Identification of Semiparametric Model Coefficients, With an Application to Collective Households

Arthur Lewbel and Xirong Lin

Boston College

September 2019

Abtract

We first prove identification of coefficients in a class of semiparametric models. We then apply these results to identify collective household consumption models. We extend the existing literature by proving point identification, rather than the weaker generic identification, of all the features of a collective household (including price effects). Moreover, we do so in a model where goods can be partly shared, and allowing children to have their own preferences, without observing child specific goods. We estimate the model using Japanese consumption data, where we find new results regarding the sharing and division of goods among husbands, wives, and children.

Keywords: Identification, Semiparametric, Collective Household Model, Cost of Children, Bargaining Power, Sharing Rule, Demand Systems

JEL codes: C21, C31, D12, D13.

1 Introduction

A long literature now exists on the identification and estimation of consumption behavior by Pareto efficient collective households. These are households with multiple members, each of whom maximizes a utility function, subject to their claims on the household's resources and a household budget constraint. Almost all of the theoretical results in this literature only show identification of a subset of the model's parameters (e.g., omitting price effects), or only establishes generic identification rather than point identification. In this paper we extend existing collective household identification theorems by proving point identification of all the features of the household's optimization problem, including resource shares and price effects. Moreover, we do so in a model that allows goods in the household to be partly shared, and we identify the extent to which each good is shared. Further, our model reduces data requirements relative to some existing theorems, such as identifying the complete demand functions of children without observing any child specific goods.

To obtain these identification results, we first propose and apply some general methods for proving identification of coefficients in a class of nonlinear and semiparametric models. A few alternative sets of identifying assumptions are provided, thereby giving researchers multiple means by which such models can be identified.

We parameterize and empirically estimate the model using Japanese consumption data on single men, single women, and couples with zero to four children. Among other empirical results, we find that multi-person households save the equivalent of about one fourth of their total expenditures through shared and joint consumption of goods, that wives consume between one fourth to one half of household resources (depending on factors like number of children), and that, to provide for the children, wives forego far more resources relative to husbands when there are children in the household. Failure to account for the extent to which goods are shared, and hence consumed jointly, leads to underestimates of the decrease in the wife's resources relative to that of the husband's when the number of children increases.

We also find that a single adult would need to spend between one third to two thirds as much as a family to attain the same standard of living by themselves as they could attain as a member of a multi-person household (the exact amount depends on the composition of the household). We find that adding one more child to a household with one or two children requires increasing household expenditures by about 8 percent to maintain the children's standard of living.

We begin by considering models of the form $M(p, s) = G(a_{s1}p_1, ..., a_{sJ}p_J)$ where the function M is identified (e.g., M could be a conditional mean function estimated by nonparametric regression), $p = (p_1, ..., p_J)$ is a vector of observed covariates (prices in our application), and s is an observed discrete variable or index. We wish to identify the vector of coefficients $a_s = (a_{s1}, ..., a_{sJ})$ for each value that s can take on. It is important to note that this is NOT an index model. Many results exist for identifying the relative values of coefficients a_{s1}/a_{s2} in linear index models, e.g., models containing terms like $a_{s1}p_1 + a_{s2}p_2$. But those results are not applicable to this context.

Here each p_j appears separately and potentially nonlinearly in the function G, though each p_j appears with a coefficient a_{sj} that varies by the observed discrete s. The relative terms we will be identifying are terms like a_{sj}/a_{tj}, not terms like a_{sj}/a_{sk}.

We first give some alternative sets of assumptions that suffice to point identify the relative coefficients a_{sj}/a_{tj} for $j = 1, ..., J$, or equivalently to identify the coefficients a_{sj} in cases where a_{tj} for some t can be normalized to equal one. These identification results employ variants of methods described by Matzkin (2003, 2007, 2012), Lewbel (1998, 2018), and Lewbel and Pendakur (2017). A useful feature of these identification results is that they do not impose monotonicity on G.

The collective household models we consider have a more general structure than the above G function. In our application, we will have an observable vector of household demand functions that depend on a J-vector of prices of goods p, and where the vector a_s is a set of coefficients that summarize how much goods are shared or jointly consumed. In addition, the household's demands will depend on terms like $\widetilde{\eta}_s^k(p)y$ where $\widetilde{\eta}_s^k(p)$ is a resource share function for household member k, and the index s is a generalization of what the collective household consumption literature calls "distribution factors." In our collective household application, we will need to identify the functions $\widetilde{\eta}_s^k(p)$ as well as the coefficients a_{sj}. The identification of this household model will proceed in multiple steps that repeatedly apply and extend the above general coefficient identification results involving M and G.

Expenditure surveys generally collect consumption data at the level of households. Standard poverty and welfare measurements based on such data are also typically calculated at the household level. But well-being and utility apply to individuals, not households. As a result, there now exists a long literature on identifying features of the consumption, bargaining, and sharing behavior of individuals inside households, using a combination of utility theory, optimization methods, and household level consumption data. These models start with the assumption that household members each have their own utility functions over goods, and that households allocate goods to their members in some way that is Pareto efficient. Important early examples of such models are Becker (1965, 1981) and Chiappori (1988, 1992). Applying standard decentralization results arising from Pareto efficiency, the latter papers show that, regardless of the bargaining or social welfare process the household uses to allocate resources, the behavior of the household is equivalent to the behavior of each household member maximizing his or her own utility function, subject to shadow prices and shadow incomes that reflect the household's chosen resource allocation method.

Of particular interest in these models are resource shares, defined as the fraction of the household's total expenditures (i.e., its budget) that is allocated to each household member. The earlier literature on such models, including Browning, Bourguignon, Chiappori, and Lechene (1994), Browning and Chiappori (1998), Vermeulen (2002), and Chiappori and Ekeland (2006, 2009), showed that, even if one knew all of the demand functions of a household (that is, how much the household would buy of every good as a function of prices, income, and other observed covariates), without additional information one still cannot identify the level of each household member's resources.[1] However, this earlier work also shows that one can usually identify how these resource shares would change in response to a change in observed covariates called distribution factors. Distribution factors are variables that affect the bargaining power of household members, and so affect their resource shares, but do not affect the tastes of household members. Papers that make use of this identification result include Bourguignon and Chiappori (1994), Chiappori, Fortin, and Lacroix (2002), and Blundell, Chiappori, and Meghir (2005).

One limitation of these resource share identification theorems is that they are based on household models that constrain goods to be either purely private or purely public within a household, meaning that each good is either completely jointly consumed by all household members (like heat) or completely privately consumed (like food, e.g., no two people can eat the same apple). We relax this restriction by working with a more general model based on Browning, Chiappori, and Lewbel (2013), that allows goods to be partly shared. An example of a partly shared good could be gasoline, which is privately consumed when one person uses a car by him or herself, but is jointly consumed by more than one household member when those members ride in the car together.

A second limitation of these earlier resource share identification theorems is that they only prove generic identification. Roughly, generic identification means that models are usually identified, but there exists special situations where identification fails (see McManus 1992 and Lewbel 2018 for the formal definition of generic identification).

Our first collective household identification theorem extends the classical resource share identification theorem proving identification of changes in resource shares in response to changes in distribution factors, by relaxing the above restrictions. That is, we show point identification rather than just generic identification, and we prove the result in the context of the Browning, Chiappori, and Lewbel (2013) (hereafter BCL) model that allows goods to partly shared. Moreover, we further allow the same covariates to affect both resource shares and the extent to which goods are shared.

In response to the results from the earlier literature that only changes in resource shares and not levels can be identified from household data, a more recent literature has focused on adding additional assumptions to the model to identify the levels of resource shares.[2] For example, in a model without children, BCL obtain generic identification by assuming common preferences over goods for individuals whether single or married. Other papers that impose additional functional form or behavioral restrictions to gain point identification include Lewbel and Pendakur (2008),

[1]These results are often presented regarding Pareto weights (the weights placed on the utility functions of each household member in the household's optimization problem) rather than resource shares. Likewise, distribution factors can be equivalently defined in terms of such weights rather than in terms of resource shares. However, as Browning, Chiappori, and Lewbel (2013) show, resource shares are monotonic in Pareto weights and vice versa, so each, along with individual member utility functions, contain comparable information about the household. However, resource shares have both more direct economic implications, and do not depend on the arbitrary cardinalization of utility functions.

[2]One response to the nonidentification of resource share levels has been the collection of costly (and hence small) data sets of extremely detailed within household consumption. Examples of constructing resource share estimates using such detailed data include Menon, Perali, and Pendakur (2012) and Cherchye, De Rock, and Vermeulen (2012). Another response has been to construct revealed preference based set identification bounds on resource shares. Examples include Cherchye, De Rock, and Vermeulen (2011), Cherchye, De Rock, Demuynck, and Vermeulen (2017), and Cherchye, De Rock, Lewbel, and Vermeulen (2015).

Couprie, Peluso, and Trannoy (2010), Bargain and Donni (2009, 2012), Lise and Seitz (2011), and Dunbar, Lewbel, and Pendakur (2013).

All of the above results that obtain point identification, rather than just generic identification, depend either on very strong functional form restrictions, or only show point identification of some features of the household's behavior (e.g., identifying resource shares but not price effects). Our second collective household identification theorem point identifies all the features of the household's behavior, again allowing for partial sharing of goods. The theorem includes point identification of the demand functions and resource shares of children, without requiring observation of any child specific goods.

2 The Collective Household Model With Cooperation Factors

Resource shares are functions that describe the allocation of a household's total budget to each of the household's members. A distribution factor is a covariate that affects resource shares, but does not affect household member's tastes for goods. Lewbel and Pendakur (2019) propose a generalization of distribution factors called "cooperation factors." A cooperation factor affects both resource shares and the extent to which each good the household consumes is shared or jointly consumed.[3] Earlier collective household models maintained the unrealistic assumption that all goods were either purely public or purely private within a household. These models therefore could not permit any variation across households in how much each good was shared or jointly consumed, and so could not contain cooperation factors.

The covariate s in our model is the value of a cooperation factor. An example of a cooperation factor might be the number of children in the household, where we assume a single utility function for all children. The sharing of goods within a household varies with the number of children, as does the fraction of the household's total resources that are devoted to children.

