Children's Resources in Collective Households: Identification, Estimation and an Application to Child Poverty in Malawi

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

June 2010

Abstract

The share of household resources devoted to children is hard to identify, because consumption is measured at the household level, and goods can be shared. Using semiparametric restrictions on individual preferences within a collective model, we identify how *total* household resources are divided up among household members, by observing how each family member's expenditures on a single private good like clothing varies with income and family size. Using data from Malawi we show how resources devoted to wives and children vary by family size and structure, and we find that standard poverty indices understate the incidence of child poverty.

JEL codes: D13, D11, D12, C31, I32. Keywords: Collective Model, Cost of Children, Bargaining Power, Identification, Sharing rule, Demand Systems, Engel Curves. We would like to thank Martin Browning, Pierre-Andre Chiappori, Federico Perali, and Frederic Vermeulen for helpful comments. Corresponding Author: Arthur Lewbel, Department of Economics, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA, 02467, USA. (617)-552-3678, lewbel@bc.edu, http://www2.bc.edu/~lewbel/

1 Introduction

Most measures of economic well-being rely, to some degree, on individual consumption. Yet the measurement of individual consumption in data is often confounded because consumption is typically measured at the household, not the individual, level. Dating back at least to Becker (1965, 1981), 'collective household' models are those in which the household is characterised as a collection of individuals, each of whom has a well-defined objective function, and who interact to generate household level decisions such as consumption expenditures. Given household data, useful measures of individual consumption expenditures are *resource shares*, defined as each member's share of total household consumption. If there is intra-household inequality, these resources shares will be unequal so standard per-capita calculations (assigning equal resource shares to all household members) are invalid measures of individual well-being.

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

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

Dauphin et al (2008) and Cherchye, De Rock and Vermeulen (2008) test whether observed household demand functions are consistent with children having separate utility functions. They find some evidence that households behave as if children do have separate utility functions. Cherchye, De Rock and Vermeulen (2010) consider estimation, but their method generally only yields bounds on resource shares.

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

Based on the collective household model of Chiappori (1988, 1992), a series of papers starting from

Bourguignon and Chiappori (1994), Browning, Bourguignon, Chiappori, and Lechene (1994), and Browning and Chiappori (1998) show identification of changes in resource shares as functions of *distribution factors*, defined as variables which affect bargaining power, but which do not affect preferences over goods or scale economies. 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 observable distribution factors for children will generally not exist. In contrast, we identify the levels of resource shares, and do not require distribution factors.

Some more recent collective household models can identify levels of (not just changes in) resource shares under some difficult to verify conditions (see, e.g., Chiappori and Ekeland 2008), but all the models in this class impose strong restrictions on how goods may be shared within households. Specifically, they assume that all goods are either purely private or purely public within the household. Browning, Chiappori and Lewbel (2007) (hereafter BCL) provide a model that nonparametrically identifies the levels of resource shares of all individual household members and which allows for very general forms of sharing of goods. We extend BCL, and so allow for these more general types of scale economies.

BCL show identification only when the demand functions of individuals can be separately observed, which is not the case for children since they are always in households along with adults. In practice, BCL observe the demand functions of individuals by observing data from single men and single women living alone, and combine those demand functions with data on the demands of men and women living together as (childless) couples. Accordingly, they assume very limited differences between the utility functions of single and married men and between those of single and married women. BCL also identify indifference scales, which are a variant of equivalence scales.

Lewbel and Pendakur (2008) propose some restrictions on BCL that permit identification of adult's resource shares and indifference scales in a numerically simpler model based on Engel curve data, without requiring observation of prices and price variation. Bargain and Donni (2009) extend the Lewbel and Pendakur (2008) model to identify the resource shares and indifference curves of children.

Our contribution is to extend the model of BCL to include children, semiparametrically identifying children's resource shares, employing a different identification strategy. Our identification assumes that resource shares do not vary with total expenditure (at least over a range of expenditure)[1], and assumes one of two semiparametric restrictions on individual preferences. With the first semiparametric restriction, we assume that preferences for a particular good are similar in certain limited ways across people (within household types), and use this similarity to help identify resource shares within households with a given number of children. In the second, we assume that a person's preferences for a particular good are similar across household types, and compare the consumption choices of people across households with varying

[1] Samuelson (1956) shows that resource shares cannot in general be constant for a large class of household social welfare functions. While we assume resource shares independent of household expenditures y, we do not require them to be constant, but instead allow them to vary arbitrarily with prices. So, for example, a welfare function that sums the utilities of individuals will satisfy our assumptions, making resource shares independent of y but not prices, if indirect utility functions are linear in y^k for any k (PIGL). We would like to thank Martin Browning for alerting us to this issue.

numbers of children.

In comparison with BCL (and with Lewbel and Pendakur 2008 and Bargain and Donni 2010), we do not need to use information on childless households (either couples or singles). In that respect, our identification strategies impose milder conditions on preference stability across household types, since e.g. we would assume that fathers of two children have similar preference to fathers of three children, rather than assume that either are similar to single men. Related identification ideas go back at least to Lazear and Michael (1988, chapter 4). We also impose milder functional restrictions on demands and preferences than Pendakur and Lewbel (2008) and Bargain and Donni (2010). In particular, we only place restrictions on the demand functions for one set of goods like clothing, instead of imposing shape restrictions on the demand functions for all goods.

Our identification uses private assignable goods. A good is defined to be *private* if it cannot be shared or consumed jointly by more than one person, and is defined to be *assignable* if it is consumed by one individual household member that is known to the researcher. Examples could include toys and diapers which are private goods assignable to children, or alcohol and tobacco which are private goods assignable to adults. In models where goods are purely private or purely public, what we call private, assignable goods are known as exclusive goods. See, e.g. Bourguignon, Browning, and Chiappori (2009). Chiappori and Ekeland (2008) and Cherchye, De Rock and Vermeulen (2008) among others show how assignable goods can aid in the identification of resource shares. Our strategy follows this line in assuming the presence, and observability, of a small number of private assignable goods, and uses these to identify childrens' resource shares. The end result is that we identify how total household expenditures on *all* goods are divided up among household members, just by observing how family member expenditures on one private good like clothing vary with total expenditures and family size.

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

We show that, exploiting semi-parametric restrictions on individual preferences similar to, but weaker than, Pendakur (1999) and Lewbel and Pendakur (2008), we can identify resource shares using Engel curves. This greatly facilitates empirical application of the model since we do not require price data and do not model price effects. Basing identification and estimation on Engel curves substantially reduces both model complexity and data requirements. Model and data requirements are also reduced by only needing

the demand functions for one type of private of good, like clothing.

We present empirical results for children's resource shares in Malawi using data from the Second Integrated Household Survey (IHS2), conducted by the National Statistics Office in conjunction with the International Food Policy Research Institute and the World Bank. We use the Malawi data for two reasons: Malawi is one of the poorest countries in the world, with per-capita (2005 PPP) GDP of US$773 in 2008, and; the IHS2 data are particularly rich in terms of household-level detail, which we exploit in our empirical work. Given the extreme poverty of most Malawian households, one may suspect that children are exceedingly vulnerable to intra-household inequality. We find that children command a reasonably large share of resources – roughly 20 percent for the first child – and that this share rises with the number of children – 5-10 percentage points per additional child. Moreover, fathers command a larger share of resources than mothers, and mothers seem to sacrifice more resources than fathers to their children. These patterns are evident even if household size is taken as endogenous and the model is estimated using instrumental variable techniques. Our findings are in the spirit of Duflo (2003), who finds evidence that male household heads tend not to allocate additional resources to children while female household heads do.

We find some evidence of gender discrimination against girls within the household, similar to Rose (1999).[2] We find that the mother's resource shares rises, and the children's resource share falls, as the proportion of children that are girls rises. Indeed, if all children are girls, then the mother's resource share rises, and the children's share falls, by roughly five percentage points. We also find that higher mother's education is associated with higher resource shares for women and children.

Finally, we use our estimates of resources shares to construct estimates of the poverty incidence of men, women and children in Malawi. Using the World Bank $2/day per-capita poverty measure, which assumes equal resource shares across people, yields a poverty rate of 91%. We find that allowing for unequal resource shares across people shows sharp differences in the incidence of poverty. In particular, we find that the incidence of poverty is roughly 60% for men and 85% for women whereas it is more than 95% for children.

