j + C_j' \ln p + b\widetilde{y}_j + q\widetilde{y}_j^2 / d(\ln p),$$

where a_j, b and q are scalar parameters, A_j are parameter vectors, $\widetilde{y}_j$ is the deflated log-real-expenditure for each person given their QAI demands, and $d(p)$ is a QAI price deflator common to all persons. Like in BCL, we assume that the matrix A which governs scale economies and within-household goods complementaries is a diagonal matrix, with diagonal elements given by the vector α. So, substituting these into (2), we have

$$\begin{aligned}
W_{cs}(y) &= s\eta_{cs}\left[a_c + C_c'(\ln p + \ln\alpha) + b\left(\ln\eta_{cs} + \widetilde{y}_c\right) + q\left(\ln\eta_{cs} + \widetilde{y}_c\right)^2 / d(\ln p + \ln\alpha)\right], \\
W_{ms}(y) &= \eta_{ms}\left[a_m + C_m'(\ln p + \ln\alpha) + b\left(\ln\eta_{ms} + \widetilde{y}_m\right) + q\left(\ln\eta_{ms} + \widetilde{y}_m\right)^2 / d(\ln p + \ln\alpha)\right], \\
W_{fs}(y) &= \eta_{fs}\left[a_f + C_f'(\ln p + \ln\alpha) + b\left(\ln\eta_{fs} + \widetilde{y}_f\right) + q\left(\ln\eta_{fs} + \widetilde{y}_f\right)^2 / d(\ln p + \ln\alpha)\right],
\end{aligned}$$

for $s = 1, 2, 3, 4$.

Theorem 1 tells us that these 3 budget-share equations are sufficient to identify the resource shares of all household members. Budget-share equations for food-in and food-out are also included in the model, but are not necessary for identification. These budget-share equations are more complicated, and are described in detail in the Appendix. The key features of this QAI demand system are: (1) children, men and women have different preferences from each other which are captured in the a_j and C_j parameters; and (2) Engel curves are quadratic in log real expenditure $\widetilde{x}_j$; and (3) Engel curves can be linear, and if they are, then $q = 0$.

Table 2 gives estimated resource shares from 6 different models, which vary in their model specification and the sample of data used. Standard errors are presented for men and children, but not for women, because the woman's resource share is computed from the summation restriction on resource shares. We estimate both QAI and Almost Ideal (AI, see Deaton and Muelbauer 1980) models. The AI models equal the QAI after imposing the parametric restriction that $q = 0$ (and imposes the same restriction on the food-in and food-out equations). The key feature of these models is that the QAI model of consumer demand is a rank 3 (see Lewbel 1991) model which features Engel curves that are quadratic in log-expenditure whereas the AI model is a rank 2 model with Engel curves that are linear in log-expenditure. In terms of our work, this distinction is important, because Theorems 2 and 3 offer identification results only for rank 2 demands like those of the AI model.

Table 2: Estimated Resource Shares from a full Demand System													
		including singles				excluding singles				excluding childless			
		QAI		AI		QAI		AI		QAI		AI	
	household	η	std err	η	std err	η	std err	η	std err	η	std err	η	std err
man	coup no kids	0.502	0.005	0.503	0.005	0.373	0.047	0.361	0.016				
	coup w/1 kid	0.427	0.020	0.420	0.024	0.382	0.038	0.344	0.019	0.328	0.038	0.328	0.023
	coup w/2 kids	0.434	0.028	0.410	0.028	0.435	0.039	0.404	0.024	0.330	0.029	0.324	0.023
	coup w/3 kids	0.476	0.042	0.433	0.042	0.478	0.043	0.450	0.034	0.338	0.034	0.364	0.034
	coup w/4 kids	0.472	0.072	0.444	0.071	0.498	0.075	0.448	0.056	0.371	0.080	0.429	0.070
wom.	coup no kids	0.498		0.497		0.627		0.639					
	coup w/1 kid	0.520		0.530		0.572		0.608		0.621		0.628	
	coup w/2 kids	0.420		0.460		0.435		0.458		0.559		0.568	
	coup w/3 kids	0.293		0.362		0.324		0.332		0.506		0.456	
	coup w/4 kids	0.121		0.202		0.159		0.183		0.381		0.237	
kids	coup w/1 kid	0.053	0.009	0.050	0.008	0.046	0.011	0.049	0.010	0.051	0.017	0.043	0.013
	coup w/2 kids	0.146	0.016	0.129	0.014	0.129	0.022	0.138	0.016	0.111	0.027	0.108	0.021
	coup w/3 kids	0.231	0.027	0.205	0.024	0.198	0.033	0.218	0.028	0.156	0.038	0.181	0.034
	coup w/4 kids	0.406	0.051	0.354	0.044	0.343	0.063	0.369	0.051	0.247	0.072	0.334	0.068

Notes: Table presents selected parameters from each estimated equation system. Estimates based on Canadian price and expenditure data, from FAMEX 1969-1996 and Browning and Thomas (1999).

In the rightmost columns of Table 2, we present estimates from models using data on only households with 1-4 children, excluding childless couples, which is the weakest information environment supported by the identification results in Theorem 1.

In the middle column of Table 2, we include couples without children in the model. This is valid if we assume that married men and women without children have the same preferences (individual utility functions) as those with children, though they need not have the same resource sharing rules η or economies of scale parameters A as couples with children. This same preferences of adults is not necessary for our identification, but if this assumption holds then it should improve the precision of our estimates by providing information on adults preferences that do not need to be disentangled from children's demand functions.

The first column of Table 2 includes data on single men and single women as well, which is valid if the individual utility functions of men or women living alone are the same as those of men or women who live with a spouse. This even stronger assumption is also not necessary for identification, but if it holds then using singles data provides further information for distinguishing men's demands from women's demands within households. By comparing the columns in Table 2, we can assess the gains in precision from these strengthened assumptions and associated data.

Consider the first whether the additional flexibility offered by the QAI model versus AI is important for estimating resource shares. Likelihood ratio tests strongly favor the QAI model: the LRT test statistics for the the AI restrictions against a QAI alternative are 1918, 337 and 135 for models including singles, excluding singles and excluding childless households, respectively. Thus, as found in many other papers, the quadratic terms in the QAI model are statistically significant.

However, in our data the resource share estimates remain almost the same regardless of whether or not the linearity restriction of the AI model is imposed. This is not a surprising result. Resource shares are essentially multiplicative scale parameters, so e.g. if the data on mother's shares are on average (for every price and total expenditure level) 10% higher in one household type than another, then we should expect to find that the fitted curve in the one household type is 10% higher than the fitted curve in the other type, regardless of whether the curve being fit is a best linear approximation or a best quadratic approximation to the data.

Of course we cannot guarantee that resource share estimates will always be robust in this way to specification error, but this result is at least suggestive that when we later apply Theorems 2 and 3 to identify resource shares from Engel curve data, we may well get reasonably accurate estimates of the shares even if we misspecify the Engel curve shapes. Note also that the range of total expenditure levels in Malawi is much smaller than in Canada, and much of detectable quadratic curvature in the Canadian data comes from including both rich and poor households in the model, so linear AI demands may well be adequate for modeling the Malawi data.

Next consider now how the estimates vary across whether or not singles or childless couples are included. The key feature that is notable here is that estimated resource shares differ quite substantially between the left-hand and middle panels of Table 2, but not very much between the middle and right-hand panels. That is, including singles changes the estimates, but including childless couples does not. This may be due to single adults having very different preferences from married adults, in which case it is unwise to pool them into the estimation. In contrast, we find no similar evidence that men and women living in childless couples have very different preferences from those living in families with children. Thus, pooling childless couples and couples with children to improve estimation precision may be acceptable.