The economic motivation for uncovering the unobserved resource share and allowing goods to be partly shared is to enable individual-level welfare analysis. In particular, we are interested in comparing individuals' welfare under different economic environments (like household size and composition), or the so-called "indifference scale" proposed by Browning et al. (2013). It measures the fraction of household total expenditure y required by an individual household member k purchasing goods privately, to be as well off materially as he or she is while living with others in a household that has joint income y. This is in contrast to the previous literature on the household-level welfare analysis, or "equivalence scale", which directly compares welfare across households. This method suffers from the cardinalization of utility functions as well as the critique that there is no household utility but only individual utility. Instead, indifference scale

[3] In Lewbel and Pendakur (2019), cooperation factors can also directly affect the utility functions of household members. That additional feature of cooperation factors is irrelevant for the present paper, because that feature only affects the value of s, but does not affect the household's demand functions or resource shares as functions of p, y, and s.

does not suffer from these problems. Indifference scales can be used for poverty, life insurance, and wrongful death calculations. For example, Cherchye et al. (2012) apply indifference scales to Dutch data and study the poverty and economic well-being among the elderly widows and widowers.

This section summarizes the derivation of our collective household consumption model, which is the Browning, Chiappori, and Lewbel (2013) model with resource shares independent of budget y, Barten consumption technology, and without observing singles.[4] The resource share functions $\widetilde{\eta}_s^k(p)$ for household member k and Barten coefficients a_{sj} for good j each depend on the cooperation factor s.

A household consists of K members. Let subscript j denote a good and superscript k denote a household member. Let z index the continuous quantities of goods purchased by the household. The household solves the following optimization problem

$$\max_{x^1,\ldots,x^K} \sum_{k=1}^{K} \mu^k(p/y)U^k(x^k) \tag{1}$$

$$\text{such that } z = Ax,\ x = \textstyle\sum_{k=1}^{K} x^k,\ p'z = y$$

U^k is member k's utility function. We allow household members to have different preferences. μ^k is the so-called "Pareto weight" of each member. It summarizes the member's bargaining power in a collective model. A higher Pareto weight implies that the household demand is represented more by the member's preferences. $p'z = y$ is the household's budget constraint.

Each member's utility function depends on the private good equivalents that the member consumes. $x^k = (x_1^k, ..., x_J^k)$ is the vector of member k's private equivalent consumption of goods. They are the quantities of transformed goods that are consumed by each member. x is the sum of private good equivalents for all members, i.e., $x = \sum_{k=1}^{K} x^k$. $z = Ax$ is the household "consumption technology function". The difference between z and x is due to the sharing and jointness of consumption. The square matrix A summarizes how much goods are shared or jointly consumed. The diagonal elements of A represents how much each good can be shared by itself. For example, suppose that the first element of x^k is the quantity of gasoline consumed by member k. If all household members shared their car (riding together) by 1/3 of their time, then in terms of the total distance traveled by each household member, it is as if member 1 consumed a quantity of g_1^1 of gasoline and member 2 consumed a quantity of g_1^2, where $z_1 = (2/3)(g_1^1 + g_1^2)$. In this example, the upper left corner of matrix A would be 2/3 and the remaining first row and first column of A would be zero. The off-diagonal element of A represents how much the sharing of one good depends on the consumption of another good. For example, a household that consumes more public transportation will have a lower degree of sharing in gasoline. For simplicity, we assume the off-diagonal elements of A to be zero.

[4] The consumption technology is called a Barten technology after Barten (1956), who proposed an analogous construction to model preference heterogeneity across consumers in unitary (not collective) models.

The key assumption in the collective household literature is that the household outcomes are Pareto efficient. From the second welfare theorem, any Pareto efficient outcomes can be implemented as an equilibrium of the economy, possibly after some lump sum transfers between members. Hence, the duality of the above household program is equivalent to a two-stage process. In stage one, household's total expenditure is divided between members according to some sharing rule $\eta_s^k(p)$, which is the fraction of resources enjoyed by member k. In stage two, each member k chooses her or his private equivalent consumption x^k to maximize her or his own utility U^k given a Lindahl (Lindahl 1958) type shadow price vector π and resource share. The Lindahl shadow price π is the market price p discounted by the degree of sharing or jointness of consumption, that is, $\pi = p'A$. The stage two can be formalized using the optimization problem below

$$\underset{x^k}{max}\, U^k(x^k) \text{ such that } (a_{s1}p_1, ..., a_{sJ}p_J)x^k = \eta_s^k(p)y \tag{2}$$

and the solution to the above individual's problem is

$$x^k = g^k(a_{s1}p_1, ..., a_{sJ}p_J, \eta_s^k(p)y) \tag{3}$$

The function $g^k\,(p, s, y)$ is the Marshallian demand function by member k, obtained by maximizing member k's utility function $U^k(x^k)$ subject to a linear budget constraint.

From this derivation, the household's demand functions have the form

$$\omega_j\,(p, s, y) = \sum_{k=1}^{K} \widetilde{\eta}_s^k(p) h_j^k\left(a_{s1}p_1, ..., a_{sJ}p_J, \widetilde{\eta}_s^k(p)y\right) \tag{4}$$

The function $\omega_j\,(p, s, y)$ is the household's demand function for good j, defined as the total expenditures of the household on good j, divided by the household's expenditures on all goods y. h_j^k is member k's demand function for good j, defined as the member's expenditure on good j, if faced with the shadow price and resource share, divided by the member's control of resources $\eta_s^k(p)y$.

The household's resource share functions $\widetilde{\eta}_s^k(p)$ and the Barten scales a_{sj} vary by s. Since each a_{sj} must be strictly positive, and the functions $\widetilde{\eta}_s^k$ can vary with s, we can without loss of generality define resource share functions $\eta_s^k(p)$ such that

$$\eta_s^k(a_{s1}p_1, ..., a_{sJ}p_J) = \widetilde{\eta}_s^k(p) \tag{5}$$

Writing resource shares in the form of η_s^k rather than $\widetilde{\eta}_s^k$ simplifies some of our later identification proofs.

Recall that A_s is a diagonal matrix with the vector of coefficients $(a_{s1}, ..., a_{sJ})$ on the diagonal.

We can then rewrite equation (4) as

$$\omega_j(p,s,y) = \sum_{k=1}^{K} \eta_s^k(A_s p) h_j^k\left(A_s p, \eta_s^k(A_s p) y\right). \tag{6}$$

The general problem to be considered is identification of the Barten scales a_{sj}, the resource share functions η_s^k, and the individual member demand functions h_j^k, given the observable (or consistently estimable) household demand functions ω_j. Note that if the functions η_s^k and the vector of Barten coefficients $(a_{s1}, ..., a_{sJ})$ are identified, then the alternative way to represent resouce shares given by the functions $\widetilde{\eta}_s^k$ are also immediately identified by equation (5).

For both identification and estimation, we will make use of the concept of assignable goods. A good is assignable if it is only consumed by one household member. Suppose, e.g., that for some household member k, good $j = k$ is assignable to member k. Then for $j = k$, equation (6) simplifies into

$$\omega_k(p,s,y) = \eta_s^k(A_s p) h_k^k\left(A_s p, \eta_s^k(A_s p) y\right) \tag{7}$$

Chiappori and Ekeland (2009) show that, without partially shared goods, one can obtain generic identification of the model given assignable goods. Dunbar, Lewbel, and Pendakur (2013) combine having assignable goods with some preference similarity restrictions to point identify resource shares (but not other features of the model, in particular, they use Engel curve data and so cannot identify any price effects).

3 Semiparametric Coefficients Identification

Before we tackle the general problem of identifying the components of equation (6), we first consider a simpler problem. Let $a_s = (a_{s1}, ..., a_{sJ})$ be a J-vector of coefficients we wish to identify. Let A_s be the J by J diagonal matrix that has the vector a_s on the diagonal. Let $P = (P_1, ..., P_J)$ be a J-vector of continuous covariates and let S be a discrete covariate.

Assume we can identify a function $M(P,S)$, e.g., $M(P,S)$ might be a conditional mean, conditional density, or conditional quantile function that we could consistently estimate. In our applications, P is a vector of prices, S is a so-called cooperation factor, and M is derived from household demand functions that can be estimated from observed consumption data. The goal is to identify the unknown vector of coefficients $a_s = (a_{s1}, ..., a_{sJ})$ in the model

$$M(p,s) = G(a_{s1}p_1, ..., a_{sJ}p_J) = G(A_s p) \tag{8}$$

for some unknown function G.

In this section we will provide a simple theorem that provides three alternative sets of conditions that suffice for point identification of the vector of coefficients $(a_{s1}, ..., a_{sJ})$ for each value s that S can equal. An attractive feature of these identification results is that they will not impose monotonicity on the function G. These simple results will form alternative building blocks we then use to identify the collective household model. These theorems are variants of identification methods described by Matzkin (2003, 2007, 2012), Lewbel (1998, 2018), and Lewbel and Pendakur (2017).

ASSUMPTION A1: Let the support of (P, S) be $\Omega_p \times \Omega_s$. For each $(p, s) \in \Omega_p \times \Omega_s$, equation (8) holds for some unknown function G and some vector of constants $a_s = (a_{s1}, ..., a_{sJ})$. The function $M(p, s)$ is identified for all $(p, s) \in \Omega_p \times \Omega_s$.

ASSUMPTION A2: Assume for some $t \in \Omega_s$ that $a_{tj} = 1$ for $j = 1, ..., J$.

Assumption A1 essentially just lays out the model. Assumption A2 is a scale normalization. In some contexts, Assumption A2 can be made without loss of generality (as long as a_{tj} is not identically zero). This is because, if $a_{tj} \neq 1$ then we can redefine the function G to make $a_{tj} = 1$, by replacing G with $\widetilde{G}$ defined by $\widetilde{G}(p) = G(a_{t1}p_1, ..., a_{tJ}p_J)$ and replacing each a_{sj} with $\widetilde{a}_{sj}$ defined by $\widetilde{a}_{sj} = a_{sj}/a_{tj}$.

We will propose three assumptions (Assumption A3, A4, and A5), each of which can be used to obtain identification. Assumption A3 is a high level assumption, which may therefore be hard to verify in practice. Assumption A4 is more restrictive, but is simple and low level, and therefore could be easier to verify or justify in applications. Assumption A5 is discussed below.

ASSUMPTION A3: Assume $G(p)$ is continuously differentiable. Let $m_j(p, s) = \partial M(p, s) / \partial p_j$ and let $g_j(p) = \partial G(p) / \partial p_j$. For any J-vector $\alpha = (\alpha_1, .., \alpha_J)$, define the J-vector valued function $\zeta(\alpha, p, s)$ as having the elements

$$\zeta_j(\alpha, p, s) = \frac{m_j(p, s)}{g_j(\alpha_1 p_1, ..., \alpha_J p_J)} \text{ for } j = 1, ..., J$$

For each $s \in \Omega_s$, assume there exists a $\widetilde{p} \in \Omega_p$ such that $A_s\widetilde{p} \in \Omega_p$ and $\zeta_j(\alpha, p, s)$ is a contraction on a.