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 total resource income constraint and a vector of Lindahl (1919) type shadow prices for goods. Each household member's resource share may differ from those of other members, but all members face the same shadow price vector. The resource share of a person and

[2]By discrimination here we only mean unequal treatment regarding allocation of resources. We do not claim that these allocations are necessarily unfair or imply inequality in welfare. For example, a large fraction of total expenditures in Malawi are devoted to food, so if women and girls are on average smaller and have lower caloric requirements, then they might be equally well off in a welfare sense to men and boys despite having smaller resource shares. We would like to thank Frederic Vermeulen for pointing this out.

shadow price vector of the household together define a shadow budget constraint faced by each individual within the household. Each household member then determines their own demand for each consumption good by maximizing their own utility function.

These shadow prices differ from market prices because of economies of scale to consumption. In particular, shadow prices will be lower than market prices for goods that are shared or consumed jointly by multiple household members. Goods that are not shared (i.e., private goods) will have shadow prices equal to market prices. Each member faces the same shadow prices because the degree to which a good can be shared is an attribute of the good, rather than an attribute of the consumer.

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

Resource shares for each individual may also be of interest even without knowledge of shadow prices. The resource share times the household expenditure level gives the extent of the individuals' budget constraint for consuming resources within the household, and is therefore an indicator of that individual's material well-being. For example, Lise and Seitz (2004) use estimated resource shares 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 "pareto weights" on individual utility, which are also used as measures of member bargaining power. Since we focus on the estimation of children's resource shares, we do not interpret our results in terms of bargaining power.

2.1 The Model

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

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

Let 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 a subutility function over goods this period, which may be just one component of member t's total utility. For children, $U_c\left(x_c\right)$ could either represent a child's actual utility function over the bundle of goods x_c that the child consumes, or 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 is assigned the same utility function $U_c\left(x_c\right)$. The model can be easily extended to include parameters that allow $U_c\left(x_c\right)$ to vary by, e.g., the age and sex of the child, but these like other observed household and individual characteristics are omitted for the time being. However, up to the inclusion of such observable characteristics, we assume that the individual household member utility functions $U_f\left(x_f\right)$, $U_c\left(x_c\right)$, and $U_m\left(x_m\right)$ 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 matrix A_s into a weakly larger (in magnitude of each element) bundle of 'private good equivalents' x, which is then divided among the household members, so $x = x_f + x_m + x_c$. Specifically, there is assumed to exist a K by K matrix A_s such that $x_f + x_m + x_c = x = A_s^{-1} z_s$. This "consumption technology" allows for much more general models of sharing and jointness of consumption than the usual collective model that categorizes goods only as purely private or purely public.

For example, suppose that a married couple without children ride together in a car (sharing the consumption of gasoline) half the time the car is in use. Then the total consumption of gasoline (as measured by summing the private equivalent consumption of each household member) is 3/2 times the purchased quantity of gasoline. Equivalently, if there had been no sharing of auto usage, so every member always drove alone, then the couple would have had to purchase 50% more gasoline to have each member travel the same distance as before. In this example, we would have $x^k = (3/2)\, z^k$ for k being gasoline, so the k'th row of A would consist of $2/3$ in the k'th column and zeros elsewhere. This $2/3$ can be interpreted as the degree of "publicness" of good k within the household. A purely private good k would have $x^k = 1$. Nonzero off diagonal elements of A_s may arise when the extent to which one good is shared depends upon other goods, e.g., if leisure time is a consumption good, then the degree to which auto use is shared may depend on the time involved, and vice versa.

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

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

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

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

In this paper, we lean heavily on existence of private assignable goods for identification of resource shares. A *private good* for our purposes is defined as one where its corresponding diagonal element of A is equal to 1 and all off-diagonal elements in that row or column are equal to 0. This means that private goods are goods that do not have any economies of scale in consumption. For example, food is private to the extent that any unit consumed by one person cannot also be eaten by another.[3] A private good is *assignable* if it is consumed exclusively by one known household member. So, e.g., a sandwich would be assignable if we could observe who ate it. In our application we observe separate expenditures on men's, women's, and children's clothing, which we take to be private and assignable.

Suppose there exists a private assignable good for a person of type t. This good is not jointly consumed, and so appears only in the utility function U_t, not in the utility functions of any other type of household member. Let $W_{ts}(y, p)$ be the share of total expenditures y that is spent by a household with s children

[3]This ignores possible economies of scale in food from reduced waste associated with preparation of larger quantities.

on the type t private good. For example $W_{cs}(y, p)$ could be the fraction of y that the household spends on toys or children's clothes. Also let $w_t(y, p)$ be the share of y that would be spent buying the type t private good by a (hypothetical) individual that maximized $U_t(x_t)$ subject to the budget constraint $p'x_t = y$. Unlike in BCL, these individual demand functions need not be observable.

While the demand functions for goods that are not private are more complicated (see the Appendix for derivations and details, especially equation (12)), the household demand functions for private assignable goods, derived from equation (1), have the simple forms

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

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

Two problems prevent us from using the BCL identification strategy in our setting with children. First, unlike adults, we cannot observe the demand functions for children living alone. BCL exploited data on adults living alone by assuming that single and married individuals have the same underlying utility functions. We replace this questionable assumption with the milder assumption that parents (and individual children) have utility functions over goods that do not depend on whether the number of children in the household is one, two, or three. (Our formal assumptions are even weaker, as described below, and in the Appendix.)

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

We get around these two problems in two steps. First, we restrict the resource share functions η_{fs} to be independent of household expenditures y, at least at low expenditure levels (though they may depend arbitrarily on p). This restriction has real bite, but one can at least write down sensible parametric household objective functions over reasonable parametric utility functions whose resulting resource shares satisfy this restriction (we present a class of such models in an online Appendix). Moreover, while resource shares cannot depend on total expenditures y, they can depend on closely related variables such as income,

wages, or wealth. Similar to Lewbel and Pendakur (2008) and Bargain and Donni (2009), this restriction allows us to recast the BCL model into an Engel-curve framework where price variation is not exploited for identification.

Second, we invoke some semiparametric restrictions on the shapes of individual Engel curves. These restrictions allow us to identify individual resource shares by comparing household demands for private assignables across people within households, or by comparing these demands across households for a given type of person. Unlike Bargain and Donni (2009), who also identify children's resource shares from Engel curves, we only place restrictions on the shapes of Engel curves for the assignable goods rather than on all goods, and we only need to assume similarity of preferences of individuals in households with varying numbers of children, rather than equality of preference of all adults regardless of whether they are single, couples without children, or couples with children.

3 Identification of Children's Resource Shares Using Engel Curves

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

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

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

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

The main difficulty for identification is that for every observable budget share function subscripted by ts on the left side of (3), there are two unobservable functions subscripted by ts on the right side. BCL achieve identification by assuming that w_{ts} on the right-hand side is observable via the behaviour of single people, leaving just one subscripted unobserved function to worry about: the resource shares η_{ts}. There are no single children, so we cannot use this method.

One extreme alternative would be to assume that people have identical preferences so that w_{ts} does not vary across t. In this case, for any household size s, we would use the 3 observable functions W_{ts} (for $t = m, f, c$) to identify 2 resource shares η_{ts} (the third may be computed because they add up to

1) and 1 budget share function w_s. A different extreme alternative would be to assume that people have preferences which do not vary across household type, so that w_{ts} does not vary across s. In this case, if we had enough household sizes, we would similarly have enough observable household budget share functions W_{ts} to identify the unobserved resource shares η_{ts} and unobserved individual budget share functions w_t. Unfortunately, both of these extreme restrictions are unreasonable. The first assumes that preferences are completely identical across people. The second is roughly equivalent to forcing w_t to be unresponsive to prices.