Since the assumption that childless adults have the same preferences as adults with children is not necessary for identification, we will consider the right-hand panel as our preferred set of estimates. We see two important patterns evident in the results which exclude childless households. First, the estimated resource shares are weakly monotonic in the number of children in the household. For example, in the estimates which exclude all childless housholds, men's resource shares are about one-third of expenditure, and are roughly invariant to the number of children. In contrast, women's resource shares decline steeply with the number of children, and (the sum of) children's shares rise sharply with the number of children. In particular, children's resources are approximately 5 percent of household expenditure per child.

Second, it seems that as households get larger, and the total needs of children rise, and the resources allocated towards children are for the most part taken away from women. For example, comparing households with 1 and 3 children, we see that children's resources rise by about 10 percentage points, from approximately 5 per cent to 15 per cent of household expenditures, and that women's resources fall by just about the same amount. This pattern in the allocation of household resources can only be observed in a context where we can model children's resource shares, which illustrates the usefulness of our identification results.

5 Engel Curve Estimation

In this section, we estimate Engel curve systems in an environment without price variation using the identification result provided in Theorem 2. 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. Thus selection effects are of minimal concern 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-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 are conducted over a period of months which encompasses the wet and dry agricultural seasons. Auxilliary price data are used to standardize the food consumption expenditures across households, and we acknowledge that these prices may reflect transitory shocks and thus may be a source of measurement error (given the recall structure of the data).

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. 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.

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.[3] 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).

Our sample consists of 7731 households comprised of married couples with 0-4 children aged less than 15, of which 1404 households are childless couples which are used in only some specifications. 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. Table 3 gives summary statistics

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

[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] The World Bank imputed total expenditure for the IHS1 using Povmap.

for our sample.

Table 3: Data Means, Malawian micro-data		childless couples	couples with 1 child	2 children	3 children	4 children
Number of Observations		1404	2062	1914	1414	937
clothing share	women	1.77	1.53	1.49	1.30	1.18
(in per cent)	men	1.44	1.13	1.09	1.00	0.73
	children		0.75	1.05	1.20	1.43
footwear share	women	0.33	0.24	0.20	0.17	0.13
(in per cent)	men	0.33	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, and because have we many more observations than in the Canadian case, 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, e.g., Browning and Chiappori (1998), the presence of distribution factors may help identification of resource shares, but 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); the average age of children; the minimum age of children; and the proportion of children who are girls. We include 2 distribution factors: the difference in age between husband and wife; and the difference in years of education between husband and wife. We allow all demographic factors to affect the preferences of every household member.

We estimate models corresponding to individuals with log-linear Engel curves given by (4) and identified household Engel curves given by (3). Noting that log-linear individual Engel curves have slopes that do not depend on the resource share,

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

we can substitute these into (3) to get

$$\begin{aligned} W_{cs}(y) &= s\eta_{cs}\left(\alpha_{cs} + \beta_c \ln \eta_{cs}\right) + \eta_{cs}\beta_c \ln y, \\ W_{ms}(y) &= \eta_{ms}\left(\alpha_{ms} + \beta_m \ln \eta_{ms}\right) + \eta_{ms}\beta_m \ln y, \\ W_{fs}(y) &= \eta_{fs}\left(\alpha_{fs} + \beta_f \ln \eta_{fs}\right) + \eta_{fs}\beta_f \ln y. \end{aligned} \tag{6}$$

Here, Engel curves are linear in the log of expenditure, with intercepts that depend on everything, and slopes that are equal to the product of the resource share and the (latent) individual preference parameter β_t multiplied by log-expenditure. Rather than estimating a complex model which imposes the above nonlinear restrictions on the intercept term, we instead estimate index models of the following form:

$$W_{ts}(y) = a_{ts}(d, z_1, z_2) + \eta_{ts}(d, z_1, z_2)\beta_t(z_1, z_2) \ln y,$$

where d are the 2 distribution factors, z_1 are the 4 area-of-residence demographic variables, and z_2 are the 3 demographic variables describing the age and sex of children. We implement the model with linear indices for a_{ts}, η_{ts} and β_t. In all estimated models, a_{ts} depends on everything, but we vary what appears in η_{ts} and β_t across specifications.

Table 4 presents our results. We present estimates for models which vary in 3 dimensions. Models vary by: (1) whether or not they are based on 1 assignable (just clothing) or 2 assignables (both clothing and footwear); (2) whether or not they include childless couples; and (3) the list of variables included as distribution factors and demographic controls.

Table 4: Estimated Resource Shares, Malawian Engel Curves, Households with Children

		Clothing Only				Clothing and Footwear			
		without controls		with controls		without controls		with controls	
person	household	share	std err	share	std err	share	std err	share	std err
man	childless couple								
	couple w/1 kid	0.345	0.043	0.320	0.084	0.409	0.038	0.293	0.075
	couple w/2 kids	0.412	0.039	0.393	0.089	0.417	0.040	0.327	0.081
	couple w/3 kids	0.459	0.052	0.474	0.093	0.481	0.045	0.443	0.078
	couple w/4 kids	0.286	0.054	0.247	0.099	0.341	0.049	0.252	0.089
woman	childless couple								
	couple w/1 kid	0.439	0.039	0.484	0.091	0.385	0.033	0.505	0.070
	couple w/2 kids	0.320	0.044	0.362	0.096	0.279	0.034	0.381	0.077
	couple w/3 kids	0.314	0.043	0.329	0.091	0.246	0.035	0.290	0.069
	couple w/4 kids	0.346	0.052	0.345	0.100	0.270	0.039	0.359	0.073
child(ren)	couple w/1 kid	0.216		0.196		0.206		0.203	
	couple w/2 kids	0.268		0.245		0.303		0.292	
	couple w/3 kids	0.227		0.197		0.273		0.268	
	couple w/4 kids	0.367		0.408		0.389		0.389	
η contain		s		s,y		c		s,y	
β contain		t		t,z1,z2		t		t,z1,z2	

Notes: Table presents selected parameters from each estimated Engel curve system. Estimates based on Malawian expenditure data from IAHS1 and IAHS2 (World Bank).

Taken as a whole, the resource share estimates in Table 4 imply that children and (presumably) their mothers split roughly 50-60 per cent of household resources. Men's share of household resources is non-decreasing until the addition of the fourth child. Women, in contrast, see a near monotonic decrease in their share of household resources until the fourth child. This monotonicity does not seem to extend to the fourth child, but this could be due estimation imprecision. In Table 5 below we investigate the role of certain demographic and distribution factors in these estimates, and we find there that the share of resources devoted to children declines monotonically in per-child terms in nearly all cases.

Banerjee and Duflo (2008) examine household expenditure data for what they define as 'middle-class' households in 13 developing countries. They define middle-class in terms of per-capita household expenditure and describe differences in per-capita expenditure across lower- and upper-middle families in order to assess the role of the middle class in the development process.[4] Our findings suggest that such per-capita measures, and in particular, inter-generational comparisons based on them, omit an important margin, which is variation in intrahousehold resource shares across households of different sizes.

Our results also touch on a second strand of the development literature. Anderson and Baland (2002), in a survey of Kenyan households, report that married women in couples are more likely to use informal savings (rotating savings and credit associations) and propose a model of intra-household conflict between spouses to explain the difference in participation rates. Our results are supportive of this view. Our results are also consistent with the argument that married women have an incentive to participate in these informal

[4] Weber (1905) and Landes (1998), among others, argue that the middle class is central to economic growth – the middle class provides the human capital and physical capital necessary for growth.

savings institutions as insurance against fertility. While we lack sufficient data to assess this argument, we conjecture that participation rates likely fall with the number of children.

The implications of our results for development policy are twofold. First, increasing the level of household resources benefits all household individuals. While targeted programs, such as a school lunch program, have undoubted direct effects for children, simply increasing household income will similarly increase the resources devoted to children. Secondly, in terms of female poverty, our results indicate that, within the household, children are a consumption expense for women. To the extent that fertility decisions themselves are inefficient, then fertility is a consumption externality for women (and not a consumption choice). Hence, development programs aimed at fertility, such as sexual education, condom usage, etc. may help to reduce gender consumption inequality.