ASSUMPTION A4: Assume Ω_p includes a neighborhood of zero, and that $G(p)$ is continuously differentiable for all p in that neighborhood of zero. Assume $\partial G(p) / \partial p_j$ does not equal zero when $p = 0$.

In Assumption A4, the neighborhood of Ω_p containing zero can be one sided, by just using one sided derivatives and limits in the proof of Lemma 1 below. So, e.g., p in our later application will be prices, which are nonnegative. But if arbitrarily low prices (relative to expenditure levels) can be observed in theory, then Lemma 1 can be applied, taking one sided limits and derivatives as p goes to zero.

Define the random vector V by $V = (V_1, ..., V_J)$ where $V_j = a_{Sj}P_j$. Let Ω_v denote the support of V.

LEMMA 1: Let Assumptions A1 and A2 hold. If either Assumption A3 or Assumption A4 also holds then the coefficients $a_{s1}, ...a_{sJ}$ and the function $G(v)$ are point identified for all $v \in \Omega_v$ and $s \in \Omega_s$.

The identification in Lemma 1 is what Khan and Tamer (2010) call "thin set" identification. Thin set identification is when identification is based on a measure zero subset of the support of the data. In this example, identification is based either on the point p that makes Assumption A3 hold, or the point $p = 0$ for Assumption A4. Either such point is observed with probability zero if P is continuous. The more well known concept of "identification at infinity" as in Chamberlain (1986) and Heckman (1990) is another example of thin set identification. Many of the identification theorems given in Matzkin (2003, 2007, 2012) assume a normalization that otherwise unknown functions take on known values at one point, such as zero. Such normalizations typically imply thin set identification. In practice, estimators of parameters that are only thin set identified will usually converge at slow rates[5].

One way to avoid thin set identification is to assume that Assumption A3 holds at a mass point p. Another way would be to assume that Assumption A3 holds for all points p in some convex positive measure subset of Ω_p. However, this is an additional strong high level assumption that could be difficult to verify.

To avoid issues associated with thin set identification, we now give an alternative identification result that integrates over the support of p. This identification however, requires a large support assumption. However, unlike identification at infinity or other thin set identification (and associated convergence rate issues), here the large support assumption is only needed to avoid the presence of boundary terms in a change of variables argument.

For a given function ψ_j, define c_j by

$$c_j = \int_0^\infty ... \int_0^\infty \psi_j \left[G(p)\right] p_1^{-1} ... p_{j-1}^{-1} p_{j+1}^{-1} ... p_J^{-1} dp_1 ... dp_J \tag{9}$$

ASSUMPTION A5: Assume Ω_p is the positive orthant. $G(p)$ is continuous for all $p \in \Omega$. All a_{sj} are positive. For each $j \in \{1, ..., J\}$, we can find a continuous function ψ_j such that the constant c_j defined by equation (9) exists, is finite, and non-zero.

Having Ω_p be the positive orthant is the large support assumption. The assumption that all a_{sj} are positive is testable, using the estimated average derivatives with respect to p_j of $M(p, s)$ relative to average derivatives of $M(p, t)$. In our empirical application, the a_{sj} coefficients will be sharing parameters that are positive by construction. It is assumed that we can find a continuous function ψ_j that makes the integral given by equation (9) convergent. Note that $G(p)$ is identified by $G(p) = M(p, t)$, so knowing G, the assumption is that we can construct a continuous function ψ_j that goes to zero sufficiently quickly whenever any element of P goes to zero, and grows

[5] See Khan and Tamer (2010) and Lewbel (2018) for details regarding thin set identification.

sufficiently slowly, or not at all, when any element of P goes to infinity. [6]

LEMMA 2: If Assumptions A1, A2, and A5 hold, then the coefficients $a_{s1}, ...a_{sJ}$ and the function $G(v)$ are point identified for all $v \in \Omega_v$ and $s \in \Omega_s$.

Both Lemmas 1 and 2 have proofs by construction, so semiparametric estimators could be readily constructed by mimicking the steps of either proof. Combining Lemmas 1 and 2, and separately considering the scale normalization of Assumption A2 gives us our first identification theorem.

THEOREM 1: Let Assumption A1 hold. If either Assumption A3, A4, or A5 also holds, then the relative coefficients $a_{s1}/a_{t1}, ..., a_{sJ}/a_{t1}$ are point identified for all $v \in \Omega_v$, $s \in \Omega_s$, and $t \in \Omega_s$. If Assumption A2 also holds then the coefficients $a_{s1}, ..., a_{sJ}$ and the function $G(v)$ are point identified for all $v \in \Omega_v$ and $s \in \Omega_s$.

4 Identification of the Collective Household Model

We now consider identification of the collective household model. The results here are variants and applications of Theorem 1. As with Theorem 1, we will present pairs of results that allow us to obtain point identification either making use of large support assumptions or entailing possible thin set identification.

ASSUMPTION B1: Household budget share demand functions $\omega_j(p, s, y)$ for $j = 1, ..., J$ are given by equation (6), where for all $(p, s, y) \in \Omega_p \times \Omega_s \times \Omega_y$, the functions $h_j^k(p, y)$ and $\eta_s^k(p)$ are continuous for each member k and cooperation index s. The Barten technology constants a_{sj} are bounded and strictly positive for each cooperation index s and good j.

Assumption B1 essentially lays out the collective household model as discussed in the previous section. The continuity conditions follow naturally from smooth utility and household bargaining or social welfare functions. Similarly, having Barten scales be bounded and positive follows from physically feasible sharing.

Our first goal is to identify the Barten constants a_{s1},...,a_{sJ}. We cannot immediately apply Theorem 1 to equation (6) or equation (7) (taking G to be any of the household demand functions ω_j), because the resource shares $\eta_s^k(A_s p)$ vary by s. We therefore will first construct a function G out of a demand function ω_j using Theorem 2 below, and then apply Theorem 1 to the result.

Assumptions B2, B3, and B4 below are alternatives; only one needs to hold for Theorem 2. These assumptions each resemble either Assumption A3 or A5 from Theorem 1, and Theorem 2 correspondingly uses similar machinery to that of Theorem 1. But instead of directly identifying coefficients, Theorem 2 is used to modify the demand function of equation (6) into a form needed to apply Theorem 1.

[6] A similar construction appears in Lewbel and Pendakur (2017), who also provide some examples. However, their application involved much stronger conditions than ours, because in their model the coefficients were random rather than constants.

ASSUMPTION B2: Assume that Ω_y includes a neighborhood of zero. Assume there exists a good j that is assignable to some household member k. Assume that for this assignable good j, for all $(p,s) \in \Omega_p \times \Omega_s$, the function $M(p,s)$ defined by the following equation is finite and nonzero

$$M(p,s) = \lim_{y \to 0} \frac{1}{\omega_j(p,s,y)^2} \frac{\partial \omega_j(p,s,y)}{\partial y}$$

ASSUMPTION B3: Assume that Ω_y includes $(0,\infty)$. Assume there exists a good j that is assignable to some household member k. Assume that for this assignable good j, for all $(p,s) \in \Omega_p \times \Omega_s$, the function $M(p,s)$ defined by the following equation is finite and nonzero for some real nonzero constant c.

$$M(p,s) = \int_0^\infty [\omega_j(p,s,y)]^c y^{c-1} dy.$$

ASSUMPTION B4: Assume that Ω_y includes $(0,\infty)$. Assume there exists a good j such that, for all $(p,s) \in \Omega_p \times \Omega_s$, the function $M(p,s)$ defined by the following equation is finite and nonzero.

$$M(p,s) = \int_0^\infty \omega_j(p,s,y)\, dy.$$

THEOREM 2: Let Assumptions B1 hold. If Assumption B2 or B3 or B4 also holds, then there exists a function $G(p)$ such that $M(p,s) = G(A_s p)$.

COROLLARY 1: Let Assumption B1 hold. If Assumption B2 or B3 or B4 also holds, and if corresponding $M(p,s) = G(A_s p)$ equation from Theorem 2 satisfies Assumption A1 and either Assumption A3 or A4 or A5, then a_{sj}/a_{tj} is identified for every $s \in \Omega_s$, every $t \in \Omega_s$ and every $j \in \{1, ..., J\}$.

Theorem 2 shows that equation (8) holds, and so Theorems 1 and 2 can be combined as in Corollary 1. Corollary 1 shows that all the relative Barten scales a_{sj}/a_{tj} are identified. In this context the scale normalization of Assumption A2 is not a free normalization, because each a_{sj} has a physical economic meaning as the extent to which good j is shared in a household with cooperation factor s. However, we will later use additional information to identify the levels of the Barten scales and not just their relative values.

Note for Theorem 2 that Assumptions B2 and B3 require an assignable good, while B4 does not. However, our identification of resource shares below will require an assignable good regardless, so B4 is mainly useful if the primary goal is just identification of the Barten scales, or if some other mechanism like functional form restrictions are used to identify the resource shares. Assumption B4 is weaker than B3 in that it doesn't require an assignable good, but is stronger in that it requires the constant c to equal one.

Assumption B2 implicitly assumes that for the given good j, $\lim_{y\to 0}\omega_j(p,s,y)$ is nonzero. This limit would be zero if ω_j was a quantity, but ω_j is a budget share. This condition is not a strong constraint for a budget share, and so would hold if, e.g., the budget share for good j was bounded away from zero for $y>0$ (given continuity). We can expect this condition to hold for most goods, but in particular for necessities, since such goods are, definitionally, necessary and hence comprise a nonzero share of the household's budget.

A notable feature of Theorem 2 is that it gets identification from the demand function of just one good that the household consumes. Since we can estimate household demand functions for many goods, we can expect the Barten scales to be greatly over identified in practice. Another feature is that these results do not require monotonicity of demands, which is useful because empirically the effects of both p and y on budget shares can change signs.

Given identification of the Barten technology, our next goal is identification of relative resource shares. Define the vector $\phi_{st}(p)$ to be the vector of elements $\phi_{stj}(p_j)$ defined by

$$\phi_{stj}(p_j) = \frac{p_j}{a_{sj}/a_{tj}}$$

where $t \in \Omega_s$ is any nonzero cooperation factor value chosen by the econometrician.

ASSUMPTION C1: Assume that Ω_y includes a neighborhood of zero, that there exists a good j that is assignable to some household member k, and for that good j the budget share function $\omega_j(s,\phi_{st}(p),0)$ is finite and nonzero for all $(p,s)\in\Omega_p\times\Omega_s$.