Our identification is based on the insight that one does not need the entire function w_{ts} to be independent of t or of s. It is enough for a separable part of w_{ts} to be independent of t or of s. Consider budget share functions w_t that are linear in functions of expenditure:

$$w_t(y, p) = h_{t0}(p) + h_{t1}(p)g_1(y) + h_{t2}(p)g_2(y)...h_{tL}(p)g_L(y),$$

where $h_{tl}(p)$ are price-varying functions which multiply the functions of expenditure $g_l(y)$. Then, observed private assignable budget share equations would be given by

$$W_{ts}(y) = \eta_{ts}h_{ts0} + \eta_{ts}h_{ts1}g_1(\eta_{ts}y) + \eta_{ts}h_{ts2}g_2(\eta_{ts}y)...\eta_{ts}h_{tsL}g_L(\eta_{ts}y),$$

where $h_{tsl} = h_{tl}(A_s p)$ for m, f and with $W_{cs}(y)$ defined analogously. We could achieve identification if any $h_{tl}(p)$ was independent of t so the coefficent h_{tsl} would drop its dependence on t. In this case, preferences would not be identical across people (indexed by t), but would be *similar* across people, due to the fact that one separable part of the budget share function is the same for all people. Identification would be analogous to the case where people had completely identical budget share functions.

Alternatively, we would achieve identification if any $h_{tl}(p)$ was independent of p so that the corresponding coefficient h_{tsl} would drop its dependence on s. In this case, preferences would not be identical across household types, but for any given person they would be *similar* across household types. Identification would be analogous to the case where preferences don't vary across household types.

Although the formulation above is useful for seeing how identification works, it is well-known that not all such formulations can be rationalised with a utility function (that is, not all are integrable). In the next sections, we describe restrictions which give individual budget share functions that can be rationalised with individual utility functions, and which permit identification of individual resource shares.

3.1 Identification if Preferences are Similar Across People

Here, we consider identification when people have similar preferences. We restrict how preferences for the private assignable goods vary across people, so we consider the same good for all people. For example, the private assignable good could be clothing, so that the demand function $w_t(y, p)$ gives person t's (unobserved) budget-share function for clothing when facing the constraint defined by y, p. In particular, we impose the restriction that Engel curves for the private assignable good have the same shape across

people, at least at low expenditure levels:

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

where $y^*(p)$ is a real expenditure threshold. The budget share functions for all people have the same shape, given by the function g, and differ only by the person-specific additive term $d_t(p)$ and the person-specific expenditure deflator $G_t(p)$. If $d_t(p)$ and the person-specific expenditure deflator $G_t(p)$ were the same for all people t, then preferences would be identical across people. These functions may differ across people, and the restriction binds only at expenditure levels below a threshold $y^*(p)$, so we say that preferences are *similar across people* (SAP) if equation (4) holds.

SAP is similar to the shape-invariance restriction of Pendakur (1999), except that we apply it only to the Engel curves for the private assignable goods and we apply it only at low expenditure levels. Pendakur (1999) shows that if people have costs that differ only by (price-dependent) multiplicative equivalence scales, then budget share functions must satisfy a condition like SAP for all goods and at all expenditure levels. When SAP is applied to all goods and at all expenditure levels, the result is a much stronger condition, known in the consumer demand literature as "shape-invariance". A vast amount of empirical consumer demand analysis imposes this shape-invariance restriction on budget share functions. See, e.g., Blundell, Duncan, and Pendakur (1998), Blundell, Chen, and Kristensen (2007), and Lewbel (2010). In contrast, we only assume SAP for a single good and only at real expenditure levels below a threshold $y^*(p)$.

Substituting the SAP restriction (4) into (3) we get, for $y \leq y^*$,

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

where $\delta_{ts} = d_t(A_s'p)$, $\gamma_s(y) = g(y, A_s'p)$ and $\Gamma_{ts} = G_t(A_s'p)$. The key here is that g does not vary across people. All these functions are evaluated at the same shadow price vector $A_s'p$, and as a result the function γ_s does not vary across people either (it does not have a t subscript). Theorem 1 in the Appendix shows the class of individual utility functions that satisfy SAP, and shows that if the function g has sufficient nonlinearity, then the resource shares η_{ts} are identified from the Engel curve functions $W_{ts}(y)$ for any household size s.

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

sufficient restriction for SAP is $b_t(p) = b(p)$.

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

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

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

Plugging these budget share functions into (3) yields

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

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

With more complex Engel curves for private assignable goods, identification is achieved by taking higher-order derivatives of the household Engel curves with respect to y or $\ln y$, but the spirit of the identification is the same. By assuming that individuals have budget share functions for their private goods that have the same shape across people for a given price vector, we are able to compare the shape of household Engel curves across people when they face the common within-household shadow price vector. Formal identification theorems are provided in the Appendix.[4]

3.2 Identification if Preferences are Similar Across Types

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

Let $p = \left[p_m, p_f, p_c, \overline{p}, \widetilde{p}\right]$ where $\overline{p}$ is the subvector of p corresponding to purely private goods other

[4]The not-for-publication Appendices, 7.1 and 7.2, also provide more details regarding the construction of PIGLOG preference models and household models that are consistent with all of our assumptions, including, e.g., that resource shares be independent of y.

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

Now we invoke the restriction that preferences are "similar across types" (SAT) as follows:

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

Again, $y^*(p)$ is a real expenditure threshold, so the restriction is applied only at low expenditure levels. Here, the scale-economies associated with non-private goods load onto the person-specific expenditure deflator $G_t(\widetilde{p})$. If $G_t(\widetilde{p}) = 1$, then preferences would be identical across household types. But, we allow preferences to vary through the expenditure deflator $G_t(\widetilde{p})$, so we say that preferences are only similar across types.

If SAT were applied to all price effects, rather than just the cross-price effects of nonprivate goods, so that $w_t(y, p) = g_t\left(\frac{y}{G_t(p)}\right)$, and if it were applied to all goods at all expenditure levels, then preferences would be homothetic, which is clearly undesirable. Here, we apply it only to the cross-price effects of non-private goods on the private assignable good, and we apply it only at low expenditure levels.

Lewbel and Pendakur (2008) apply a restriction like SAT to all price effects for all goods at all expenditure levels. They avoid the implication of homotheticity by requiring that the restriction hold for just one set of price changes rather than for all possible price vectors. They find that this restriction, which is much stronger than SAT, is consistent with the observed behaviour of single men, single women and childless married couples.

Substituting the SAT restriction (6) into (3), we get

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

where $\gamma_t(y) = g_t(y, p_t, \overline{p})$ and $\Gamma_{ts} = G_t\left(\widetilde{A}_s'\widetilde{p}\right)$. The key here is that the functions g_t, and therefore $\gamma_t(y)$, do not depend on household size s. We show in Theorem 2 in the Appendix that if private assignable good budget shares don't asymptote to zero when expenditures get too low (that is, if $\lim_{u\to 0}\gamma_t(u) \neq 0$) and there is sufficient variation in resource shares across individuals and household sizes, then the resource

shares η_{ts} are identified from the Engel curve functions $W_{fs}(y)$ for any three household sizes.

To illustrate, suppose again that each person has PIGLOG preferences over goods, so the indirect utility is given by $V_t(p, y) = b_t(p)\left[\ln y - \ln a_t(p)\right]$. This utility function satisfies SAT if $b_t(p) = \overline{b}_t(\overline{p}/p_t)$ and $a_t(p) = \overline{a}_t(\widetilde{p})$, so $\overline{b}_t$ is some function of private good prices and $\overline{a}_t$ is some function of the prices of other goods.[5] By Roy's identity, the corresponding budget share functions for each person's private assignable good are given by

$$w_t(y, p) = \beta_t(\overline{p}/p_t)\left[\ln y - \ln a_t(p)\right] + \alpha_t(p)$$

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

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

where $\delta_{ts} = -\beta_t(\overline{p}/p_t)\ln a_t(A_s' p) + \alpha_t(A_s' p)$ and $\beta_t = \beta_t(\overline{p}/p_t)$ These Engel curves are linear in $\ln y$, with slopes that vary across household size s for any person t. The coefficient of $\ln y$ for person t in a household with s children (which can be identified by linearly regressing W_{ts} on a constant and on $\ln y$) is $\eta_{ts}\beta_t$. The ratio of $\ln y$ coefficients for person t's assignable good in two different households equals the ratio of that person's resource shares in the two households. Given three household sizes we have a total of twelve equations (three Engel curves for each of three households, plus three sets of resource shares summing to one) in twelve unknowns (three sets of three resource shares, plus three β_t parameters), so the order condition for identification is satisfied. The corresponding rank condition for identification is provided in the Appendix. A nice feature of the SAT restriction is that with more than 3 household sizes, the model is overidentified. Thus, the information from additional household sizes can be used to test the model, or to improve the precision of the estimates.