Finally, our results, while not definitive, suggest that the share of resources for children in Malawi are similar in magnitude (and perhaps even larger) than the share of resources for children in Canada. Although these results may not generalize, they do suggest that differences in child poverty do not result from differences in preferences regarding children. As well, our results may also shed some light on Sen's (1990) claim (echoed by Duflo (2005)) that 60-100 million women are missing in developing countries. Taken at face value, our finding that the women's share of household resources declines similarly with fertility in Canada and Malawi suggests that the explanation of the missing women may be relative incomes, not differences in preferences towards women.

Table 5: Estimated Resource Shares, Malawian Engel Curves, Clothing and Footwear									
		Child info in Resource Share only				Child info in Both			
		without childless		with childless		without childless		with childless	
person	household	share	std err	share	std err	share	std err	share	std err
man	childless couple			0.570				0.602	
	couple w/1 kid	0.386	0.103	0.444	0.083	0.364	0.110	0.530	0.082
	couple w/2 kids	0.462	0.135	0.492	0.106	0.448	0.146	0.592	0.112
	couple w/3 kids	0.594	0.165	0.602	0.133	0.583	0.181	0.695	0.141
	couple w/4 kids	0.469	0.209	0.499	0.165	0.446	0.230	0.591	0.179
	(hs age - wife age)/10	-0.054	0.048	-0.021	0.040	-0.058	0.050	-0.011	0.039
	(hs ed - wife yrsed)/10	-0.430	0.219	-0.399	0.200	-0.461	0.221	-0.421	0.200
	avg kid age -5	0.029	0.025	0.019	0.020	0.026	0.028	0.027	0.022
	min kid age -2	-0.033	0.026	-0.028	0.021	-0.030	0.029	-0.038	0.023
	prop girl kids	0.050	0.059	0.016	0.055	0.044	0.064	-0.055	0.058
woman	childless couple			0.430	0.092				
	couple w/1 kid	0.422	0.092	0.318	0.076	0.440	0.094	0.398	0.091
	couple w/2 kids	0.273	0.112	0.158	0.095	0.292	0.116	0.296	0.077
	couple w/3 kids	0.184	0.130	0.061	0.116	0.192	0.137	0.126	0.098
	couple w/4 kids	0.224	0.150	0.084	0.130	0.237	0.157	0.024	0.122
	(hs age - wife age)/10	0.074	0.046	0.023	0.035	0.076	0.047	0.037	0.138
	(hs ed - wife yrsed)/10	0.267	0.180	0.195	0.166	0.290	0.182	0.018	0.035
	avg kid age -5	-0.020	0.018	-0.031	0.016	-0.018	0.019	0.220	0.171
	min kid age -2	0.017	0.017	0.025	0.016	0.014	0.019	-0.036	0.017
	prop girl kids	0.044	0.052	0.054	0.047	0.036	0.053	0.028	0.017
kids	couple w/1 kid	0.192		0.238		0.196		0.071	
	couple w/2 kids	0.265		0.350		0.260		0.112	
	couple w/3 kids	0.222		0.337		0.224		0.180	
	couple w/4 kids	0.307		0.416		0.317		0.385	
	(hs age - wife age)/10	-0.020		-0.002		-0.018		-0.026	
	(hs ed - wife yrsed)/10	0.163		0.204		0.170		0.403	
	avg kid age -5	-0.008		0.013		-0.008		-0.247	
	min kid age -2	0.016		0.003		0.016		0.073	
	prop girl kids	-0.095		-0.070		-0.081		0.026	
resource shares	contain	c,y,z2				c,y,z2			
preferences	contain	t,z1				t,z1,z2			

Notes: Table presents selected parameters from each estimated Engel curve system. Estimates based on Malawian expenditure data from IAHS1 and IAHS2 (World Bank).

In Table 5, we examine the robustness of our estimated resource shares to different specifications of the way in which demographic information enters the model. Effectively we are confronted by a set of discrete choices – should we include demographic information for children in the resource and preferences or just in resource shares? We do not know the extent to which children's utility functions vary by gender and age. To assess this, our models in Table 5 examine how the resource share estimates change across different model specifications. In general, the results of this robustness exercise imply that there is not much difference in the estimates across specifications. That gender differences in children do not affect resource shares for children is consistent with Deaton (1989 and 1997) who reports a similar conclusion for households in Cote d'Ivoire and India, respectively.[5] The main differences appear in the final column when we include childless couples and demographic information in preferences. In this specification, the resource shares of both men and children are no longer monotonic in the number of children while the resource shares for women are – indeed, the resource shares estimates for women suggest they receive an even smaller share of household resources. Thus, our main conclusion from this robustness exercise is not to select a particular specification but rather to argue that our findings are robust across specifications.

6 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. In this paper we propose and estimate a collective household model with the identification of children's resource shares.

Using household consumption data for Canada and Malawi, we find that men's share of household resources is relatively unaffected by the presence of children. The implication then, is that women and children essentially split a fixed share of roughly 50 to 60 per cent of household resources. Our estimates also suggest that there are not sizeable differences in children's resource shares between Canada and Malawi, despite the gross disparity in income.

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. Finally, we find that children affect primarily the resource share of women and that fertility and intra-household inequality are linked.

7 Appendix: Theorems and Proofs

Define inverse expenditure scaled prices

$$r^k = \frac{y}{p^k} \text{ and } r = \left(r^1, ..., r^K\right)'. \tag{7}$$

Let $h_t(r)$ denote the Marshallian demand function associated with the utility function $U_t(x_t)$, so an individual that chooses x_t to maximize $U_t(x_t)$ under the usual linear budget constraint $p'x_t = y$ would choose $x_t = h_t(r)$. For their identification, BCL assumed that for a person of type t, $U_t(x_t)$ was the utility function of a single person of type t living alone, and so $h_t(r)$ would be that single person's observed demand functions over goods. We do not make this assumption.

Using this r notation which will simplify later results, equation (1) can be written as

$$\max_{x_f, x_m, x_c, z_s} \widetilde{U}_s \left[U_f\left(x_f\right), U_m\left(x_m\right), U_c\left(x_c\right), r\right] \quad \text{such that} \quad z_s = A_s\left[x_f + x_m + x_c\right] \text{ and } 1 = \Sigma_{k=1}^K z_s^k / r^k \tag{8}$$

[5] Duflo (2005) argues that girls are treated differentially from boys during periods of household distress, such as droughts or disease epidemics. Our results shed no light under this stone.

The demand functions for the household s arising from the household's maximization problem, equation (8), can be written as follows. Let A_s^k denote the column k vector of the matrix A_s and let $\widetilde{A}_s^k$ denote the row vector given by the k'th row of the matrix A_s. Define

$$r_s^k = \frac{y}{p' A_s^k} \text{ and } r_s = \left(r_s^1, ..., r_s^K\right)'. \tag{9}$$

This r_s will turn out to be the vector of inversed Lindahl shadow prices for household type s, analogous to the vector of inversed expenditure scaled market prices r. Define $H_s^k(r)$ to be the demand function for each good k in a household with s children. We could also include A_s as an argument in η_{ts}, but this is superflous since the s subscript shows that η_{ts} can already vary in arbitrary ways with s, including any dependence on A_s. Including r_s as an argument in H_s^k is similarly unnecessary, since r_s can be expressed as a function of r and A_s. 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

$$z_s^k = H_s^k(r) = \widetilde{A}_s^k \left[h_f\left(r_s\eta_{fs}\right) + h_m\left(r_s\eta_{ms}\right) + s h_c\left(r_s\eta_{cs}\right)\right] \tag{10}$$

where η_{ts} denotes the resource share of a person of type t in a household with s children. Note that each child gets a share η_{cs}, so the total share devoted to children is $s\eta_{cs}$. By definition, resource shares must sum to one, so for any s

$$\eta_{fs} + \eta_{ms} + s\eta_{cs} = 1 \tag{11}$$

The resource shares η_{ts} 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).