ASSUMPTION C2: Assume that Ω_y includes $(0,\infty)$, that there exists a good j that is assignable to a household member k, and for that good j, for all $(p,s)\in\Omega_p\times\Omega_s$, the function $m(p,s)$ defined by the following equation is finite and nonzero for some real constants c_1 and c_2 where $c_2 \neq c_1 - 1$ and $c_1 \neq 0$.

$$m(p,s) = \int_0^\infty \left[\omega_j(s,\phi_{st}(p),y)\right]^{c_1} y^{c_2} dy.$$

THEOREM 3: Let the Assumptions of Corollary 1 hold for some $s\in\Omega_s$ and let them also hold replacing s with some other value $r\in\Omega_s$. If in addition either Assumption C1 or C2 holds, then the relative values of resource shares $\eta_s^k(A_tp)/\eta_r^k(A_tp)$ are identified for all p such that $(a_{s1}\phi_{st1}(p_1),...,a_{sJ}\phi_{stJ}(p_J))$ and $(a_{r1}\phi_{rt1}(p_1),...,a_{rJ}\phi_{rtJ}(p_J))$ lie in Ω_p. If Ω_p is the positive orthant, then $\eta_s^k(A_tp)/\eta_r^k(A_tp)$ is identified for all $p\in\Omega_p$.

The classical identification result in the collective household literature discussed earlier, and given in its most general form by, e.g., Chiappori and Ekeland (2006, 2009) is that, without additional information, the level of resource shares were not identified, but the changes in the resource shares resulting from changes in distribution factors are generically identified. However, this classical model required that all goods be either completely private within the household, or completely public. Theorems 2 and 3 together generalize this classical result to the model where goods can be partly shared as in BCL, and where the extent to which goods are shared can vary across households. In particular, these theorems show identification of relative Barten scales and relative resource shares, and so show that changes in these functions are identified given a change in the cooperation factor s. Moreover, these theorems here give explicit conditions for point identification of these relative values, rather than just the generic identification of earlier results (see, e.g., Lewbel 2019 for the difference between point and generic identification).

One potential limitation of Theorem 3 is that, if Ω_p is not the positive orthant, there could exist values of p for which identification of the relative resource shares is not shown. However, the identification in Theorem 3 uses just the demand function of one good for each household member. Since the demand functions for many goods are observed, as with Theorem 2 we can in general expect substantial overidentification, based on information in the other goods the household consumes.

For many applications, identification of relative values, particularly of resource shares, does not suffice to answer some questions of economic significance. E.g., as stressed by Dunbar, Lewbel, and Pendakur (2013), identification of poverty rates and of relative bargaining power of household members requires identifying the levels of resource shares, not just their relative values.

Therefore, for the last part of this section, we consider using additional information to obtain identification of the entire model, including levels of resource shares, levels of Barten scales, and the demand functions of each household member. These results will also allow us to relax the assumption that a private assignable exists for every household member.

ASSUMPTION D1: For some household member $k \in \{1, ..., K\}$ assume there exists an assignable good j. Without loss of generality let $j = k$. Assume that the demand function $h_k^k(p, y)$ is identified.

Letting $j = k$ in Assumption D1 is a free index normalization. What Assumption D1 says is that the demand function for one household member's assignable good is identified. The easiest way for Assumption D1 to hold is if our data includes single person households, and the demand function for an assignable good consumed by member k is the same whether that person lives alone or with other people. For example, if member k is a middle aged man, then let s_k denote the value of s that indexes households consisting of a middle aged man living alone. Since there is no one to share with when living alone, $\eta_{s_k}^k$ must equal one and $a_{s_k j}$ must equal one for all goods j. It follows that $\omega_k(p, s_k, y) = h_k^k(p, y)$, which then identifies $h_k^k(p, y)$.

THEOREM 4: Let the Assumptions of Corollary 1 hold for all $s \in \Omega_s$, let either Assumption C1 or C2 hold, and let Assumption D1 hold. Then the Barten constants a_{s1},...,a_{sJ} are identified for all $s \in \Omega_s$, and the relative resource shares $\eta_s^k(p)/\eta_r^k(p)$ are identified for all $p \in \Omega_p$. If in addition Assumption D1 holds for $k = 1, ..., K-1$, then each $\eta_s^k(p)$ function is identified for $k = 1, ..., K$.

COROLLARY 2: Let the Assumptions of Corollary 1 hold for all $s \in \Omega_s$. If either Assumption C1 or C2 holds, and if Assumption D1 holds for $k = 1, ..., K-1$, then the entire model is identified.

What we mean by the entire model being identified in Corollary 2 is that all the Barten scales a_{sj}, all the resouce share functions $\eta_s^k(p)$, and all the demand functions $h_j^k(p, y)$ are identified. Note that Corollary 2 follows immediately from Theorem 4, because once all the Barten constants and resource share functions are identified, we may then from equation (6) obtain the demand functions $h_j^k(p, y)$ for each good j.

By Theorem 4, only one assignable good for one household member is needed to identify the levels of the Barten constants. To identify the levels of the resource shares, and hence identify the entire model by Corollary 2, we require $K - 1$ assignable goods. So, e.g., if $K = 3$ where $k = 1$ is the father, $k = 2$ is the mother, and $k = 3$ is the children, then we only need to have one identified, assignable good for the mother and for the father. As discussed above, these could come from observing single men and single women, assuming that one's taste for the assignable good does not differ between those living with others versus those living alone. In this example we do not need to observe or identify any child assignable goods, which is very useful because we would not expect to observe households consisting of children living alone (the original BCL model did not include children because, unlike the present paper, it did not overcome this obstacle to identification with children).

5 Empirical Application

5.1 Japanese Expenditure Data

We use Japanese household expenditures and demographic data. The data come from the Keio Household Panel Survey (KHPS) and the Japan Household Panel Survey (JHPS), made available to us by the Panel Data Research Center at Keio University. The KHPS has been implemented continuously since 2004, and consists of 4,000 households and 7,000 individuals nationwide. An additional survey on a cohort of about 1,400 households and 2,500 individuals was initiated in 2007. In 2009, the Panel Data Research Center at Keio University began implementing the JHPS, a new survey targeting 4,000 male and female subjects nationwide in parallel with the KHPS.

The survey questionnaires cover comprehensive topics such as household structure, individual attributes, academic background, employment or education status, distribution of living hours, and matters related to cohabitation with parents, etc. Households are asked the following questions regarding household expenditures, "Enter the amount your household spent on each of the following living expenditures last month (January)." The expenditure categories that we include in this paper are food (at-home or eating-out and school lunches), utilities, clothing and shoes, transportation, communication, and entertainment, giving us a total of $J = 6$ different goods.

The consumption data separately reports household expenditures (in January) on clothing and shoes for the household head, spouse(s), and children. The sum of expenditures on clothing and shoes for each household member type (men, women, and children) are our private assignable goods. Note that while the data include assignables for all $K = 3$ types of household members, our identification theory only requires observation of $K - 1 = 2$ assignable goods. This provides over identifying information.

We select households (single men, single women, and married couples) according to the following criteria: (1) couples with children aged 15 or over are excluded (since adult clothing purchases could be consumed by older children); (2) for married couples, households with members over 50 are excluded; (3) single women and men are restricted to be between 22 to 65 years old; (4) households with members as students are excluded; (5) observations where expenditures on four or more of the six goods is zero are excluded; (6) To mitigate the possible effects of outliers, we further trim the three samples with respect to key variables (the budget share of each aggregate good and log real total expenditure) by dropping observations in the lower and upper 1 percentile. After applying these criteria, we are left with a sample consisting of 277 single women, 361 single men, and 1070 married couples having from zero to four children.

Table 1 provides summary statistics of demographic characteristics, general expenditure, and our private assignable goods for our sample of singles and couples with 0-4 children.

5.2 Price Data

We use price data from the 2015 based Consumer Price Index (CPI) available from e-Stat, the Portal Site of Official Statistics of Japan. The goal is to construct a price index for each aggregate good for each household in our sample. It is challenging to merge this CPI data into the JHPS/KHPS because the two datasets divide the country somewhat differently. JHPS/KHPS provides the region and city size of the residence of each household. The CPI divides Japan into 10 regions, whereas the JHPS/KHPS divides it into 8 regions. We first reduce the number of regions in the CPI by merging some of the CPI regions to match the definitions in JHPS/KHPS. While most prefectures belong to the same region between the CPI and JHPS/KHPS data after merging, the three prefectures of Yamanashi, Nagano, and Mie are classified to different regions between the CPI and JHPS/KHPS data.[7]

In addition to regional prices, the CPI dataset provides price data for each "designated city," that is, each major city with a population of more than half million that is designated as such by order of the Cabinet of Japan.[8] Combining these city level prices using CPI weights, we construct price indices for designated cities within each of the eight regions, except for the Shikoku region where there is no designated city. Using each regional price index and the price indices for

[7] To match the JHPS/KHPS definition of Kyushu region (Fukuoka, Saga, Negasaki, Miyazaki, Kagoshima, Kumamoto, Oita, and Okinawa prefectures), we merged Kyushu and Okinawa regions in CPI. To match the JHPS/KHPS definition of Chubu region (Yamanashi, Nagano, Niigata, Fukui, Toyama, Ishikawa, Shizuoka, Gifu, and Aichi prefectures), we also merge Hokuriku and Tokai prefectures. With these merging, most prefectures belong to the same region between the JHPS/KHPS and CPI datasets with the following exceptions: Yamanashi and Nagano prefectures belong to Kanto [Chubu] region in CPI [JHPS/KHPS] dataset, and Mie prefecture belongs to Chubu [Kinki] region in CPI [JHPS/KHPS] dataset. About 3.7 percent of the Japanese population live in these three prefectures, according to the 2015 population census. See also, http://www.stat.go.jp/english/data/kokusei/2015/final_en/final_en.html. This procedure follows Fujii and Lin (2018).

[8] There are 20 designated cities in Japan as of January 1, 2019.