One drawback of using the SAT restriction is that the identification hinges on the summation restriction on the resource shares, and hence may not be very strong in practice. To see this, observe that SAT with PIGLOG preferences identify resource shares by having derivatives of observable budget shares that satisfy

$$\begin{aligned} \partial W_{cs}(y)/\partial \ln y &= s\eta_{cs}\beta_c \\ \partial W_{ms}(y)/\partial \ln y &= \eta_{ms}\beta_m \\ \partial W_{fs}(y)/\partial \ln y &= \eta_{fs}\beta_f. \end{aligned}$$

[5] Assumption B3 of Theorem 2 in the Appendix provides a general class of utility functions that yield equation (6). For PIGLOG preferences, Assumption B3 holds if $b_t(p) = \overline{b}_t(\overline{p}/p_t)$ and $a_t(p) = \overline{a}_t(\widetilde{p})$. However, Assumption B3 is sufficient but not necessary for equation (6), and in the case of PIGLOG, this equation will hold under the weaker restriction that $b_t(p) = \overline{b}_t(\overline{p}/p_t)\widetilde{b}_t(\widetilde{p})$ and $a_t(p)$ is unrestricted, so the only required restriction for PIGLOG is that $b_t(p)$ be multiplicatively separable into a function of private goods $\overline{b}_t(\overline{p}/p_t)$ and a function of public goods $\widetilde{b}_t(\widetilde{p})$. Either way, the Engel curve system to be estimated takes the form (8).

for multiple values of s. Since the β_t coefficients are also unknown, the only thing that identifies the levels of η_{ts} from the observed budget share functions is the restriction that the resource shares η_{ts} sum to 1. If we instead had the restriction that the product of η_{ts} was 1, then identification would fail, because then we could for example replace each η_{ts} and β_t with $\widetilde{\eta}_{ts} = \eta_{ts}\lambda_t$ and $\widetilde{\beta}_t = \beta_t/\lambda_t$ for any positive constants λ_t such that $\lambda_m\lambda_f\lambda_c = 1$, without changing any of the observed budget share derivatives. This suggests that although identification is possible given the SAT restriction alone, it may take a lot of data to get precise estimates just from SAT.

3.3 Combining restrictions

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

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

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

for some functions $d_t(p)$, and household demands for the private assignables are then

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

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

It is important to stress that by invoking either or both of our identifying restrictions, we identify the levels of the resource shares themselves, not just how they vary with distribution factors, and we identify children's resource shares, not just those of adults. These features are not provided in the existing liter-

[6] Either restriction is compatible with large classes of indirect utility functions as described in the Appendix, though obviously the intersection of these restrictions is smaller. Using both restrictions together should provide more efficient estimates, assuming both restrictions are true.

ature on resource share/pareto-weight identification (as discussed in the introduction). Both are crucially important for our policy analysis, which is to measure the relative welfare of children in households of varying composition.

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

4 Engel Curve Estimation

In this section, we estimate Engel curve systems in an environment without price variation using the identification results provided in Theorems 1 and 2 and summarized in the previous section. The data come from the second Malawi Integrated Household Survey, conducted in 2004-2005. The Survey was 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. The survey includes roughly 11,000 households, drawn randomly from a stratified sample of roughly 500 strata.[7]

Malawi is a former British protectorate in southern Africa which achieved independence in 1964. The population of Malawi is roughly 16 million as of 2009 with a population density of approximately 120 persons per sq. km. It is one of the most densely populated countries in Africa. Half of Malawians live in the Southern region, 40% in the Central region and 10% in the Northern region, with more than 90% of the population living in rural areas. The economy of Malawi is largely based on agriculture and fishing with its chief exports being tobacco and sugarcane. It's recent political history has been remarkable for the absence of military coups and occasional multi-party elections, most recently in 2009. Despite its relative political stability, Malawi has numerous socio-economic tensions including extreme poverty (over 90% living under two US dollars per person per day), a high incidence of HIV/AIDs, high infant mortality and one of the lowest life expectancies in the world (51 years). In 2005, Malawi received almost $600 million

[7] For computational reasons, we do not use the complex sampling information associated with stratification in our estimation. This means that our estimates are unbiased and consistent, but not efficient. However, the robust nonlinear SUR and GMM estimated standard errors that we report remain consistent.

in foreign aid, equivalent to roughly 50 percent of government spending. Malawi is a good case study for our empirical exercise of measuring intra-household inequality, because with so much of the population having low household expenditure, inequality within households could in principle change our assessment of individual-level poverty.

4.1 Malawian Expenditure Data

Our sample consists of 2794 households comprised of married couples with one to four children all under 15 years of age. These households (drawn from the database of approximately 11,000 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; (4) households with any member over 65 are excluded; and (5) urban households are excluded. Our private assignable good is the sum of clothing and footwear expenditures. Table 1 gives summary statistics for our sample of nonurban families with 2 parents and 1-4 children.

Table 1: Data Means, Malawian micro-data

	couples with				all
	1 child	2 children	3 children	4 children	
Number of Observations	845	825	667	457	2794
clothing plus men	1.46	1.34	1.21	1.00	1.29
footwear women	2.10	1.92	1.61	1.52	1.84
(in per cent) children	1.06	1.50	1.69	1.89	1.48
log-total-expenditure	-0.13	-0.06	0.01	0.11	-0.04

Because the Malawian data are very rich, we also include some demographic variables, which may affect preferences and/or resource shares. If they were to affect resource shares and not preferences, they would be called "distribution factors" in much of the collective household model literature. Our theorems show identification for models without these variables, so one can apply the theorems conditioning on each value these additional variables can take on, and thereby prove identification when these variables are included. As in Browning and Chiappori (1998) the presence of distribution factors may help identification of resource shares, but, unlike Browning and Chiappori (1998) (and most other empirical collective household models), we do not require distribution factors for identification. This also means that we do not have to take a stand on whether any particular demographic variable affects only resource shares and hence is a distribution factor, versus affecting either resource shares, preferences or both.

We include 14 demographic variables in our models: region of residence (non-urban North and non-urban Central with non-urban South as the left-out category); the average age of children less 5; the minimum age of children less 5; the proportion of children who are girls; the age of the man less 28 and the age of the woman less 22 (the average ages of men and women in the sample); the education levels

of the household head and spouse (ranging from -2 to 4, where 0 is the modal education level); the log of the distance of the village to a road and to a daily market; a dummy indicating that the 3 month recall period for consumption occurred over the dry season; and dummy variables indicating that the household is christian or muslim (with animist/other as the left-out category). We allow all demographic factors to affect both the preferences and the resource shares of every household member.

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

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

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

Let a be a vector of 4 dummy variables for the 4 household types (indexed by s), and let z indicate the 14 demographic variables. For each person t, the resource shares η_{ts} and the intercept preference parameters δ_{ts} are specified as linear in a and z, so they have 18 coefficients each. The slope preference parameters β_{ts} are specified according to the identifying restriction: given SAP, β_{ts} is linear in a and z for a total of 18 coefficients; and given SAT, β_{ts} is linear in only a constant and z (15 coefficients) for each person t (3 people) for a total of 45 coefficients. Given both SAP and SAT, β_{ts} is linear in a constant and z for a total of 15 coefficients.

We implement the model by adding an error term to each equation of (10). These errors may covary across equations, so in the case with exogenous regressors, we estimate the model via nonlinear Seemingly Unrelated Regression (SUR), and with endogenous regressors, we use Hansen's (1982) Generalised Method of Moments (GMM). Both SUR and GMM estimators are iterated until the estimated parameters and error/orthogonality condition covariance matrices "settle". Iterated SUR is equivalent to maximum likelihood with multivariate normal errors. All models use the sum of clothing and footwear expenditures for each person as the private assignable good.

4.2 Results

We present estimates for η_{ts} in Table 2. Estimates of other parameters are not presented to save space, but they are available on request. Asymptotic standard errors are robust to heteroskedasticity of unknown form, and are given in *italics*. We note that all estimated values of the constant terms in β_{ts} are statistically significantly different from zero (nonzero latent slopes are required for identification for all models). The leftmost block of Table 2 gives estimates using the SAP restriction, the middle block gives estimates using the SAT restriction, and the rightmost block imposes both SAT and SAP restrictions. We report only coefficients relating to the levels of resource shares in different household sizes, and coefficients relating to a few key demographic factors which potentially relate to policy levers: age and gender composition of children, and the education level of the parents (full estimation results are available on request from the authors). Parameters related to children's resource shares are computed off of the estimated values for adult resource share parameters, based on the restriction that resource shares sum to one.