Our first assumption is that the BCL model as described above holds for households with one, two, or three children, that is,

ASSUMPTION A1: Equations (8), (10), and (11) hold.

BCL show generic identification of their model by assuming the demand functions of single men, single women, and married couples (that is, the functions $H_0(r)$, $h_f(r)$, and $h_m(r)$) 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 wormen we cannot observe demand functions for children living alone. We instead will obtain identification by looking at households with varying numbers of children.

For our results, By Assumption A1 we assume that married adults with one child have the same utility functions over goods as adults with two or three children, which we feel is much less restrictive than the BCL assumption that single adults have the same utility functions as adults living as married couples. We also do not assume that childless adults have the preferences as adults with children. In both our results and BCL, households of different composition all vary in their consumption technologies and resource sharing rules.

7.1 Types of Goods

Definition: A good k is a private good if the vector A_s^k has a one in row k and has all other elements equal to zero, and if $\widetilde{A}_s^k = A_s^{k\prime}$.

This is equivalent to the definition of a private 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 adult good if $U_c(x_c)$ does not depend on x_c^k, and a good k is a child good if $U_f(x_f)$ and $U_m(x_m)$ do not depend on x_f^k and x_m^k.

By this definition, an adult good is a good that appears only in the utility functions of adults (individual types m and f), and a define a child good is a good that appears only in the utility functions of children. Adult goods are not necessarily assignable, since both types of parents could consume them.

ASSUMPTION A2: Assume that the demand functions include a private child good, denoted as good c, a private adult good which we will denote as good j, and another private good which we will denote as good i.

Our identification results will only require observing the demand functions for the three goods listed in Assumption A2. Examples of private child goods could be toys or children's clothes, while examples of private adult goods could be alcohol, tobacco, or men's and women's clothing. Private or assignable goods are often used in this literature to obtain identification, or to increase estimation efficiency.

7.2 Identification With Price Variation

We now provide our first identification result, which assumes we have data from multiple price regimes, so household demands as functions of prices and total expenditures can be observed.

ASSUMPTION A3: The household demand functions $H_s^k(r)$ are identified, for $s = 1, 2, 3$ and for the three goods indexed by $k = i, j, c$.

As with ordinary revealed preference theory, we assume we can observe the functions $H_s^k(r)$, corresponding to the observed purchased bundles of goods by households having s equal to one, two, or three children, in varying price and total expenditure regimes. Our identification will only require observing the demand functions for the three goods we index by i, j, and c in Assumption A2. Based on these observed household demand functions, our primary goal will be identification of the child resource shares η_{sc}.

ASSUMPTION A4: Assume that resource shares η_{ts} are independent of r.

Assumption A4 imples that the resource shares are independent of prices p and total household expenditures y. Note that this still allows resource shares to depend on related measures such as the income, wages, or savings of the household members, and they only need to be independent of p and y after conditioning on these other measures. Later we will provide alternative assumptions that permit the resource shares to vary with p, though in that case we will only identify the shares in one price regime. Lewbel and Pendakur (2008) also assume resource shares do not depend on total expenditures y, and most empirically implemented models of collective households do not include p in η_{ts}.

Finally, we require some technical assumptions. Recall that $h_t^k(r)$ denotes the demand function of individual t for good k. Let $h_t^{kj}(r) = \partial h_t^k(r) / \partial r^j$ and $h_t^{kji}(r) = \partial h_t^{kj}(r) / \partial r^i$. Also let h_{t0}^k denote $h_t^k(0)$, h_{t0}^{ki} denote $h_t^{kj}(0)$, and h_{t0}^{kji} denote $h_t^{kji}(0)$. Similarly, for household demands let $H_s^{kj}(r) = \partial H_s^k(r) / \partial r^j$ and $H_{s0}^k = H_s^k(0)$, etc., so, e.g., $H_{20}^{cc} = H_2^k(0)$ is the derivative with respect to r^c of the demand function for the children's good of a household that consists of a mother, father, and two children.

ASSUMPTION A5: Assume $h_t^k(r)$ is twice continuously differentiable in the neighborhood of $r = 0$, with finite first and second derivatives, for $k = c, j, i$ and $t = m, f, c$.

ASSUMPTION A6: At least one of the following four sets of conditions (i), (ii), (iii), or (iv), hold:

(i) $h_m^{jj} \neq h_f^{jj}$, $h_m^{jj} \neq 0$, $h_m^{ji} \neq h_f^{ji}$, $h_m^{ji} \neq 0$, $h_m^{jj} / \left(h_f^{jj} - h_m^{jj}\right) \neq h_m^{ji} / \left(h_f^{ji} - h_m^{ji}\right)$, and the following matrix is nonsingular

$$\begin{pmatrix} H_{30}^{cc}/H_{10}^{cc} & H_{30}^{jj} & H_{30}^{ji} \\ H_{20}^{cc}/H_{10}^{cc} & H_{20}^{jj} & H_{20}^{ji} \\ 1 & H_{10}^{jj} & H_{10}^{ji} \end{pmatrix}. \tag{12}$$

(ii) Good i is an adult good, $h_m^{jj} \neq h_f^{jj}$, $h_m^{jj} \neq 0$, $h_m^{ii} \neq h_f^{ii}$, $h_m^{ii} \neq 0$, $h_m^{jj} / \left(h_f^{jj} - h_m^{jj}\right) \neq h_m^{ii} / \left(h_f^{ii} - h_m^{ii}\right)$, and the following matrix is nonsingular

$$\begin{pmatrix} H_{30}^{cc}/H_{10}^{cc} & H_{30}^{ij} & H_{30}^{ii} \\ H_{20}^{cc}/H_{10}^{cc} & H_{20}^{ij} & H_{20}^{ii} \\ 1 & H_{10}^{ij} & H_{10}^{ii} \end{pmatrix}. \tag{13}$$

(iii) $h_m^{jj} = h_f^{jj}$, $h_m^{jj} \neq 0$, and the following matrix is nonsingular:

$$\begin{pmatrix} H_{20}^{cc}/H_{10}^{cc} & H_{20}^{jj} \\ 1 & H_{10}^{jj} \end{pmatrix}. \tag{14}$$

(iv) Assume good i is an adult good, $h_f^{jj} = 0$, $h_m^{jj} \neq 0$, $h_m^{ii} = 0$, $h_m^{ii} \neq 0$ and the matrix (13) is nonsingular.

Conditions (i), (ii), (iii), or (iv) in Assumption A6 are alternatives; only one of these four conditions is required to hold. Identification will not require knowing which one holds, but only that at least one does. Condition (iii) does not require good i, conditions (i), (ii), and (iii) all use the fact that good i is private (from Assumption A2) and consumed by adults, but only (ii) or (iv) requires it to be an exclusively adult good.

Identification requires some differences between the demand functions of adults and children, and some nonlinearity. These requirements are most easily described in terms of the shapes of the demand functions for some goods in the neighborhood of zero. For Assumption A6, note that in any price regime $r = 0$ when total expenditures $y = 0$. However, by continuity of demand functions, we do not literally need to observe the demand functions of households with zero expenditures. Identification will generally hold even if we only observe the demands of households far away from zero, but sufficient technical conditions analogous to Assumption A6 appear to be much more difficult to construct there.

A technicality is that it is possible for well behaved demand functions to have derivatives that go to infinity as y goes to zero, violating Assumption A5, for example, Assumption A5 could be violated by a demand system that is quadratic in y. In this case Theorem 1 below can be modified by dividing the demand functions by power functions y before evaluating them at $y = 0$.