Table 1: *Summary Statistics, JHPS/KHPS 2004 - 2016*

	Single Men	Single Women	Couples with			
			0 child	1 children	2 children	3 - 4 children
Number of observations	1,194	830	379	711	1,376	396
Number of unique households	361	277	195	281	458	139
Household income	346.37	.	750.01	590.89	652.72	640.20
Total expenditures (month)	124.17	114.43	182.80	180.35	192.35	201.81
Budget share (food)	0.45	0.40	0.34	0.35	0.36	0.38
Budget share (clothing)	0.05	0.08	0.09	0.08	0.07	0.07
Budget share (communication)	0.11	0.12	0.12	0.13	0.12	0.13
Budget share (entertainment)	0.18	0.16	0.23	0.22	0.23	0.22
Budget share (transportation)	0.08	0.09	0.09	0.09	0.07	0.07
Budget share (utility)	0.13	0.15	0.13	0.14	0.14	0.14
Husband clothing&shoes share	-	-	0.04	0.02	0.01	0.01
Wife clothing&shoes share	-	-	0.06	0.02	0.02	0.01
Children clothing&shoes share	-	-	0.00	0.03	0.04	0.04
Female age	-	47.13	38.33	37.79	38.36	38.25
Female unemployed	-	0.11	0.10	0.23	0.23	0.22
Female college graduate or above	-	0.20	0.07	0.10	0.10	0.07
Female some college	-	0.40	0.33	0.30	0.28	0.21
Male age	48.05	-	39.23	39.10	39.89	39.29
Male unemployed	0.46	-	0.01	0.00	0.00	0.00
Male college graduate or above	0.19	-	0.07	0.10	0.07	0.10
Male some college	0.46	-	0.39	0.27	0.26	0.30
Child 1 age	-	-	-	6.79	9.71	11.42
Child 2 age	-	-	-	-	6.50	8.69
Child 3 age	-	-	-	-	-	5.35
Child 4 age	-	-	-	-	-	-
Child average age	-	-	-	6.79	8.11	8.34
Home ownership	0.36	0.41	0.49	0.59	0.73	0.80

Notes: income and expenditures are in thousand yen. JHPS/KHPS covers years 2004 - 2016. Expenditures are for January. Definition of aggregate goods in JHPS/KHPS: food expenditure includes eating out. Transportation includes automobile expenses, fares, commuting passes, taxes, and tolls. Communications includes postage, fixed-line, and mobile phone charges. Culture & amusement includes stationery, sporting goods, travel, hobbies. Utility includes electricity, gas, water (supply & sewage). Clothing includes both clothese and shoes. All sources of income are before tax in the past year. "." means observations are all missing for this variable. "-" means information not available/not applicable. For education variable, college graduate or above in JHPS/KHPS includes junior college or technical college, univeristy, or graduate school. Household income refers to annual take-home income (after tax and social insurance deductions).

designated cities, we additionally back out price indices for the areas outside each designated city in each region. Thus, for each aggregate good, we obtain price data for 15 (8 regions $\times$ 2 (designated city or not) $-$ 1 (no designated city in Shikoku region) combinations of regions and city sizes, which we then assign to households in the JHPS/KHPS dataset.

In the food category, the CPI dataset has separate price indices for food-at-home and eating-out. We construct household-level price indices for food using a Stone price index, by taking a weighted average of the log of the price of eating-out and the log price of food-at-home, where the weights are the household's food budget shares of eating-out and of food-at-home. By employing each household's own within food relative consumption weights, this construction more accurately reflects the price for food faced by individual households than the total food index provided by the CPI.

5.3 Model Specification

We have proven identification of the model where all the component functions are nonparametric. However, these functions are high dimensional, so nonparametric estimation is not practical with modest sample sizes. We will therefore instead estimate the model parametrically, but make use of relatively flexible functional forms. Estimation will be based on moments implied by the model, and so will not entail specifying or estimating the distribution of error terms.

5.4 Budget Shares for Individuals

Our model starts with a utility derived functional form for the budget shares of individuals. We specify individual preferences using the Quadratic Almost Ideal Demand System (QUAIDS) developed by Banks et al. (1997).

Let p denote the J-vector of price indices of aggregate consumption goods. In our application, $J = 6$. Let y denote total expenditures. Let h index households, and let k denote a household member. The household member types k are f for female, m for male, and c for children. For member k of household h, let ω^{jhk} denote the fraction of member k's total resources in the household that he/she spends on good j, and let ω^{hk} be the J-vector of budget shares ω^{jhk} for $J = 1, ..., J$. Note that we can only observe ω^{jhk} in households h that have just one member k (since for those households observed purchased budget shares equal the shares consumed by member k).

The QUAIDS demand system, for a single individual of type k, living in the household h, takes the $J-$ vector form

$$\omega^{hk}\left(\frac{p}{y}\right) = \alpha^{hk} + \Gamma^k \text{ln}p + \beta^{hk}[\ln(y) - c^{hk}(p)] + \frac{\lambda^k}{b^{hk}(p)}[\ln(y) - c^{hk}(p)]^2. \tag{10}$$

Here $b^{hk}(p)$ and $c^{hk}(p)$ are price indices defined as

$$\ln[b^{hk}(p)] = (\text{ln}p)'\beta^{hk}, \tag{11}$$

$$c^{hk}(p) = c_0^{hk} + (\text{ln}p)'\alpha^{hk} + \frac{1}{2}(\text{ln}p)\Gamma^{k\prime}\text{ln}p, \tag{12}$$

α^{hk}, β^{hk}, and λ^k are J-vectors of preference parameters, Γ^k is a $J \times J$ matrix of preference parameters, and c_0^{hk} is a scalar parameter which we set to equal to zero based on the insensitivity reported in Banks et al. (1997). By definition, budget shares must add up to one, i.e., $\mathbf{1}'_J\omega^{hk} = 1$ for all p/y, where $\mathbf{1}_J$ is a J-vector of ones. This, in turn, implies that $\mathbf{1}'_J\alpha^{hk} = 1$, $\mathbf{1}'_J\beta^{hk} = 0$, $\mathbf{1}'_J\lambda^k = 0$, and $\Gamma^{k\prime}\mathbf{1}_J = \mathbf{0}_J$, where $\mathbf{0}_J$ is a J-vector of zeros. Slutsky symmetry requires that Γ^k be a symmetric matrix.

As the indices above show, we let the parameter vectors α^{hk} and β^{hk} vary by household h as well as by individual k. In particular, we specify these parameter vectors by

$$\alpha^{hk} = \alpha_0^k + \sum_{m=1}^{M_\alpha} \alpha_m^k d_{m,\alpha}^{hk} \tag{13}$$

$$\beta^{hk} = \beta_0^k + \sum_{m=1}^{M_\beta} \beta_m^k d_{m,\beta}^{hk}, \tag{14}$$

where $d_{m,\alpha}^{hk}$ and $d_{m,\beta}^{hk}$ are observed demographic characteristics, and M_α and M_β are the number of such covariates we observe. Each α^{hk} and β^{hk} is a J-vector, which from the above adding up restrictions must satisfy $\mathbf{1}'_J\alpha_0^k = 1$, $\mathbf{1}'_J\alpha_m^k = 0$ for $m = 1, ..., M_\alpha$, and $\mathbf{1}'_J\beta_m^k = 0$ for $m = 0, ..., M_\beta$.

In our application $d_{m,\alpha}^{hk}$ consists of 7 region dummies and the age of member k, making $M_\alpha = 8$, while $d_{m,\beta}^{hk}$ is an indicator for homeownership, so $M_\beta = 1$. Taken together, we have 17 preference parameters for each of $J - 1 = 5$ distinct equations, yielding a total of 85 parameters for each type of individual k. Note that the model for households with more than one member will also have additional parameters associated with resource shares and Barten scales.

5.5 The Estimator for Singles

The demand functions for households h consisting of just a single man or a single woman are given by equation (10). Such households have either $k = f$ if the household h is a single woman or $k = m$ if the household h is a single man (there are of course no single children households). In this subsection we describe how these demand functions for singles are estimated. The demand functions and associated estimators for households consisting of multiple members are given in the next subsection.

For households h consisting of singles, we append a J-vector valued additive error term U^{hk} (consisting of elements U^{jhk}) to equation (10). This introduces unobserved heterogeneity in the singles' demand functions. We assume that the error vectors U^{hk} are uncorrelated across households. Adding up requires $\mathbf{1}'_J U^{hk} = 0$, which implies that nonzero correlations must exist among the elements of each U^{hk}, that is, across goods j. We estimate the budget share demand equations for single men and for single women separately using GMM, allowing for arbitrary correlations in the errors across goods.

Let $u^{jhk}\left(\theta^k\right) = U^{jhk}$ denote ω^{jhk} minus the j'th element of the right hand side of equation (10), where θ^k is the vector of all the parameters in that equation. Note that $u^{jhk}\left(\theta^k\right)$ is implicitly a function of ω^{jhk} and of all the regressors in the model. The moments used for GMM estimation take the form $E\left(u^{jhk}\left(\theta^k\right)\tau^{hk}\right) = 0$, with τ^{hk} being a vector of covariates as defined below. To impose the adding-up constraints we apply the standard practice of dropping one demand equation, and we recover the estimated parameters for that last equation using the adding-up constraints. The choice of which demand equation to drop is numerically irrelevant, because by the adding-up constraints, the parameters of the dropped equation are all deterministic functions of the parameters in the remaining equations. The full set of moments for estimating the model of singles of type k is therefore $E\left(u^{jhk}\left(\theta^k\right)\tau^{hk}\right)=0$ for $j=1,...,J-1$. Letting $u^{hk}\left(\theta^k\right)$ be the $J-1$ vector of elements $u^{jhk}\left(\theta^k\right)$ for $j=1$,...,$J-1$, we equivalently write these moments as $E\left(\left(I_{J-1}\otimes\tau^{hk}\right)u^{hk}\left(\theta^k\right)\right)=0$.

The set of covariates τ^{hk} (for single households h) consists of region dummies, age, log relative prices, log real total expenditure (defined as the log of total expenditures divided by a Stone price index computed for our six nondurable goods) and its square, and the product of log real total expenditures with the home ownership dummy and with log prices. The number of moments therefore consists of $J-1=5$ distinct demand equations times the number of elements of τ^{hk}, which is 22, for a total of 110 moments for $k=f$ and for $k=m$.

We apply GMM for estimation separately for single women and for single men. For $k=f$ and for $k=m$, let H^k denote the set of households that consist of a single member of type k, and let n_k denote the number of elements of H^k. Denote the sample moments for GMM estimation by

$$v^k(\theta^k)=\frac{1}{n_k}\sum_{h\in H^k}\left(I_{J-1}\otimes\tau^{hk}\right)u^{hk}\left(\theta^k\right), \tag{15}$$

The GMM criterion function is then

$$\widehat{\theta}^k=\arg\min_{\theta^k}v^k(\theta^k)'W^kv^k(\theta^k) \tag{16}$$

where W^k is a weighting matrix. We apply standard two step GMM, where W^k is an estimate of the efficient GMM weighting matrix, given by

$$W^k=\left(\frac{1}{n_k}\sum_{h\in H^k}\left(I_{J-1}\otimes\tau^{hk}\right)u^{hk}\left(\widetilde{\theta}^k\right)u^{hk}\left(\widetilde{\theta}^k\right)'\left(I_{J-1}\otimes\tau^{hk}\right)\right)^{-1} \tag{17}$$

where $\widetilde{\theta}^k=\arg\min_{\theta^k}v^k(\theta^k)'v^k(\theta^k)$.