Define a reference household as one in which $z = 0$, which is the case for animist/other households living in a village with both a daily market or a road, whose consumption recall was during the wet season, in which the man is aged 28 and woman is aged 22, and both have the modal level of education, and the children are all boys aged 5 (so that the average and minimum are both 5). For a reference household, the resource share is given by the number-of-children term in η_{ts}. In the Table, we report the level of the resource share in households of various sizes for the man η_{ms}, woman η_{fs}, all children $s\eta_{cs}$, and each child η_{cs}. For the demographic factors, we report only the effect on the resource shares of all children.

Consider first the rightmost block which presents the estimates given both the SAP and SAT (SAP&SAT) restrictions. Looking at the coefficients giving the level of resource shares in reference households of different sizes, we see that, roughly speaking, as the number of children increases, the total share of household resources devoted to children goes up, but the average share devoted to each child declines. A reference household with one child directs 22.7 per cent of its expenditures to children's consumption. With two children, this share rises to 31.7 per cent, and four children, to 43.4 per cent. Even with three or four children, the resource share for each child is still about 11 per cent.

Table 2: Estimates from Malawian Clothing (inc Footwear) Budget Shares

		SAP		SAT		SAP&SAT	
		Estimate	*Std Err*	Estimate	*Std Err*	Estimate	*Std Err*
one child	man	0.443	*0.048*	0.378	*0.076*	0.400	*0.045*
	woman	0.308	*0.041*	0.368	*0.062*	0.373	*0.042*
	children	0.249	*0.037*	0.254	*0.072*	0.227	*0.036*
	each child	0.249	*0.037*	0.254	*0.072*	0.227	*0.036*
two children	man	0.423	*0.051*	0.436	*0.090*	0.462	*0.051*
	woman	0.222	*0.042*	0.212	*0.056*	0.221	*0.043*
	children	0.355	*0.045*	0.352	*0.100*	0.317	*0.045*
	each child	0.177	*0.022*	0.176	*0.050*	0.158	*0.023*
three children	man	0.427	*0.057*	0.437	*0.099*	0.466	*0.053*
	woman	0.185	*0.046*	0.166	*0.054*	0.176	*0.044*
	children	0.388	*0.050*	0.397	*0.114*	0.358	*0.050*
	each child	0.129	*0.017*	0.132	*0.038*	0.119	*0.017*
four children	man	0.318	*0.070*	0.352	*0.112*	0.384	*0.063*
	woman	0.214	*0.054*	0.168	*0.062*	0.182	*0.052*
	children	0.468	*0.061*	0.479	*0.133*	0.434	*0.059*
	each child	0.117	*0.015*	0.120	*0.033*	0.109	*0.015*
min. age	man	-0.005	*0.010*	0.007	*0.010*	0.008	*0.009*
of children	woman	-0.005	*0.008*	-0.014	*0.008*	-0.014	*0.008*
	children	0.010	*0.006*	0.007	*0.007*	0.007	*0.006*
avg. age	man	0.006	*0.010*	-0.007	*0.010*	-0.008	*0.009*
of children	woman	0.006	*0.008*	0.017	*0.008*	0.017	*0.008*
	children	-0.012	*0.006*	-0.010	*0.008*	-0.009	*0.006*
proportion	man	0.006	*0.029*	0.001	*0.031*	-0.003	*0.028*
girl children	woman	0.053	*0.024*	0.058	*0.027*	0.056	*0.026*
	children	-0.059	*0.020*	-0.059	*0.025*	-0.053	*0.019*
man	man	0.021	*0.009*	0.008	*0.010*	0.008	*0.010*
education	woman	-0.009	*0.008*	0.003	*0.009*	0.002	*0.009*
	children	-0.012	*0.006*	-0.011	*0.007*	-0.010	*0.006*
woman	man	-0.022	*0.012*	-0.050	*0.012*	-0.049	*0.011*
education	woman	0.007	*0.010*	0.030	*0.012*	0.032	*0.011*
	children	0.015	*0.008*	0.020	*0.010*	0.017	*0.008*

Although the total resources of parents roughly decline with the number of children, this is not spread evenly across men and women. Men absorb between 40 per cent and 47 per cent of household resources if there are 3 or less children. Given the standard errors, this is a relatively small amount of variation. In contrast, women see their resource shares drop by about 20 percentage points as the number of children goes from 1 to 3. These patterns are reasonably consistent under the SAP and SAT assumptions individually. One difference is that there is a large (but statistically insignificant) drop in men's resource share in households with 4 children under just the SAP assumption. A second difference is that the estimated levels of resource shares are much less precisely estimated under SAT than under the SAP or SAT&SAP

cases, with standard errors that are almost twice as large. This is consistent with our earlier discussion regarding the comparative weakness of SAT identification.

Turning to the covariates, three observations stand out. First, the coefficients relating to the proportion of children who are girls are important. In particular, if all children in the household are girls, then their combined resource share is about 6 percentage points lower than if the children are all boys. These resources are almost fully diverted to the woman (the man's resource share is almost unaffected).[8] Thus, unlike Deaton (1989, 1997) but similar to Rose (1999), we find statistically significant evidence of gender discrimination in consumption within the household. One difference between our finding and that of Rose (1999), is that we find that gender discrimination is the *status quo* and does not arise only in response to household income shocks.

Second, the higher the mother's level of education the more resources are diverted from fathers – with these extra resources being allocated $2/3$ to mothers and $1/3$ to children. These effects have reasonably large magnitudes. If a woman moves from the median to the top decile of education (from 0 to 2), the man's resource share declines by 10 percentage points. In contrast, we see little difference in resource shares from differences in men's education. The magnitude of education effects depends on which identifying assumption is used: the effects of women's education are much smaller (though still statistically significant) given the SAT assumption alone.

Third, a higher variance in the age distribution of children tends to increase the mother's share of resources. If the minimum age of children in the household rises by one year, women lose a 1.4 percentage point share of resources. Conversely, when the average age of children in the household rises by one year, women gain 1.7 percentage points. The estimates also suggest that these resources are diverted to men and children in roughly equal measure although this division is not statistically significant. These estimates imply that women tend to receive higher shares of household resources when both young and old children are present. As with the education effects, the estimated magnitudes are smaller given the SAT assumption alone.

4.3 SAP, SAT or Both?

A natural question is whether the SAT, SAP, or both assumptions are valid. Given the sensitivity of some of our findings, the answer to this question can have policy relevance. We can easily test downwards,

[8] We note that part of this finding is somewhat specification dependent. For various specifications of the list of included demographic variables, we find that children's resources always respond negatively to the proportion of girls. However, whether these resources are diverted to the man or the woman is specification-dependent.

and assess the joint restriction that *both* SAP and SAT hold. For example, we can estimate the model (for a single private assignable good) under SAP and conduct a Wald test on the hypothesis that the the coefficients on the household size dummies are identical for the 4 household types. The sample value of this test statistic is 1.1, and it is distributed as a χ^2_3 which has a 5 per cent critical value of 6 under the null hypothesis that SAP and SAT both hold. Alternatively, we can estimate under SAT and conduct a Wald test of the hypothesis that the β_t are the same for all persons t. Since each of the 3 person-specific β_t functions has 15 parameters, this amounts to testing 30 restrictions. The sample value of this test statistic is 7.4, and it is distributed as a χ^2_{30} with a 5 per cent critical value of 43.8. Thus, the combination of SAP and SAT is not much worse than either SAP or SAT separately, so we recommend combining SAP and SAT.

Some aspects of the SAP and SAT identifying restrictions themselves are testable. Since resource shares are exactly identified given SAP, we cannot test upwards with only 1 assignable good (though, we can with 2 assignable goods, as we show below). However, since resource shares are overidentified given SAT if there are more than 3 household sizes, our setting with 4 household sizes allows us to test the overidentifying restriction. In particular, given SAT and four household sizes, there are 12 identifiable slopes of W_{ts} with respect to $\ln y$ (4 household sizes times 3 goods), and they depend on 8 resource share functions (4 household sizes times 2 resource share functions, where the third is given by the summation restriction) and 3 latent slopes β_t. Thus, we can add an additional slope parameter to the model, which is "on" for one household size for one person's assignable good, and test the exclusion restriction on this additional parameter. Of course, this additional parameter must be interacted with the 14 demographic parameters as well, yielding a total of 15 parametric restrictions. The sample value of the Wald test statistic for this restriction is 0.4, and it is distributed as a χ^2_{15} with a 5 per cent critical value of 25. Thus, we find no evidence that SAT does not hold against a more general (linear) alternative.