Turning to Assumption A6, having $h_m^{ji} \neq h_f^{ji}$ and $h_m^{ji} \neq 0$ means that men consume the adult good j, and their usage of the good varies with the price of good i at low income levels, and that women have a different price elasticity for the good j than men do at low income levels. Similarly, $h_m^{jj} \neq h_f^{jj}$ and $h_m^{jj} \neq 0$ means that men's usage of good j has a nonzero own price elasticity at low income levels, and women have a different own price elasticity for good j. For example, suppose j is adult clothing and i is tobacco. Then one way these conditions will all hold is if women don't smoke, men's clothing purchases depends in part on the price of tobacco, and men and women have different own price elasticites for clothing, which together imply the h restrictions in condition (i). Another simpler possibility, which satisfies the corresponding restrictions in condition (ii) instead, is if women don't smoke, men have a nonzero own price elasticity for smoking, and men and women have different own price elasticities for clothing. Or if j is alcohol and i is men's clothing, then (ii) holds if men and women have different, nonzero own price elasticities for alcohol and men have a nonzero own price elasticity for clothes.

Condition (iii) is simpler still (though perhaps more restrictive), since it only requires that there exist some private adult good j for which both women and men have the same nonzero own price elasticity near zero.

Perhaps the simplest of all is condition (iv). This will hold if goods i and j have nonzero own price elasticities and are assignable to adult women and men respectively

All the alternative A6 assumptions also impose nonsingularity of a matrix. This essentially requires some variation in demand price derivatives across households of different sizes. Note that each of these alternatives provides sufficient but far from necessary conditions for identification. These are just four of potentially many other possible sets of nonlinearities and variations that suffice.

THEOREM 1: Let Assumptions A1, A2, A3, A4, A5, and A6 hold. Then children's resource shares η_{cs} are identified.

Since the household of type s has s children, the total share of resources devoted to children is identified from Theorem 1 as $s\eta_{cs}$, and the total parent's share is $1 - s\eta_{cs}$. Under either mild additional assumptions or no additional assumptions (depending on which version of Assumption A6 holds), the separate shares of the husband and wife are also identified. For example, under Assumption A6(iv), $1/h_f^{ii}$ is identified from the same matrix inversion that yields η_{c1}, and η_{fs} is then identified from $\eta_{fs} = \mu_s^{ii}/h_f^{ii}$ in this case.

PROOF OF THEOREM 1: It follows immediately from equations (9), (10) and the definition of private goods, that for any private goods k and i,

$$H_s^k(r) = h_f^k(r_s\eta_{fs}) + h_m^k(r_s\eta_{ms}) + sh_c^k(r_s\eta_{cs}) \tag{15}$$

and, given Assumption A4,

$$H_s^{ki}(r) = \eta_{fs}h_f^{ki}(r_s\eta_{fs}) + \eta_{ms}h_m^{ki}(r_s\eta_{ms}) + s\eta_{cs}h_c^{ki}(r_s\eta_{cs})$$

Evaluating these derivatives in the neighborhood of $y = 0$, which implies $r = 0$ and $r_s = 0$, gives

$$H_{s0}^{ki} = \eta_{fs}h_{f0}^{ki} + \eta_{ms}h_{m0}^{ki} + s\eta_{cs}h_{c0}^{ki}. \tag{16}$$

By Assumption A3, $H_s^{ki}(r)$ and so H_{s0}^{ki} are identified. By Assumption A2, for the goods i, j, and c equation (16) gives

$$H_{s0}^{ji} = \eta_{fs}h_{f0}^{ji} + \eta_{ms}h_{m0}^{ji}, \quad H_{s0}^{jj} = \eta_{fs}h_{f0}^{jj} + \eta_{ms}h_{m0}^{jj}, \tag{17}$$

and

$$H_{s0}^{cc} = s\eta_{cs}h_{c0}^{cc}. \tag{18}$$

Equation (18) yields

$$\eta_{cs} = \frac{H_{s0}^{cc}}{sH_{10}^{cc}}\eta_{c1} \tag{19}$$

and by equations (17), (11), and (19) we have

$$\begin{aligned} H_{s0}^{ji} &= \eta_{fs}h_{f0}^{ji} + \left(1 - \eta_{fs} - s\eta_{cs}\right) h_{m0}^{ji} \\ &= \left(h_{f0}^{ji} - h_{m0}^{ji}\right)\eta_{fs} + h_{m0}^{ji}\left(1 - s\eta_{cs}\right) \\ &= \left(h_{f0}^{ji} - h_{m0}^{ji}\right)\eta_{fs} + h_{m0}^{ji}\left(1 - \frac{H_{s0}^{cc}}{H_{10}^{cc}}\eta_{c1}\right) \end{aligned}$$

Now assume that Assumption A6i holds. The above equation can be solved for η_{fs} to obtain

$$\eta_{fs} = \frac{1}{h_{f0}^{ji} - h_{m0}^{ji}}H_{s0}^{ji} - \frac{h_{m0}^{ji}}{h_{f0}^{ji} - h_{m0}^{ji}}\left(1 - \frac{H_{s0}^{cc}}{H_{10}^{cc}}\eta_{c1}\right)$$

and the same derivation replacing i with j yields

$$\eta_{fs} = \frac{1}{h_{f0}^{jj} - h_{m0}^{jj}}H_{s0}^{jj} - \frac{h_{m0}^{jj}}{h_{f0}^{jj} - h_{m0}^{jj}}\left(1 - \frac{H_{s0}^{cc}}{H_{10}^{cc}}\eta_{c1}\right).$$

Equating these two expressions for η_{fs} gives

$$\frac{1}{h_{f0}^{ji} - h_{m0}^{ji}}H_{s0}^{ji} - \frac{h_{m0}^{ji}}{h_{f0}^{ji} - h_{m0}^{ji}}\left(1 - \frac{H_{s0}^{cc}}{H_{10}^{cc}}\eta_{c1}\right) = \frac{1}{h_{f0}^{jj} - h_{m0}^{jj}}H_{s0}^{jj} - \frac{h_{m0}^{jj}}{h_{f0}^{jj} - h_{m0}^{jj}}\left(1 - \frac{H_{s0}^{cc}}{H_{10}^{cc}}\eta_{c1}\right)$$

which simplifies to

$$\left(\frac{h_{m0}^{jj}}{h_{f0}^{jj} - h_{m0}^{jj}} - \frac{h_{m0}^{ji}}{h_{f0}^{ji} - h_{m0}^{ji}}\right)\left(1 - \frac{H_{s0}^{cc}}{H_{10}^{cc}}\eta_{c1}\right) = \frac{1}{h_{f0}^{jj} - h_{m0}^{jj}}H_{s0}^{jj} - \frac{1}{h_{f0}^{ji} - h_{m0}^{ji}}H_{s0}^{ji}.$$

Define

$$\rho_1 = \frac{1}{h_{f0}^{jj} - h_{m0}^{jj}} \Big/ \left(\frac{h_{m0}^{jj}}{h_{f0}^{jj} - h_{m0}^{jj}} - \frac{h_{m0}^{ji}}{h_{f0}^{ji} - h_{m0}^{ji}}\right) \text{ and } \rho_2 = \frac{-1}{h_{f0}^{ji} - h_{m0}^{ji}} \Big/ \left(\frac{h_{m0}^{jj}}{h_{f0}^{jj} - h_{m0}^{jj}} - \frac{h_{m0}^{ji}}{h_{f0}^{ji} - h_{m0}^{ji}}\right)$$

which by Assumption A6i are nonzero. Then

$$1 = \frac{H_{s0}^{cc}}{H_{10}^{cc}}\eta_{c1} + H_{s0}^{jj}\rho_1 + H_{s0}^{ji}\rho_2.$$

Evaluating this expression for households having s equal 1, 2, and 3 gives

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} H_{30}^{cc}/H_{10}^{cc} & H_{30}^{jj} & H_{30}^{ji} \\ H_{20}^{cc}/H_{10}^{cc} & H_{20}^{jj} & H_{20}^{ji} \\ 1 & H_{10}^{jj} & H_{10}^{ji} \end{pmatrix} \begin{pmatrix} \eta_{c1} \\ \rho_1 \\ \rho_2 \end{pmatrix}$$

and since this matrix is invertible this can be solved for η_{c1}, ρ_1, and ρ_2

$$\begin{pmatrix} \eta_{c1} \\ \rho_1 \\ \rho_2 \end{pmatrix} = \begin{pmatrix} H_{30}^{cc}/H_{10}^{cc} & H_{30}^{jj} & H_{30}^{ji} \\ H_{20}^{cc}/H_{10}^{cc} & H_{20}^{jj} & H_{20}^{ji} \\ 1 & H_{10}^{jj} & H_{10}^{ji} \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Everything on the right side of the above equation is identified by revealed preference, so the variables on the left are also identified, and in particular, the resource share of the child in one child households', η_{c1}, is identified. This can then be substituted in equation (19) to obtain the resource share of each child, η_{cs} in household's with any number of children s.