Although we do not use it for our main analysis, in addition to estimating the above model for single men and for single women, we for comparison also estimated if for other households (couples with 0-4 children). For multiple member households, this corresponds to what is known in the collective household literature as a unitary model, that is, a model that treats a household as if it was a single maximizing agent. We provide this unitary model just for comparison to singles, and to our later collective model estimates. Iillustrating the differences in demands of single women, single men, and other households, Figure B1 presents fitted Engel curve plots for our six goods, with total expenditures y ranging from the 1st to the 99th percentile. We shift the plots for couples with 0-4 children to the left in these figures to make them comparable to the singles plots. We find that food (at home and eating-out), utility, and communication are necessities while clothing and shoes, transportation, and entertainment are luxuries. Single women have a steeper Engel curve slope for clothing and shoes compared to other households. Couples with 0-4 children have a steeper Engel curve slope for entertainment compared to singles.

Elasticity estimates for single women and single men are reported in Table B1 in Appendix B.

5.6 The Joint Model

For our empirical application of the joint model, we assume that men, women, and children each have demands given by the QUAIDS functional form described in the previous section. The Barten type linear consumption technology for households with multiple members is

$$z_j = a_{sj} x_j \tag{18}$$

for each good j, or equivalently, $z = Ax$ with a diagonal matrix A. For households having total expenditure level y and facing market prices p, the resulting shadow prices for this technology are

$$\pi(p/y) = \frac{A'p}{y} \tag{19}$$

We parameterize each household member's resource shares with the functional form

$$\eta^f = \frac{\exp(\delta^{f\prime} s)}{1 + \exp(\delta^{f\prime} s) + \exp(\delta^{m\prime} s)}, \tag{20}$$

$$\eta^m = \frac{\exp(\delta^{m\prime} s)}{1 + \exp(\delta^{f\prime} s) + \exp(\delta^{m\prime} s)}, \tag{21}$$

where f denotes female and m denotes male, and the children's resource share is $1 - \eta^f - \eta^m$. If there are no children in the household, then

$$\eta^f = \frac{\exp(\delta^{f\prime} s)}{1 + \exp(\delta^{f\prime} s)}, \tag{22}$$

the husband's share is $1 - \eta^f$.

We allow the same cooperation factors s to affect the resource shares of every household member (wives, husbands, and children). The vectors of coefficients of s are δ^f and δ^m. The vector of cooperation factors s consists of the difference in age between the wife and husband, the difference in log income between the wife and husband, number of children, the minimum age of children less 5, the age of the wife less 39 (the average age of wives in the sample), and indicators of whether the wife has some college education, and whether the husband has some college education.

With the technology function (18), the corresponding shadow prices (19), and the sharing rule (20) and (21), we end up with the following expression for the budget shares of couples with one to four children:

$$\omega_j^h(p,s,y) = \eta_s^{hf}\omega_j^{hf}\left(\frac{\pi}{\eta_s^{hf}}\right) + \eta_s^{hm}\omega_j^{hm}\left(\frac{\pi}{\eta_s^{hm}}\right) + (1-\eta_s^{hf}-\eta_s^{hm})\omega_j^{hc}\left(\frac{\pi}{1-\eta_s^{hf}-\eta_s^{hm}}\right). \quad (23)$$

Couples with no children have the same expression but with $\omega^{h_j^c}$ (the budget share demand function of children c for good j) set equal to zero.

Equation (23) shows that the budget share of couples with zero to four children is equal to a weighted average of the budget share of its members (wives, husbands, and children), evaluated at shadow prices, with weights given by their respective resource share. The resource share η_s^{hk} represents both the fraction of the total expenditures controlled by member k and the extent to which the household's demand is represented by that member's preferences.

Unlike singles, who have budget share equations for six goods, couples have budget shares $\omega_j^h(p,s,y)$ for seven or eight goods, since they include men's clothes, women's clothes, and (when present) children's clothes as separate goods, while singles just consume one type of clothing.

The parameters of the joint model consists of all the QUAIDS parameters of budget shares, ω^{hf}, ω^{hm}, and ω^{hc}, the Barten scales A_j, and the parameters of the sharing rules η_s^{hf} and η_s^{hm}. We jointly estimate all the parameters of the model using data from both singles and couples.

We have 150 preference parameters (5 × 17 - 10 = 75 symmetry constrained QUAIDS parameters for each of men and women). We also have 6 Barten scale parameters and 16 sharing rule parameters (the 7 listed above plus the constant for each of men and women); this gives a total of 172 parameters. We have 335 instruments (for each of the 5 goods there are 22 instruments for single men, 22 for single women, and 23 for couples), giving a maximum degrees of freedom of 163 for the most general model. The GMM weighting matrices for singles, W^f and W^m, are obtained from the QUAIDS estimates for singles in the previous subsection; see equation (17). The weighting matrix for children, W^c is derived using two-step GMM on the full system, starting with an initial identity weighting matrix. The GMM criterion is:

$$\min_\theta (v^c(\theta)'W^c v^c(\theta) + v^f(\theta)'W^f v^f(\theta) + v^m(\theta)'W^m v^m(\theta)), \quad (24)$$

where θ is the full parameter vector of the joint model and the instrument matrices are defined as in equation (15).

5.7 Resource Shares and Barten Scales

The main results for our preferred model are displayed in Table 2. Panel A in Table 2 reports the estimates of sharing rule parameters. We find that wives' resource share decreases significantly with the number of children. In terms of percent change, as the number of children increases by 1, the wife's resource share decreases by 35.7%. In contrast, the number of children does not significantly affect the husband's resource share. Dunbar et al. (2013) similarly found that husband's resource shares were little affected by the number of children.

We find that education is a significant cooperation factor. Specifically, the resource share of wives who have some college education is 92.5% higher than those who do not. Even in families where husbands have some college education, wives enjoy a 52.4% higher resource share than in families where husbands do not have any college education.

The estimated resource shares for each type of household member (wives, husbands, and children) are reported in Table 3. The mean value of the wife's resource share is 0.51 in couples without children, 0.3 in couples with 1 child, 0.24 in couples with 2 children, and 0.17 in couples with 3 or 4 children. The mean value of the husband's resource share is 0.49 in couples without children, 0.24 in couples with 1 child, and stays almost constant as the number of children increases to 3 or 4. The results suggest that when there are no children present in the household, wives and husbands have similar resource shares or bargaining power. However, when the number of children rises, mothers on average devote much more of their own resource shares to children compared to fathers.

Table 2: *Estimation Results: the Sharing Rule and Barten Scales*

	Wife		Husband	
Panel A: the Sharing Rule	Coef	Std Error	Coef	Std Error
Constant	-0.701***	0.239	-0.913**	0.434
Difference in log income (female - male)	-0.069	0.057	-0.106	0.075
Difference in age (female - male)	-0.003	0.009	0.017	0.013
Number of children	-0.357***	0.058	-0.101	0.181
Minimum child age less 5	-0.375	0.256	-0.151	0.306
Female age less 39	0.223	0.152	0.409	0.244
Wife some college	0.925***	0.333	-0.869	0.793
Husband some college	0.524**	0.247	0.347	0.343
Panel B: Estimates of Barten Scales	Barten scale		Std Error	
Food	0.838***		0.017	
Clothing and shoes	1.000		-	
Communication	0.845***		0.020	
Entertainment	0.665***		0.015	
Transportation	0.760***		0.014	
Utility	0.562***		0.014	

Notes: Barten Scales are assumed to be homogeneous across all households. The Barten scale of clothing and shoes is assumed to be 1. *p < 0.10, **p<0.05, ***p<0.01.

Table 4 reports wives' resource share conditional on household characteristics. The benchmark household are ones in which neither the wife nor the husband has college education and are renters with median total expenditure. The first row shows that at our benchmark values, wives' resource share is 0.21. Rows 2 and 3 suggest that the education of both wives and husbands has a strong impact on wives' resource share. Wives in households who are home owners have slightly lower resource share. This is because home-owner households also tend to have children, and wives' resource share is lower in families with children.

For identification in their model, Dunbar et al. (2013) required that resource shares not depend on total expenditures. Our model does not require this restriction, and so can be used to test if it is valid. The last two rows of Table 4 show that resource shares do not change by total expenditure, providing empirical support to the assumption required by Dunbar et al. (2013). Our finding is also consistent with Menon, Pendakur, and Perali (2012) who find that the assumption also holds with data from the Italian International Center of Family Studies (CISF).

Estimates of Barten scales are reported in Panel B of Table 2. We restrict Barten scales to be between 0.5 and 1, as in Browning et. al. (2013). Because we assume clothing and shoes to be private assignable goods, we set their Barten scales equal to 1. We find that food and communication are highly private (having Barten scales close to one), while communication and utility are highly public (with Barten scales well below one). Transportation is found to be partially public. Note the transportation here includes both private cars and public transportation, where the former is more public and the latter is less public (here public means within households). The findings on Barten scales here are generally consistent with findings from previous literature, including Browning et al. (2013), Cherchye et al. (2012), Solvejg (2016), Lin (2018), and Fujii and Lin (2018).

Table 3: *Estimated Resource Shares*

		Zero child	One child	Two children	Three/four children	All households
Woman	Mean	0.51	0.30	0.24	0.17	0.28
	Std Dev	0.12	0.11	0.10	0.07	0.14
	Min	0.25	0.14	0.11	0.08	0.08
	Max	0.80	0.56	0.49	0.38	0.80
Man	Mean	0.49	0.24	0.25	0.27	0.28
	Std Dev	0.12	0.10	0.10	0.10	0.13
	Min	0.20	0.06	0.07	0.07	0.06
	Max	0.75	0.43	0.45	0.45	0.75
Children	Mean	-	0.45	0.50	0.56	0.43
	Std Dev	-	0.07	0.07	0.07	0.19
	Min	-	0.27	0.30	0.36	0.00
	Max	-	0.69	0.75	0.76	0.76

5.8 External Validation of Model Predictions

The estimated resource shares are unobserved, and may suffer from measurement error or estimation error due to possible model misspecification (an example is Calvi et al. 2019). To verify our results, we compare our estimated resource shares to individual private consumption given by the Japanese Panel Survey of Consumers (JPSC).