Next we briefly consider testing our model and identifying restrictions upwards against a more general alternative. When there is more than one private assignable good k for each person, we can use either private good to estimate the resource shares η_{ts} for each person, and then test that the estimate resource shares do not vary by k. Alternatively, we can impose the restrictions that η_{ts} not vary with k, and these overidentifying restrictions should improve the precision of our estimates.

We implement a two-private-good model by separating clothing and footwear expenditures into two private assignable goods, each of which has a budget-share equation for each person in the household. We invoke the combined restriction SAT and SAP. A formal test that the estimated resource shares recovered from the clothing shares are the same as those estimated from the footwear shares has 36 restrictions –

the 2 resource share functions each have 18 parameters (4 household sizes and 14 demographic variables). The sample value of the likelihood ratio test statistic for this restriction is 15.4, and is distributed as a χ^2_{36} with 5 per cent critical value of 51 under the null hypothesis that our resource shares are unique. In contrast, the sample value of the Wald test statistic for this restriction is 72. Taken together, these results say that the test of the restriction may or may not be statistically significant.

The estimates given both SAP and SAT are the most precise of the estimates presented because more identifying restrictions are imposed than with either SAP or SAT alone. These estimates allow for the sharpest testing of hypotheses about the behaviour of resource shares across household size. One interesting hypothesis is whether the resource share functions depend linearly on the number of children. In this case, the resource share functions are linear in a constant, s, and the 14 demographic variables, which involves 2 exclusion restrictions in each of 2 resource share functions compared to the results in the rightmost column of Table 2. The sample value of the likelihood ratio test statistic for this hypothesis is equal to 0.6 and it is distributed as a χ^2_4 with a 5 per cent critical value of 9.5. In contrast, the sample value of the Wald test statistic for this restriction is 13.2 (and also a χ^2_4). Thus, it may or may not be reasonable to model resource shares as linear in the number of children. We conclude that imposing the restriction that resource shares are linear in the number of children does not do undue violence to the data, and provides a useful summary measure of the marginal effect of each additional child on resource shares.[9]

4.4 Dealing with Endogeneity

Our models can be readily extended to deal with endogeneity via instrumental variables. Two types of endogeneity are of particular interest in our setting. First, empirical work on consumer demand often has to contend with infrequency-driven endogeneity in the total expenditure of the household. Endogeneity of total expenditures is particularly important to deal with in our setting because, as noted above, we achieve our identification of resource shares off of estimated derivatives of budget shares with respect to the log of total expenditures.

A second type of endogeneity that we wish to address is the possible correlation of the number of children with the residuals in budget share equations. In particular, if unobserved preference heterogeneity is connected to both fertility decisions and expenditure allocation decisions, then the number of children in the household will be endogenous.

[9]This linearity will be a useful restriction when we consider dealing with possible endogeneity in the number of children. It is difficult to find sufficient instruments to control for each of three different dummy variables corresponding household size, but it may be feasible to find instruments for just a single scalar-valued number of children.

To formally account for endogeneity (and possible heteroskedasticity of unknown form), we use an instrumental variables estimator. Since we show above that the joint SAP/SAT identification works reasonably well, we implement the model, Equation (10), with $\beta_{ts} = \beta$, so that the latent slope parameters do not depend on either the number of children in the household or the person. In this case, the subscript s is absorbed into regressors and coefficients (specifically, s appears in the household size dummy variables a which are inside the resource shares η_{ts}, and in the latent intercepts δ_{ts}) so that the error terms from estimation for each person's assignable good, e_t, do not have an s subscript.

Let $q_t = q_{t1}, ..., q_{tJ_t}$ be an J_t-vector of instruments uncorrelated with the error terms, e_t. These instruments can be any functions of any variables that are conditionally exogenous with respect to e_t. Then, $E\left(e_t q_{tj}\right) = 0$ for all t, j implies for our model:

$$E\left[\left(W_{cs} - s\eta_{cs}\left(\delta_{cs} + \ln \eta_{cs}\right) - s\eta_{cs}\beta \ln y\right) q_{cm}\right] = 0,$$

for $j = 1, ..., J_c$, and

$$E\left[\left(W_{ts}(y) - \eta_{ts}\left(\delta_{ts} + \ln \eta_{ts}\right) - \eta_{ts}\beta \ln y\right) q_{tm}\right] = 0,$$

for $t = m$, f and $j = 1, ..., J_m$ and $j = 1, ..., J_f$. Given these moment conditions, the parameters may be estimated by Hansen's (1982) generalised method of moments (GMM).

We use instruments which are correlated with the optimal instruments for our model, where the optimal instruments are the derivatives of the error terms e_t with respect to the model parameters η_{ts}, δ_{ts} and β. In particular, we evaluate these derivatives at SUR pre-estimates, and plug in "hat" versions of endogenous variables rather than their true values, where "hat" versions are (first-stage) OLS predictions of the endogenous variables on the basis of all observed exogenous variables. This is essentially equivalent to the first stage of two stage least squares, when the first stage equations are nonlinear. Note that our models are overidentified, since β is found in all 3 equations, and η_{ms} and η_{fs} are each found in 2 equations (due to the summation restriction on η_{ts}).

Our exogenous variables include: the log of expenditure (except in models where we treat it as endogenous), all 14 demographic variables, the log of the value of livestock holdings, the log of the value of durable goods holdings, the log of the sum of livestock and durable holdings, the presence in the village of a HIV-prevention oriented NGO office, the distance to a doctor's office and a dummy variable indicating that the woman has a chronic illness. Our endogenous variables are either: the number of children in the

household; or both the number of children in the household and the log of total expenditure. These instruments are not very strong in predicting the number of children in the household in that, conditional on the demographic variables and the log of expenditure, the F statistic on the excluded instruments in the first stage is only 2.5. However, these instruments are very strong in predicting the log of expenditure: the F statistic on the excluded instruments is 67.

Table 3: GMM Estimates

		SUR		GMM endog: extra child		GMM endog: extra child, lny	
		Estimate	*Std Err*	Estimate	*Std Err*	Estimate	*Std Err*
one child	man	0.456	*0.045*	0.407	*0.056*	0.341	*0.074*
	woman	0.358	*0.044*	0.427	*0.054*	0.408	*0.071*
	child	0.186	*0.030*	0.166	*0.044*	0.251	*0.073*
extra child	man	-0.012	*0.018*	0.083	*0.085*	-0.008	*0.095*
	woman	-0.055	*0.015*	-0.148	*0.073*	-0.075	*0.098*
	children	0.068	*0.014*	0.065	*0.040*	0.083	*0.042*
min. age of children	man	0.003	*0.009*	0.056	*0.040*	0.004	*0.043*
	woman	-0.007	*0.008*	-0.056	*0.034*	0.000	*0.044*
	children	0.004	*0.006*	0.000	*0.019*	-0.004	*0.019*
avg. age of children	man	-0.004	*0.009*	-0.058	*0.040*	-0.010	*0.043*
	woman	0.009	*0.008*	0.058	*0.035*	0.007	*0.044*
	children	-0.005	*0.006*	-0.001	*0.019*	0.003	*0.019*
proportion girl children	man	-0.015	*0.030*	0.030	*0.033*	-0.026	*0.038*
	woman	0.063	*0.029*	0.026	*0.027*	0.090	*0.040*
	children	-0.048	*0.016*	-0.056	*0.024*	-0.065	*0.033*
man education	man	0.008	*0.010*	0.020	*0.010*	0.013	*0.012*
	woman	-0.001	*0.010*	-0.016	*0.010*	-0.006	*0.012*
	children	-0.008	*0.005*	-0.004	*0.005*	-0.007	*0.008*
woman education	man	-0.047	*0.011*	-0.044	*0.012*	-0.058	*0.014*
	woman	0.033	*0.011*	0.028	*0.012*	0.042	*0.015*
	children	0.014	*0.006*	0.016	*0.007*	0.016	*0.009*

Table 3 gives estimates of resource share parameters in models where children enter the resource shares and latent intercepts (η_{ts} and δ_{ts}) linearly, so that these functions have 16 parameters each (a constant, the number of children s, and the 14 demographic shifters). The leftmost column presents SUR estimates analogous to those presented in Table 2, and the middle and rightmost columns give GMM estimates corresponding to instrumenting either the number of children, or both the number of children and the log of expenditure. Hansen J-tests of overidentifying restrictions do not suggest that the instruments are endogenous and hence invalid (with p-values of 51% and 60% for the middle and rightmost columns, respectively).