The derivation based on A6ii is the same as for A6i above, just replacing ji superscripts with ii everywhere.

Next, consider the case where Assumption A6iii holds. In that case equations (17), (11), and (19) produce

$$\begin{aligned} H_{s0}^{ji} &= \eta_{fs} h_{f0}^{ji} + (1 - \eta_{fs} - s\eta_{cs})\, h_{m0}^{ji} \\ &= h_{m0}^{ji}\,(1 - s\eta_{cs}) = h_{m0}^{ji}\left(1 - \frac{H_{s0}^{cc}}{H_{10}^{cc}}\eta_{c1}\right) \end{aligned}$$

so

$$1 = \frac{H_{s0}^{cc}}{H_{10}^{cc}}\eta_{c1} + \frac{H_{s0}^{jj}}{h_{m0}^{ji}}.$$

which for s equal to 1 and 2 gives

$$\begin{pmatrix} \eta_{c1} \\ 1/h_{m0}^{ji} \end{pmatrix} = \begin{pmatrix} H_{20}^{cc}/H_{10}^{cc} & H_{20}^{jj} \\ 1 & H_{10}^{jj} \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

which again identifies η_{c1}, and by equation (19) we get η_{cs} identified for any s.

Finally, consider the case where Assumption A6iv holds. In that case equations (17), (11), and (19) produce $H_{s0}^{ii} = \eta_{fs} h_{f0}^{ii}$ and

$$H_{s0}^{jj} = \left(1 - \eta_{fs} - \frac{H_{s0}^{cc}}{H_{10}^{cc}}\eta_{c1}\right) h_{m0}^{jj}$$

so

$$1 = \frac{H_{s0}^{cc}}{H_{10}^{cc}}\eta_{c1} + \frac{H_{s0}^{jj}}{h_{m0}^{jj}} + \frac{H_{s0}^{ii}}{h_{f0}^{ii}}.$$

For s equal 1, 2, and 3 this gives

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} H_{30}^{cc}/H_{10}^{cc} & H_{30}^{jj} & H_{30}^{ii} \\ H_{20}^{cc}/H_{10}^{cc} & H_{20}^{jj} & H_{20}^{ii} \\ 1 & H_{10}^{jj} & H_{10}^{ii} \end{pmatrix} \begin{pmatrix} \eta_{c1} \\ 1/h_{m0}^{jj} \\ 1/h_{f0}^{ii} \end{pmatrix}$$

which again identifies η_{c1} and hence η_{cs}

7.3 Identification Without Price Variation

We have shown nonparametric identification of childrens' resource shares. To reduce data requirements and simplify empirical application, we now consider semiparametric specifications. In particular, by parameterizing how demand functions depend upon y we provide identification of childrens' resource shares in data without price variation. These results may be more convenient in application because they can be used with cross section data that does not possess price variation (instead of panels or repeated cross sections that provide price variation), and because they do not require identifying and estimating the dependence of demand functions on prices. Essentially, implementing these results only requires estimating Engel curves rather than full Marshallian demand functions.

For this identification we continue to let Assumptions A1 above hold, so we are still in the same collective model framework, but now introduce a set of alternative identifying assumptions. These alternative assumptions are more restrictive than those in Theorem 1, primarily in imposing functional form restrictions regarding y and semiparametric restrictions regarding p and A_s. However, these assumptions are less restrictive in some other ways, in particular, they permit resource shares η_{ts} to vary with prices and they only require observation of household Engel curves instead of full demand functions.

ASSUMPTION B2: Assume that the demand functions include a private child good, denoted as good c, A private male adult good which we will denote as good m, and a private female adult good which we will denote as good f.

One possible choice of private goods satisfying Assumption B2 could be men's, women's, and children's clothing.

Consider a particular price regime $p = p_0$. Define the budget share Engel curves

$$G_s^k(y) = \frac{p_0^k}{y} H_s^k \left(\frac{y}{p_0^1}, \ldots \frac{y}{p_0^K} \right)$$

So $G_s^k(y)$ is the fraction of total expenditures y that is spent buying good k by a household with s children facing prices p_0.

ASSUMPTION B3: For goods $k \in \{c, m, f\}$ the household budget share Engel curve functions $G_{s_o}^k(y)$ are identified.

ASSUMPTION B4: Resource shares η_{ks} are independent of y.

ASSUMPTION B5: For $k \in \{c, m, f\}$, there exists a function α_k such that $\alpha_k(p_0) \neq 0$ and either a function β_k such that

$$h_k^k(r_s) / r_s^k = \alpha_k(p) \ln y + \beta_k(p, A_s) \tag{20}$$

or a function δ_k such that $\delta_k(p, A_s, 0) = 0$ and

$$h_k^k(r_s) / r_s^k = \alpha_k(p) + \delta_k(p, A_s, y). \tag{21}$$

Assumptions B2, B3, and B4 are analogous to Assumptions A2, A3 and A4, but are strictly weaker since they do not require identification of how observable household demand functions vary with prices and they do not impose the condition that the resource shares not depend on prices. The price to be paid for this added generality is Assumption B5, which imposes functional restrictions on the demands for three private goods, but permits many common demand model specifications including polynomials of arbitrary degree in y. Construction of demand functions that satisfy Assumption B5 is discussed later in Theorem 3.

ASSUMPTION B6: Define $\tau_s^k = \eta_{ks}/\eta_{k1}$ for $k \in \{c, m, f\}$. Assume the six by six matrix in equation (22) in the appendix, which has elements depending on τ_s^k, is nonsingular.

The identification theorem first obtains identification of τ_s^k, and then uses Assumption B6 to recover estimates of resource shares η_{ks}. This assumption can be relaxed if we observe households of more than three sizes.

THEOREM 2: Let Assumptions A1, B2, B3, B4, B5, and B6 hold. Then children's and adult's resource shares η_{cs}, η_{ms}, and η_{cs} are identified.

PROOF OF THEOREM 2: First suppose equation (20) holds. By equations (10), (20), and Assumption A2,

$$\frac{z_s^c}{r^c} = \frac{H_s^c(r)}{r^c} = \frac{s h_c^c(r_s)}{r_s^c \eta_{cs}} = \alpha_c(p)\, s + \delta_c\left(p, A_s, \eta_{cs} y\right) s$$

so by the definition of $G_s^c(y)$,

$$G_s^c(y) = \alpha_c(p_0)\, s\eta_{cs} + \delta_c\left(p_0, A_s, \eta_{cs} y\right) s\eta_{cs}$$

and by Assumption B5

$$G_s^c(y_0) = \alpha_c(p_0)\, s\eta_{cs}$$

Define $\tau_s^c = G_s^c(y_0) / \left[s G_1^c(y_0)\right]$, which is identified because $G_s^c(y)$ is identified. Then $\tau_s^c = \eta_{cs}/\eta_{c1}$.