A unique feature of JPSC is that it asks the individual expenses and savings of each household member. Specifically, JPSC asks the following question (answered for both the wife and husband): How much expenditure, savings (including life insurance premiums etc.), and loan repayments

Table 4: *Sharing Rule Implications*

Household Characteristics	Wife's share η
	All households
Benchmark	0.21
Wife with some college education	0.45
Husband with some college education	0.32
Home owner	0.19
First quantile total expenditure	0.21
Third quantile total expenditure	0.21

did you pay this September? The answers are divided into : i) expenses/savings for all of my family ii) expenses/savings for me iii) expenses/savings for my husband iv) expenses/savings for my children v) expenses/savings for the others.

Categories ii), iii), and iv) are measures of private consumption for wives, husbands, and children. Category i) represents expenditures on goods that can be jointly consumed (like heat or gasoline). Previous studies have used inequality in private consumption to infer intra-household inequality in resource allocation.[9] However, these types of estimates, at best based on data like the JPSC, are incomplete, in the sense that they do not account for the potentially large role that shared goods may have in the actual resources consumed by each family member. In the JPSC data, over two thirds of expenditures are listed as shared goods. [10]

Comparing our results to the JPSC data, first consider children's shares. Our model predicts children's resource shares in the range of 0.45 to 0.56. This is consistent with the JPSC data, being above what JPSC reports for private children's consumption, and below the sum of JPSC's private children plus shared household expenditures. Second, our model estimates are that wives and husbands have roughly equal resource shares when there are no children present in the household. But the resource share of husbands increases up to around 1.6 times that of wives as the number of children in the household increases. This number is close to the ratio of private expenditures between husbands and wives, 1.5 - 2.3, found in the JPSC data. Taken together, these results provide evidence that our estimates are at least in ranges consistent with existing direct (albeit incomplete) measures of resource shares.

Finally we compare implications that one might draw about intra-household inequality from the JPSC data to estimates based on our model that accounts for and allocates expenditures on

[9] Lise and Yamada (2014) look at JPSC households and find that there is a substantial difference between private consumption devoted to the wife, 6.3 percent, versus the husband, 15 percent. On average, 21.3 percent of the household expenditures are reported as the private consumption of either the wife or the husband, leaving 78.7 percent of household expenditures as public (expenditures for the family, children, and others). Fujii and Lin (2018) look at JPSC couples without children and also find similar patterns. The average private consumption devoted to the wife is 10 percent, versus the husband, 15 percent. 68 percent of household expenditures are devoted to the family. The remaining 4 percent of household expenditures are devoted to others. The previous findings imply that if we only consider private expenditures, the husband's resources are about 1.5-2.3 times of the wife's. The public expenditures, including both children and family expenditures, are around 70-80 percent of total household expenditures.

[10] Note, however, that expenditures for the family in the JPSC data include some durables that our data excludes, like furniture and electronics spending.

shared goods. Our estimates are that increasing the number of children decreases the wive's share by 35 percent while the husband's share decreases by only 10 percent. In the JPSC data, these numbers are 47 percent and 15 percent (based on summary statistics reported in Table 1 of Fujii and Ishikawa 2013). By failing to allocate shared goods, the JPSC appears to underestimate the relative contribution of wives vs husbands to children's resources.

5.9 Indifference Scales and Economies of Scale

We next consider the private equivalent expenditures for household members in multi-person households, and the resulting household's economies of scale to consumption, and household member's indifference scales. The private good equivalent of good j by member k in household h, x_j^{hk}, is the quantity of good j that member k consumes, accounting for the extent to which that good is shared with other members. The more public a good is, and hence the more that good is shared, the lower is its Barten scale, and the greater is the sum of x_j^{hk} across household members k, relative to z_j, the household's purchased quantity of good j.

The household's economies of scale to consumption is how much more it would cost to buy every member's private good equivalents at market prices, relative to the household's actual total expenditure level. A member's indifference scale is defined to be the cost, at market prices, of the cheapest bundle of goods that gets member j to the same utility level (i.e., the same indifference curve over goods) that the member attains in the household by consuming his or her own vector of private good equivalents. See Browning et. al. (2013) for more details on these definitions.

Given our estimates of budget shares for singles, resource shares, and Barten scales, the private good equivalent quantities for each household member k for each good j are given by

$$x_j^{hk} = \frac{\eta_s^{hk}\omega_j^{hk}(\pi/\eta_s^{hk})}{\pi_j} = \frac{\omega_j^{hk}}{a_{sj}}\eta_s^{hk}y \tag{25}$$

The equivalent expenditures for each household member k are given by:

$$x^{hk} = \sum_j x_j^{hk} = \eta_s^{hk} y \sum_j \frac{\omega_j^{hk}}{a_{sj}} \tag{26}$$

The indifference scales for each are given by:

$$IS^{hk} = \frac{x^{hk}}{y} = \eta_s^{hk} \sum_j \frac{\omega_j^{hk}}{a_{sj}} \tag{27}$$

The relative economies of scale to consumption, R, are defined as

$$R = \frac{p'(\sum_k x^{hk})}{y} - 1 = \frac{p'(\sum_k x^{hk} - z)}{p'z} \tag{28}$$

Table 5 reports the estimates of members' private good equivalent expenditures x^k, indifference scales IS^k, and the overall economies of scale R. Row 6 in Table 5 reports the indifference scale for wives. This indifference scale can be interpreted as the fraction of the household's total expenditures that a wife would need when living alone (i.e., as a single) to attain the same indifference curve over goods that she reaches as a member of the household. The table shows that, on average, wives would require 67% of the couple's total expenditures to be as well off living alone as she is in the couple, when there are no children. This drops to only 23% in families with 3 to 4 children, reflecting how much less, relatively, women consume when children are present. The corresponding numbers for husbands (in row 7 of Table 5) are 66% without children, dropping to 36% when 3 to 4 children are present.

The interpretation of an indifference scale as the relative cost of living alone is not relevant for children, however, it still provides a measure of the savings that household's attain by sharing consumption, and we can still compare the relative values of children's indifference scales in households of different compositions. Children's indifference scales are reported in row 8 of Table 5.

The second to the last row in Table 5 gives household's overall economies of scale. On average, it ranges between 0.33 to 0.36 across different household types. This implies that it would cost the families 33% to 36% more to buy the (private equivalent) goods they consumed if there had been no shared or joint consumption.

Table 5: *Implications of Estimates*

	Couples with			
	0 child	1 child	2 children	3 - 4 children
Wife's resource share	0.51	0.30	0.24	0.17
Wife's equivalent expenditure	121.66	69.79	59.71	45.15
Husband's equivalent expenditure	119.80	56.47	62.58	69.09
Children's equivalent expenditure	-	107.47	126.26	148.15
Actual couple's expenditure	181.82	173.31	183.40	192.64
Indifference scale for women	0.67	0.40	0.32	0.23
Indifference scale for men	0.66	0.33	0.34	0.36
Indifference scale for children	-	0.62	0.69	0.77
Scale economy, R	0.33	0.35	0.35	0.36
Number of Observations	379	704	1369	392

Notes: values are in mean. Equivalent budget share is the budget share of the wife (husband) if she (he) is endowed with the fraction of resources and faced with shadow prices (market prices discounted by the Barten scales). The equivalent expenditure is the expenditure that the wife (husband) needs to obtain the same private good equivalents in marraige if she (he) is living alone, endowed with the fraction of resources in marriage and faced with market prices. Scale economy means it would cost the couple R percent more to buy the (private equivalent) goods they consumed if there had been no shared or joint consumption. The expenditures are in thousand yen.

6 Conclusions

We extend the existing theorem on collective household models by making the following contributions. First, we allow goods to be partly shared or jointly consumed. Second, we point identify all features of a collective household's optimization problem instead of weaker generic identification obtained in previous literature. Third, our model directly includes the utility of children rather than assuming it to be joint with the utility of an adult household member. Fourth, we reduce the data requirements relative to exiting theorems. Specifically, we do not require any child specific goods to identify the complete demand function of children and the resources devoted to them.

We apply our model to a Japanese data consisting of single men, single women, and married couples with zero to four children. We find that wives and husbands have similar control over resources when there are no children in the household. However, wives devote much more of their resources to children relative to husbands when there are children in the household. Around half of the household total expenditure is devoted to children. We find that it is important to allow goods to be partly shared because failure to account for that leads to underestimates of the decrease in wives' resources relative to husbands' when the number of children rises. In terms of individual welfare analysis, wives need two thirds of household total expenditure to live alone while being materially as well as living in the household. That number drops to only one fourth in households with 3-4 children. However, husbands still need one third of household total expenditure to live along while being materially as well as living in a 3-4 children household. One more child in a 1-2 children household will need an additional 7-8% of household total expenditure in order to maintain the current living standard of all children.

Our findings have important policy implications for the welfare analysis of children in multi-person households. For example, one potential application of our model is to calculate the insurance compensation for children when their parents pass away. Since we identify the preference of children, the framework can also be used to evaluate the impact of welfare programs (e.g., taxes or subsidies) that are targeted specifically at children on the individual welfare of mother, fathers, and children themselves. Lastly, we also call for future research in the development of alternative identifying assumptions and allowing for more general models of individual's preference change that results from the change of economic environments, e.g., marriage or household formation.

References

Banks, J., Blundell, R., and Lewbel, A. (1997). Quadratic Engel curves and consumer demand. The review of economics and statistics, 79(4), 527-539.

Bargain, O., and Donni, O. (2012). Expenditure on children: A Rothbarth-type method consistent with scale economies and parents' bargaining. European Economic Review, 56(4), 792-813.

Barten, A. P. (1964), "Family Composition, Prices, and Expenditure Patterns," in Hart, P., Mills, G. and Whitaker, J. K.(eds) Econometric Analysis for National Economic Planning: 16th Symposium Colston Society (London: Butterworth), 277–292.

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

Becker, G. S., (1981) A Treatise on the Family. Cambridge, MA: Harvard University Press.

Blundell, R., Chiappori, P. A., and Meghir, C. (2005). Collective labor supply with children. Journal of political Economy, 113(6), 1277-1306.

Browning, M., Bourguignon, F., Chiappori, P. A., and Lechene, V. (1994). Income and outcomes: A structural model of intrahousehold allocation. Journal of political Economy, 102(6), 1067-1096.

Browning, M., and Chiappori, P. A. (1998). Efficient intra-household allocations: A general characterization and empirical tests. Econometrica, 1241-1278.

Calvi, R. (2017). Why are older women missing in India? The age profile of bargaining power and poverty.