In the leftmost column, the SUR estimates show that most of the results in Table 2 are evident when we replace the household size dummies with the scalar-valued number of children variable. In particular,

we see that men's resource shares do not respond to the number of children, but women's decline substantially and statistically significantly with the number of children, so that each additional child increases the children's resource share by almost 7 percentage points.

The middle and rightmost columns show GMM estimates which account for the possible endogeneity of household size and both household size and the log of expenditure, respectively. In general, the patterns relating to household size are still visible, but are estimated very much less precisely (particularly so when both size and expenditure are treated as endogenous). Again, children's resources are marginally statistically significantly increasing in the number of children, with about the same magnitude as in the SUR regressions. Men's shares are not statistically significantly related to the number of children and women's shares are statistically significantly declining if only household size is considered as endogenous.

The fact that the GMM regression estimates are very similar to the OLS estimates suggests that there may not be significant endogeneity in the number of children in the household. Hausman tests support this claim: sample value of the Hausman test statistic for the hypothesis that the parameter estimates are the same in the SUR and the GMM regressions is equal to 80 in the middle regression and equal to 116 in the rightmost regressions. In either case, the test statistic is distributed as a χ^2_{95} with a 5 per cent critical value of 119 under the null hypothesis. The lack of significant endogeneity may be due in part to timing issues - resource allocations may be decided long after fertility decisions.

Our key take-away from this exercise is that the importance of the education and child gender covariates remains unchanged. As one might expect, the reduced precision of GMM relative to SUR does not take as large a toll on the estimates of parameters associated with the exogenous regressors. The gender bias in children's resources is evident across all specifications: if the children are all girls, they absorb about $5-7$ percentage points less of household resources. Whether it is men or women who gain at the expense of girls is less clear, although in all specifications the change in the men's share is not statistically significant. We also see a substantial effect of women's education, diverting resources towards women and children. As women's education increases, men's share of household resources falls by roughly 5 percentage points. These resources are shared roughly $2/3$ and $1/3$ by women and children, respectively.

4.5 Resource Shares, Poverty Rates and Child Poverty

The empirical results described so far relate to the levels of resource shares for persons in reference households, and to the marginal effects of various demographic factors. However, this does not tell us how resource shares would change in aggregate across household sizes because the demographic factors them-

selves covary with household size. To evaluate, for example, whether men or women make the larger sacrifice of consumption for their children, it is illustrative to consider the average resource shares in households of different sizes, averaging over all the values of demographic factors observed in the population.

Table 4: Estimated Resource Shares and Poverty Rates

		Mean	Std Dev	Min	Max	Pov Rate *Unequal*	Pov Rate *Equal*
one child	man	0.463	0.087	0.245	0.762	0.686	0.850
	woman	0.402	0.071	0.168	0.587	0.766	
	children	0.135	0.047	0.008	0.260	0.954	
	each child	0.135	0.047	0.008	0.260		
two children	man	0.516	0.078	0.282	0.786	0.547	0.916
	woman	0.273	0.063	0.075	0.475	0.885	
	children	0.211	0.044	0.059	0.326	0.970	
	each child	0.105	0.022	0.029	0.163		
three children	man	0.521	0.081	0.219	0.795	0.522	0.948
	woman	0.244	0.065	0.002	0.512	0.889	
	children	0.236	0.042	0.112	0.374	0.996	
	each child	0.079	0.014	0.037	0.125		
four children	man	0.441	0.080	0.170	0.701	0.538	0.972
	woman	0.267	0.066	0.043	0.532	0.838	
	children	0.293	0.037	0.178	0.402	0.989	
	each child	0.073	0.009	0.044	0.101		
All Households	man	0.489	0.088	0.170	0.795	0.582	0.913
	woman	0.304	0.093	0.002	0.587	0.842	
	children	0.207	0.070	0.008	0.402	0.974	
	each child	0.103	0.038	0.008	0.260		
All Persons	all	0.235	0.177	0.008	0.795	0.855	0.924

The leftmost columns of Table 4 presents summary statistics on the estimated values of resource shares for people in households of different sizes. It is comforting to see that the minima and maxima of estimated resource shares do not fall outside the zero to one range for any person in any household in the sample. The standard deviations are quite small in most cases. Interestingly, the standard deviations of resource shares are larger for men than for women in all household sizes. Thus, the covariates are not very important in terms of their effects on resource shares, though they do induce more variation for men than for women. Much more important than these factors are the household sizes themselves. This suggests that our ability to identify the *level* of resource shares, rather than just their response with respect to demographic or distribution factors, is particularly important.

The rightmost columns of Table 4 show the estimated poverty rates (at the household level) for households of different sizes. The rightmost column (Equal) uses the World Bank's US$2/person/day poverty

threshold applied to the entire household. The next column to the left (Unequal) uses our resource shares to construct person-level expenditures (equal to household expenditure times the resource share) and compares this to the US$2/day threshold. We use the OECD estimate of the relative needs of children (60 % that of adults), and so for children, compare their expenditure to US$1.20/day. The bottom block and bottom row give the estimated poverty rate for all households together and for all persons. Here, we see an equal poverty rate for households of 91.3% for our sample. For comparison, the World Bank reported poverty rate for all households in Malawi in 2004 was 90.5%. We now consider how accounting for intrahousehold inequality changes these per capita estimates.

There are at least three features to note in our poverty estimates. First, Table 4 shows that there are a lot more households with poor women than with poor men. For example, looking at the rows for All Households, we see that 58.2 per cent of households have a poor man, but 84.2 per cent of households have a poor woman. Second, the poverty rates of men seem to drop with household size, but the poverty rate for women and children is roughly rising with household size. Third, more households have poor children than have poor adults. In households with 3 or 4 children, nearly all children are poor. Indeed our estimates are that incidence of child poverty in Malawi is over fifty percent higher than the rate of poverty for adult men, and roughly $1/6$ higher than that of adult women.

We considered a similar exercise for the three regions of Malawi and find somewhat contrasting pictures of poverty. In the North, we observe lower incidences of poverty for men (38%), but higher incidences for women and children (93% and 99%, respectively). In contrast, in the South, we observe higher incidences of poverty for men (76%) and relatively similar incidences for women and children in comparison to the North (90% and 97%). In the Central region women tend to fare best in a relative sense (a poverty incidence of 75%) while men are slightly worse off than in the North (44%) and children are essentially identical to the South (97%).

The major conclusion here is that intra-household allocation may be very important to measuring child poverty. The fact that we can now measure children's resource shares within household may thus be very useful in measuring poverty, and in measuring the full effect of policy interventions aimed at poverty alleviation.

5 Conclusions

Child poverty is at the root of much inequality. 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 children's shares of household resources. Most collective household models either ignore children, or treat them as public or private goods for adults.

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

Using household consumption data for Malawi, we find that children command a reasonably large share of household resources (though not enough to avoid having higher rates of poverty than their parents) and that the share of resources devoted to children rises with the number of children, though the average share per child tends to decrease, at least when going from one child to two or three. Mothers appear to contribute more resources than fathers to children, and we find some evidence of gender-bias in children's resource shares. We also find that there is substantial intra-household inequality, one consequence of which is that standard per-capita poverty indices, which by construction ignore intra-household inequality, present a misleading picture of poverty, particularly for children..

6 Appendices

6.1 Appendix 1: Theorems

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

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

Rewriting equation (1) we have

$$\max_{x_f, x_m, x_c, z_s} \widetilde{U}_s \left[U_f\left(x_f\right), U_m\left(x_m\right), U_c\left(x_c\right), p/y \right] \quad \text{such that} \quad z_s = A_s \left[x_f + x_m + s x_c \right] \quad \text{and} \quad y = z_s' p \tag{11}$$

The demand functions for the household s arising from the household's maximization problem, equation (11), can be written as follows. Let A_s^k denote the row vector given by the k'th row of the matrix A_s.