Similarly, for $k \in \{m, f\}$, by equations (10), (20), and Assumption A2,

$$\frac{z_s^k}{r^k} = \frac{H_s^k(r)}{r^k} = \frac{h_k^k(r_s)}{r_s^k \eta_{ks}} = \alpha_k(p) + \delta_k\left(p, A_s, \eta_{ks} y\right)$$

$$G_s^k(y) = \alpha_k(p_0)\, \eta_{ks} + \delta_k\left(p_0, A_s, \eta_{ks} y\right) \eta_{ks}$$

$$G_s^k(y_0) = \alpha_k(p_0)\, \eta_{ks}$$

and defining $\tau_s^k = G_s^k(y_0) / G_1^k(y_0)$ for $k \in \{m, f\}$ gives, $\tau_s^k = \eta_{ks}/\eta_{k1}$which is identified because $G_s^k(y)$ is identified.

Alternatively, if equation (21) holds instead of (20) in Assumption B5, then

$$\frac{z_s^c}{r^c} = \frac{H_s^c(r)}{r^c} = \frac{s h_c^c(r_s)}{r_s^c \eta_{cs}} = s\alpha_c(p) \ln\left(\eta_{cs} y\right) + s\beta_c(p, A_s)$$

so by the definition of $G_s^c(y)$,

$$\begin{aligned} G_s^c(y) &= \alpha_c(p_0)\, s\eta_{cs} \ln\left(\eta_{cs} y\right) + \beta_c(p_0, A_s)\, s\eta_{cs} \\ & \alpha_c(p_0)\, s\eta_{cs} \ln(y) + \alpha_c(p_0)\, s\eta_{cs} \ln\left(\eta_{cs}\right) + \beta_c(p_0, A_s)\, s\eta_{cs} \end{aligned}$$

and therefore

$$G_s^c(y) - G_s^c(1) = \alpha_c(p_0)\, s\eta_{cs}\left[\ln(y) - 1\right].$$

Now defining τ_s^c by $\tau_s^c = \left[G_s^c(y) - G_s^c(1)\right] / \left[s G_1^c(y_0) - s G_s^c(1)\right]$ for any value of y again gives $\tau_s^c = \eta_{cs}/\eta_{c1}$.

Similarly, for $k \in \{m, f\}$,

$$G_s^k(y) = \alpha_k(p_0)\eta_{ks}\ln(y) + \alpha_k(p_0)\eta_{ks}\ln(\eta_{ks}) + \beta_k(p_0, A_s)s\eta_{ks}$$

$$G_s^k(y) - G_s^k(1) = \alpha_k(p_0)\eta_{ks}[\ln(y) - 1]$$

and now defining τ_s^k by $\tau_s^k = \left[G_s^k(y) - G_s^k(1)\right] / \left[G_1^k(y_0) - G_s^k(1)\right]$ for any value of y again gives $\tau_s^k = \eta_{ks}/\eta_{k1}$ for $k \in \{m, f\}$.

Combined with equations (11), we now have with Assumption B5 holding based on either equation (21) or (20) that τ_s^c, τ_s^m, and τ_s^f are identified for any interger s and that $\eta_{c1}\tau_s^c - s\eta_{cs} = 0$, $\eta_{f1}\tau_s^f - \eta_{fs} = 0$, and $(1 - \eta_{c1} - \eta_{f1})\tau_s^m - (1 - s\eta_{cs} - \eta_{fs}) = 0$. These three equations for $s = 2$ and for any other integer $s \geq 3$ provide a total of six equations in six unknown resource shares, which in matrix form are

$$\begin{pmatrix} \tau_2^c & -1 & 0 & 0 & 0 & 0 \\ \tau_s^c & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \tau_2^f & -1 & 0 \\ 0 & 0 & 0 & \tau_s^f & 0 & -1 \\ -\tau_2^m & 1 & 0 & -\tau_2^m & 1 & 0 \\ -\tau_s^m & 0 & 1 & -\tau_s^m & 0 & 1 \end{pmatrix} \begin{pmatrix} \eta_{c1} \\ 2\eta_{c2} \\ s\eta_{cs} \\ \eta_{f1} \\ \eta_{f2} \\ \eta_{fs} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \tau_2^m - 1 \\ \tau_s^m - 1 \end{pmatrix} \tag{22}$$

Given nonsingularity of this matrix, this equation can be solved for the resource shares η_{cs} for children and η_{fs} for women, and using equations (11) yields the resource shares for men. If we observe households of more than three sizes, then we may replace τ_2^k and η_{k2} (for $k = c, f$) in equation (22) with τ_s^k and η_{ks} (for $k = c, f$) for other values of s like 3, 4, etc.,. to obtain additional equations that may be combined with equation (22) to obtain identification.

Assumption B5 of Theorem 2 requires demand systems to satisfy some properties. Theorem 3 below provides general classes of preferences (specifically, indirect utility functions) that yield demand systems satisfying the conditions of Assumption B5. We first require some more definitions. Recall that the vector of Marshallian demand functions $h_t(r)$ is defined by saying that if an individual of type t chose x_t to maximize $U_t(x_t)$ under the usual linear budget constraint $r'x_t = 1$, then that individual would choose $x_t = h_t(r)$. The indirect utility function associated with $U_t(x_t)$ is then defined as the function $V_t(r) = U_t(h_t(r))$.

Also let $\widetilde{r}$ denote the vector of all the elements of r except r^c, r^m, and r^f; and let $\widetilde{r}_s$ denote the vector of all the elements of r_s except r_s^c, r_s^m, and r_s^f.

ASSUMPTION C1: For $t \in \{m, f, c\}$ let

$$V_t(r_s) = \psi_t[v_t(r_s), \widetilde{r}_s] \tag{23}$$

where ψ_t is differentiable, strictly monotonically increasing in its first argument, and homogeneous of degree zero in its remaining (vector valued) argument $\widetilde{r}$.

In Theorem 3 below, the function $v_t(r_s)$ will be an indirect utility function that is constructed to obtain required restrictions on the demand functions for the private goods $k \in \{m, f, c\}$. The role of ψ_t is to permit far more general functional forms for all the other goods the household consumes.

Theorem 3 will require either Assumption C2 or Assumption C3 below to hold, depending on whether Assumption B5 will be satisfied by equation (20) or (21).

ASSUMPTION C2: For $t \in \{m, f, c\}$, assume

$$v_t(r_s) = e^{\lambda_t' \ln(r_s) + a_t(\widetilde{r}_s)} \ln B_t(r_s) \tag{24}$$

where λ_t is a K vector of constants that sum to zero, $a_t(\widetilde{r})$ (which could be identically zero) is a homogeneous of degree zero function, and $B_t(r)$ is a homogeneous of degree one function.

The function v_t in Assumption C2 can be described as

$$v_t(r) = -e^{a_t^*(r)} \ln B_t(r) \tag{25}$$

where $a_t^*(r) = \lambda_t' \ln(r) + a_t(\widetilde{r})$. Apart from this last equality, equation (25) with $a_t^*(r)$ homogeneous of degree zero and $B_t(r)$ homogeneous of degree one equals the general form of the PIGLOG (price independent generalized logarithmic) indirect utility function (see Muellbauer 1975). The most popular example of a PIGLOG demand system is Deaton and Muellbauer's (1980) Almost Ideal Demands (AID) which is the special case of equation (25) where $a_t^*(r) = \lambda_t' \ln(r)$ and $\ln B_t(r)$ has a certain quadratic form.

ASSUMPTION C3: For $t \in \{m, f, c\}$, assume

$$v_t(r_s) = e^{-B_t(r_s)} M_t\left(\frac{1}{\sum_{k=1}^{K} \lambda_t^k / r_s^k}\right) \tag{26}$$

where $B_t(r)$ is a homogeneous of degree zero function, $1/\left[\partial \ln M_t(y) / \partial \ln y\right]$ equals zero at $y = 0$, and λ_t and ζ_t are nonzero K vectors of constants that do not depend upon s having the property that $A_s \lambda_t = \zeta_t$.

The restriction $A_s \lambda_t = \zeta_t$ in Assumption C3 makes $M_t\left(1/\sum_{k=1}^{K} \lambda_t^k / r_s^k\right) = M_t\left(y/p'\zeta_t\right)$. One way this restriction can be satisfied is if the elements λ_t^k of λ_t all equal zero except for λ_t^c, λ_t^m, and λ_t^f. It will then follow from Assumption A2 that, with these restrictions, $\zeta_t = \lambda_t$ and M_t will depend only upon y and the prices p^c, p^m, and p^f. Alternatively, M_t can depend upon the prices of all goods if linear restrictions are imposed on A_s. Linear restrictions on economy of scale of consumption parameters analogous to these have appeared elsewhere in the literature, e.g., the model of Jorgenson, Lau, and Stoker (1983) imposes the constraint that logged Barten scales in their model sum to zero.

Assumption C3 also requires that the inverse log derivative of $M_t\left(y/p'\zeta_t\right)$ equal zero when $y = 0$. There are many functional forms that satisfy this requirement, e.g., letting $\ln M_t(y) = 1/\left(c_1 y + c_2 y^2 + ..c_L y^L\right)$ will make $1/\left[\partial \ln M_t(y) / \partial \ln y\right]$ be a polynomial in y which equals zero when $y = 0$. This example will correspond to having the demand functions for the private goods in equation (21) be polynomials of arbitrary degree in y, resulting in household Engel curves for the private goods that are polynomials in y. One can similarly define $M_t(y)$ to make household Engel curves for the private goods be polynomials in $\ln y$.

THEOREM 3: Let Assumptions A1, B2, B3, B4, C1, and either C2 or C3 hold. Then the demand functions arising from the indirect utility function $V_t(r_s)$ satisfy Assumption B5. Specifically, Assumption C2 yields equation (20) and Assumption C3 yields equation (21).

PROOF of THEOREM 3: Having ψ_t be homogeneous of degree zero in $\widetilde{r}_s$ means that $\psi_t[v_t(r_s), \widetilde{r}_s] = \psi_t\left[v_t(r_s), y\widetilde{r}_s\right]$. By the definition of $\widetilde{r}_s$, each element of the vector $y\widetilde{r}_s$ equals $1/p'A_s^k$ for some good $k \notin \{c, m, f\}$, and for all such k the expression $p'A_s^k$ does not depend upon p^c, p^m, and p^f, because those goods are private (note also that . We can therefore define Ψ_t by $\Psi_t\left[v_t(r_s), \widetilde{p}, A_s\right] = \psi_t[v_t(r_s), \widetilde{r}_s]$

where $\widetilde{p}$ is the vector of all prices p^k except p^c, p^m, and p^f. Applying Roy's identity for any good k then gives

$$h_t^k(r_s) = -\frac{\frac{\partial \Psi_t\left[\upsilon_t(r_s),\widetilde{p},A_s\right]}{\partial p^k} + \frac{\partial \Psi_t\left[\upsilon_t(r_s),\widetilde{p},A_s\right]}{\partial \upsilon_t(r_s)}\frac{d\upsilon_t(r_s)}{dp^k}}{\frac{\partial \Psi_t\left[\upsilon_t(r_s),\widetilde{p},A_s\right]}{\partial \upsilon_t(r_s)}\frac{d\upsilon_t(r_s)}{dy}}$$

which for $k \in \{c, m, f\}$ simplifies to

$$h_t^k(r_s) = -\left(\frac{d\upsilon_t(r_s)}{dp^k}\right) / \left(\frac{d\upsilon_t(r_s)}{dy}\right)$$

and recalling that $r_s^k = y/p^k$ for $k \in \{c, m, f\}$ yields

$$\frac{h_k^k(r_s)}{r_s^k} = \frac{h_k^k(r_s)\,p^k}{y} = -\left(\frac{\partial \upsilon_t(r_s)}{\partial \ln p^k}\right) / \left(\frac{\partial \upsilon_t(r_s)}{\partial \ln y}\right) \tag{27}$$

Now consider Assumptions C2 and C3 separately. First let Assumption C2 hold, and define $b_t(p) = B_t(1/p_1, ..., 1/p_n)$. By the assumed homogeneity of B_t, $B_t(r_s) = yb_t\left(A_s'p\right)$. In the same way Ψ_t was constructed from ψ_t by homogeneity of degree zero, $a_t(\widetilde{r_s})$ can be written as a function of just $\widetilde{p}$ and A_s. Let $\widetilde{a}_t(\widetilde{p}, A_s)$ equal this expression for $a_t(\widetilde{r_s})$ added to the sum of terms of the form $-\lambda_t'\left[\ln\left(p'A_s^k\right)\right]$ for all $k \notin \{c, m, f\}$. We may then without loss of generality rewrite $\upsilon_t(r_s)$ in equation (24) as

$$\upsilon_t(r_s) = e^{\widetilde{a}_t(\widetilde{p},A_s)-\lambda_{ct}\ln(p^c)-\lambda_{mt}\ln(p^m)-\lambda_{ft}\ln\left(p^f\right)}\left[\ln y + \ln b_t\left(A_s'p\right)\right]$$

which, when substituted into equation (27), gives for $k \in \{c, m, f\}$,

$$\frac{h_k^k(r_s)}{r_s^k} = \lambda_{kt}\ln y - \left[\lambda_{kt}\ln b_t\left(A_s'p\right) + \frac{\partial \ln b_t\left(A_s'p\right)}{\partial \ln p^k}\right] \tag{28}$$

which is in the form of equation (20), and so satisfies Assumption B5, with $\alpha_k(p) = \lambda_{kt}$ and $\beta_k(p, A_s)$ equal to the term in square brackets in equation (28).

Now consider the case where Assumption C3 holds instead of C2. Homogeneity of $B_t(r_s)$ and the restriction that $A_s\lambda_t = \zeta_t$ makes $\upsilon_t(r_s)$ in equation (26) satisfy

$$\upsilon_t(r_s) = e^{-b_t\left(A_s'p\right)}M_t\left(y/p'\zeta_t\right)$$

which, when substituted into equation (27), gives for $k \in \{c, m, f\}$,

$$\frac{h_k^k(r_s)}{r_s^k} = \frac{\zeta_t^k}{p'\zeta_t} + \frac{\partial b_t\left(A_s'p\right)}{\partial \ln p^k}\frac{m_t(y)}{p'\zeta_t} \tag{29}$$

where the function m_t is defined by $m_t(y) = 1/\left[\partial \ln M_t(y)/\partial \ln y\right]$. Equation (29) is in the form of equation (21) with $\alpha_k(p) = \zeta_t^k/p'\zeta_t$ and $\delta_k(p, A_s, y)$ equalling the rest of the right side of equation (29), and having the required property for Assumption B5 that $\delta_k(p, A_s, 0) = 0$.

Theorem 3 shows that a wide variety of models for preferences can satisfy the functional form requirements of Theorem 2. The examples of υ_t indirect utility functions in Theorem 3 all can be written in terms of just two functions of prices, and so by Lewbel (1993) must have rank two demand functions, however, the ψ_t function permits goods other than the three private goods to have higher rank. Moreover, despite the three private goods having rank two, the actual functional forms of engel curves for these goods can be quite general, with Theorem 3 permitting budget shares to be polynomials in y or $\ln y$. While Assumption C3 permits general polynomial functional forms for Engel curves, Assumption C2 for PIGLOG demands may be more attractive in that it does not impose such strong constraints on the A_s matrix or on the price functions, and identification based on Assumption C2 may be empirically stronger in the sense that it does not depend upon evaluating budget shares in the neighborhood of zero expenditures.

To apply Theorems 2 and 3, it is not necessary to completely specify the indirect utility functions υ_t and ψ_t. Since ψ_t only affects the demand functions of the nonprivate goods, it is not necessary to specify ψ_t at all if one is only going to estimate the demand functions of the private goods. It is also not necessary to specify exactly how υ_t varies with prices, since only the Engel curves arising from υ_t are used for estimation, not the entire demand function. Examples are the budget share models we described in section 2.

8 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??? (Crossley I presume?) 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" forthcoming in *Journal of Political Economy*.

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, E. (2005), “Gender Equality in Development,” mimeo.

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..

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.