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

household s demand functions are given by

$$z_s^k = H_s^k(p, y) = A_s^k\left[h_f\left(A_s' p, \eta_{fs} y\right) + h_m\left(A_s' p, \eta_{ms} y\right) + s h_c\left(A_s' p, \eta_{cs} y\right)\right] \tag{12}$$

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

Note in equation (12) that each child gets a share $\eta_{cs}(p)$, so the total share devoted to children is $s\eta_{cs}(p)$. By definition, resource shares must sum to one, so for any s

$$\eta_{fs}(p) + \eta_{ms}(p) + s\eta_{cs}(p) = 1 \tag{13}$$

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

ASSUMPTION A1: Equations (11), (12), and (13) hold, with resource shares $\eta_{ts}(p)$ that do not depend upon y.

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

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

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

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

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

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

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

It follows immediately from Assumptions A1 and A2 that, for the private, assignable goods $k = f, m, c$, equation (12) simplifies to

$$z_s^k = H_s^k(p, y) = h_k\left(A_s' p, \eta_{ks}(p)\, y\right) \quad \text{for } k \in \{m, f\} \tag{14}$$

$$\text{and } z_s^c = H_s^c(p, y) = s h_c\left(A_s' p, \eta_{cs}(p)\, y\right) \tag{15}$$

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

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

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

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

$$V_t(p, y) = I\left(y \leq y^*(p)\right) \psi_t\left[\upsilon\left(\frac{y}{G_t(p)}\right) + F_t(p), \widetilde{p}\right] + I\left(y > y^*(p)\right) \Psi_t(y, p) \tag{16}$$

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

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

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

imal level $y^*(p)$, and therefore the demand functions for all goods can have any smooth parametric or nonparametric functional form at total expenditure levels $y > y^*(p)$.

The key restriction in Assumption A3 is that the functions υ and φ do not vary across people. The function $\upsilon(y/g_t(p)) + F_t(p)$ with $\partial F_t(p)/\partial p_t = \varphi(p)$, if it were the entire indirect utility function, would, induce shape invariance on the Engel curves of the private, assignable goods. See Pendakur (1999), Blundell, Duncan, and Pendakur (1998), Blundell, Chen, and Kristensen (2007), and Lewbel (2010). However, the demand functions that arise from equation (16) are only constrained to satisfy same invariance shape at low expenditure levels, because this restriction is only imposed for $y \leq y^*(p)$. The result of this restriction will be that the Engel curves for assignable goods can have any shape, but they will all need to have the same shape at low total expenditure levels.

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

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

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

To show this formally, apply Roy's identity to equation (16). The result is that, for person t and any good k, when $y > y^*(p)$, the demand function will be given by applying Roy's identity to $\Psi_t(y, p)$ giving $h_t(y, p) = -\left[\partial\Psi_t(y, p)/\partial p_k\right] / \left[\partial\Psi_t(y, p)/\partial y\right]$. However, when $y \leq y^*(p)$,applying Roy's identity to equation (16) gives

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

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

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

which makes the above demand function simplify to

$$h_k(y, p) = \frac{y}{G_k(p)} \frac{\partial G_k(p)}{\partial p_k} - \frac{\varphi(p) \frac{G_k(p)}{y}}{\upsilon' \left(\frac{y}{G_k(p)} \right)} y \text{ for } y \leq y^*(p) \tag{17}$$

which we can write more simply as

$$h_k(y, p) = \delta_k(p) y + g\left(\frac{y}{G_k(p)}, p \right) y \text{ for } y \leq y^*(p) \tag{18}$$

for functions δ_k and g. Substituting this into equation (14) gives household demand functions for the assignable goods

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

and, for children

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

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

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

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

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

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

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

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

Notes:

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

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

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

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

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

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

Now we consider alternative identifying assumptions, based on comparing demand functions across households of different sizes, instead of across individuals within a household. We maintain Assumptions A1 and A2, but in place of Assumption A3 now assume the following:

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

$$V_t(p, y) = I\left(y \leq y^*(p)\right) \psi_t\left[u_t\left(\frac{y}{G_t(\widetilde{p})}, \frac{\overline{p}}{p_t}\right), \widetilde{p}\right] + I\left(y > y^*(p)\right) \Psi_t(y, p) \tag{19}$$

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

The goods in the price vector $\overline{p}$ are assumed to be private, and so have no economies of scale in household consumption, but they need not be assignable, so for example $\overline{p}$ might include food products

that are consumed by all household members. Being private means that the elements of $A_s' p$ corresponding to $\overline{p}$ will just equal $\overline{p}$, so the term $\overline{p}/p_t$ will not change when p is replaced by $A_s' p$.

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

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

To obtain demand functions corresponding to the indirect utility function in Assumption B3, apply Roy's identity to equation (19). As before, for person t and any good k, when $y > y^*\left(p\right)$, the demand function will be given by applying Roy's identity to $\Psi_t\left(y, p\right)$ giving
$h_t(y, p) = -\left[\partial\Psi_t\left(y, p\right)/\partial p_k\right]/\left[\partial\Psi_t\left(y, p\right)/\partial y\right]$. However, when $y \leq y^*\left(p\right)$,applying Roy's identity to equation (19) gives

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

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

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

$$h_k(y, p) = \frac{\partial u_k\left(\frac{y}{G_k(\widetilde{p})}, \frac{\overline{p}}{p_k}\right)'}{\partial\left(\overline{p}/p_k\right)}\frac{\overline{p}}{p_k^2}\frac{G_k\left(\widetilde{p}\right)}{u_k'\left(\frac{y}{G_k(\widetilde{p})}, \frac{\overline{p}}{p_k}\right)} \text{ for } y \leq y^*\left(p\right) \tag{20}$$

which we can write simply as

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

for functions $\widetilde{f}_k$. Recalling that p_k and $\overline{p}$ do not change when p is replaced with $A_s' p$, substituting this $h_k(y, p)$ expression into equation (14) gives household demand functions for the assignable goods

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

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

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

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

Define the matrix Ω by

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

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

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

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

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

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

It is possible to have models that satisfy the restrictions of both Theorems 1 and 2, by restricting the function $G_t(p)$ in Assumption A3 to only depend on $\widetilde{p}$ and restricting $F_t(p)$ in A3 to only depend on p_t and $\overline{p}$. Such models will be able to exploit comparisons of individuals both within and across households to strengthen the identification.[10]

[10]For examples of models that satisfy such restrictions, we refer our readers to the not-for-publication appendices 7.1 and 7.1.

6.2 Appendix 2: Proofs

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

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

and, in particular,

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

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

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

and by equation (13), we can then identify the resource shares η_{ks} for each household member k by $\eta_{ks} = \rho_{ks}/\left(\rho_{ms} + \rho_{fs} + s\rho_{cs}\right)$.

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

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

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

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

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

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

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

so

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

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

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

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

7 References

Bargain, O. and O. Donni, (2009), "The Measurement of Child Costs: A Rothbarth-Type Method Consistent with Scale Economies," Institute for the Study of Labor (IZA) discussion paper No. 4654.

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., Chen, X. and D. Kristensen (forthcoming), "Semi-Nonparametric IV Estimation of Shape Invariant Engel Curves," *Econometrica*

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

Bourguignon, Browning, and Chiappori (2009), "Efficient Intra-household Allocations and Distribution Factors: Implications and Identification," Review of Economic Studies, 76, 503-528.

Bourguignon, F. and P.A. Chiappori, (1994) "The Collective Approach to Household Behavior," in *The Measurement of Household Behavior*, R. Blundell, I. Preston, and I. Walker, Eds., 70-85, Cambridge: Cambridge University Press

Browning, M., F. Bourguignon, P.A. Chiappori, and V. Lechene (1994), "Incomes and Outcomes: a Structural Model of Intrahousehold Allocation," *Journal of Political Economy*, 102, 1067-1096.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Pendakur, K. (1999). "Semiparametric Estimates and Tests of Base-Independent Equivalence Scales", *Journal of Econometrics*, 88(1), pp 1-40.

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

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

Samuelson, P. (1956), "Social Indifference Curves," The Quarterly Journal of Economics, 70, 1-22.

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