Chiappori, P. A., Fortin, B., and Lacroix, G. (2002). Marriage market, divorce legislation, and household labor supply. Journal of political Economy, 110(1), 37-72.

Browning, M., Chiappori, P. A., and Lewbel, A. (2013). Estimating consumption economies of scale, adult equivalence scales, and household bargaining power. Review of Economic Studies, 80(4), 1267-1303.

Calvi, R., Lewbel, A., and Tommasi, D. (2019). Women's Empowerment and Family Health: Estimating LATE with Mismeasured Treatment. Available at SSRN 2980250.

Chamberlain, G. (1986). Asymptotic efficiency in semi-parametric models with censoring. journal of Econometrics, 32(2), 189-218.

Cherchye, L., De Rock, B., and Vermeulen, F. (2012). Married with children: A collective labor supply model with detailed time use and intrahousehold expenditure information. The American Economic Review, 102(7), 3377-3405.

Cherchye, L., De Rock, B., and Vermeulen, F. (2012). Economic well-being and poverty among the elderly: an analysis based on a collective consumption model. European Economic Review, 56(6), 985-1000.

Cherchye, L., De Rock, B., Lewbel, A., and Vermeulen, F. (2015). Sharing rule identification for general collective consumption models. Econometrica, 83(5), 2001-2041.

Cherchye, L., Demuynck, T., De Rock, B., and Vermeulen, F. (2017). Household consumption when the marriage is stable. American Economic Review, 107(6), 1507-1534.

Chiappori, P. A. (1988). Rational household labor supply. Econometrica: Journal of the Econo-

metric Society, 63-90.
Chiappori, P. A. (1992). Collective labor supply and welfare. Journal of political Economy, 100(3), 437-467.
Chiappori, P. A., and Ekeland, I. (2006). The micro economics of group behavior: general characterization. Journal of Economic Theory, 130(1), 1-26.
Chiappori, P. A., and Ekeland, I. (2009). The Microeconomics of Efficient Group Behavior: Identification 1. Econometrica, 77(3), 763-799.
Couprie, H., Peluso, E., and Trannoy, A. (2010). Is power more evenly balanced in poor households?. Journal of Public Economics, 94(7-8), 493-507.
Dunbar, G. R., Lewbel, A., and Pendakur, K. (2013). Children's resources in collective households: identification, estimation, and an application to child poverty in Malawi. American Economic Review, 103(1), 438-71.
Fujii, T., and Lin, X. (2018). Individual Welfare Analysis in Japanese Couples without Children. Boston College Working Papers.
Heckman, J. (1990). Varieties of selection bias. The American Economic Review, 80(2), 313.
Lewbel, A. (1998). Semiparametric latent variable model estimation with endogenous or mismeasured regressors. Econometrica, 105-121.
Lewbel, A., and Pendakur, K. (2017). Unobserved preference heterogeneity in demand using generalized random coefficients. Journal of Political Economy, 125(4), 1100-1148.
Lewbel, A. (2018). Identification and estimation using heteroscedasticity without instruments: The binary endogenous regressor case. Economics Letters, 165, 10-12.
Lewbel, A., and Pendakur, K. (2019). Inefficient Collective Households: Abuse and Consumption.
Lin, X., (2018). SNAP and Food Consumption among Older Adults: A Collective Household Approach with Homescan Data. Boston College Working Papers.
Lindahl, E. (1958). Just taxation—a positive solution. In Classics in the theory of public finance (pp. 168-176). Palgrave Macmillan, London.
Lise, J., and Seitz, S. (2011). Consumption inequality and intra-household allocations. The Review of Economic Studies, 78(1), 328-355.
Lise, J., and Yamada, K. (2014). Household sharing and commitment: Evidence from panel data on individual expenditures and time use (No. W14/05). IFS Working Papers.
Matzkin, R. L. (2003). Nonparametric estimation of nonadditive random functions. Econometrica, 71(5), 1339-1375.
Matzkin, R. L. (2007). Nonparametric identification. Handbook of econometrics, 6, 5307-5368.
Matzkin, R. L. (2012). Identification in nonparametric limited dependent variable models with simultaneity and unobserved heterogeneity. Journal of Econometrics, 166(1), 106-115.
Menon, M., Pendakur, K., and Perali, F. (2012). On the expenditure-dependence of children's resource shares. Economics Letters, 117(3), 739-742.
Vermeulen, F. (2002). Collective household models: principles and main results. Journal of Economic Surveys, 16(4), 533-564.

Wewel, S. (2017). Heterogeneity in Consumption Gains from Living Together as a Couple. Boston College Working Papers.

7 APPENDIX A: Proofs

PROOF of LEMMA 1: The function $G(p)$ is identified for all $p \in \Omega_p$ by $G(p) = M(p,t)$, where t is defined in Assumption A2. Also, the functions $m_j(p,s)$ and $g_j(p)$ are identified (where the derivatives defining these functions exist) for all $p \in \Omega_p$ by construction because they are derivatives of identified functions.

Now let Assumption A3 hold. Since $m_j(p,s) = g_j(a_{s1}p_1, ...a_{sJ}p_J)$, we have that

$$\zeta_j(\alpha, \widetilde{p}, s) = a_{sj}\frac{g_j(a_{s1}\widetilde{p}_1, ...a_{sJ}\widetilde{p}_J)}{g_j(\alpha_1\widetilde{p}_1, ...\alpha_J\widetilde{p}_J)} \text{ for } j = 1, ..., J \tag{29}$$

Since this mapping is a contraction, by the Banach fixed point theorem there exists is a unique α such that $\alpha = \zeta(\alpha, \widetilde{p}, s)$. This unique α is identified, because the value of the function $\zeta(\alpha, \widetilde{p}, s)$ is identified. But by equation (29), $a_s = \zeta(a_s, \widetilde{p}, s)$, and therefore the unique identified α is the desired coefficient vector a_s.

Next, suppose instead that Assumption A4 holds. For all p in the neighborhood of zero given by Assumption A2, let $m_j(p,s) = \partial M(p,s)/\partial p_j$ and let $g_j(p) = \partial G(p)/\partial p_j$. These functions are identified by construction given that $M(p,s)$ and $G(p)$ are identified. Then, it follows from equation (29) that a_s is identified by $a_{sj} = \zeta_j(0,0,s) = \lim_{p\to 0} m(p,s)/g_j(p)$, noting that $m_j(p,s)$ and $g_j(p)$ now exist for p in a neighborhood of zero.

Finally, given identification of each a_s, the function $G(z)$ is identified not just for all $z \in \Omega_p$ but for all $z \in \Omega_z$ by $G(a_{s1}p_1, ...a_{sJ}p_J) = M(p,s)$ for all $(p,s) \in \Omega_p \times \Omega_s$.

PROOF of LEMMA 2:

First observe that, given Ω_p is the positive orthant and all a_{sj} are positive, it follows that Ω_z is also the positive orthant, and therefore $G(z)$ for all $z \in \Omega_z$ by $G(p) = M(p,0)$. It follows that c_j defined by equation (9) is also identified, since $G(p)$ is identified over the positive orthant and the function ψ_j is chosen. Next define constants C_{sj} by

$$C_{sj} = \int_0^\infty ... \int_0^\infty \psi_j[M(p,s)]\, p_1^{-1}...p_{j-1}^{-1}p_{j+1}^{-1}...p_J^{-1}dp_1...dp_J.$$

Each C_{sj} is identified, since $M(p,s)$ is identified for all p over the positive orthant and all $s \in \Omega_s$, and the function ψ_j is chosen. Notice that $c_j = C_{0j}$. Then, using the change of variables $\phi_i = a_{si}p_i$ for each good i,

$$\begin{aligned}
C_{sj} &= \int_0^\infty ... \int_0^\infty \psi_j[G(a_{s1}p_1, ...a_{sJ}p_J)]\, p_1^{-1}...p_{j-1}^{-1}p_{j+1}^{-1}...p_J^{-1}dp_1...dp_J \\
&= \int_0^\infty ... \int_0^\infty \psi[G(\phi_1, ...\phi_J)]\, \frac{a_{s1}}{\phi_1}...\frac{a_{s,j-1}}{\phi_{j-1}}\frac{a_{s,j+1}}{\phi_{j+1}}...\frac{a_{sJ}}{\phi_J}\frac{d\phi_1}{a_{s1}}...\frac{d\phi_J}{a_{sJ}} \\
&= \int_0^\infty ... \int_0^\infty \psi[G(\phi_1, ...\phi_J)]\, \frac{1}{\phi_1}...\frac{1}{\phi_{j-1}}\frac{1}{\phi_{j+1}}...\frac{1}{\phi_J}d\phi_1...d\phi_J\frac{1}{a_{sj}} = \frac{c_j}{a_{sj}}
\end{aligned}$$

so a_{sj} is identified for each $s \in \Omega_s$ and $j \in \{1, ..., J\}$ by $a_{sj} = c_j/C_{sj}$.

PROOF of THEOREM 1: This follows immediately from Lemmas 1 and 2, noting that without the normalization of Assumption A2, the coefficients a_{sj} in the proofs of Lemmas 1 and 2 correspond to a_{sj}/a_{tj} for some $t \in \Omega_s$ where the function $G(p)$ in these proofs corresponds to $G(a_{t1}p_1, ...a_{tJ}p_J)$

PROOF of THEOREM 2: If Assumption B2 holds then without loss of generality let $j = k$. Let $h_k^{k\prime}(p, y) = \partial h_k^k(p, y)/\partial y$. Then

$$\begin{aligned} M(p, s) &= \lim_{y \to 0} \frac{\partial \left[\eta_s^k(A_s p) h_k^k\left(A_s p, \eta_s^k(A_s p) y\right)\right]/\partial y}{\eta_s^k(A_s p)^2 h_k^k\left(A_s p, \eta_s^k(A_s p) y\right)^2} \\ &= \lim_{y \to 0} \frac{\eta_s^k(A_s p)\partial \left[\eta_s^k(A_s p) h_k^k\left(A_s p, \eta_s^k(A_s p) y\right)\right]/\partial \left[\eta_s^k(A_s p) y\right]}{\eta_s^k(A_s p)^2 h_k^k\left(A_s p, \eta_s^k(A_s p) y\right)^2} \\ &= \lim_{y \to 0} \frac{h_k^{k\prime}\left(A_s p, \eta_s^k(A_s p) y\right)}{h_k^k\left(A_s p, \eta_s^k(A_s p) y\right)^2} \\ &= \frac{h_k^{k\prime}(A_s p, 0)}{h_k^k(A_s p, 0)^2} = G(A_s p) \end{aligned}$$

where the last equality above defines the function G.

Alternatively, If Assumption B3 holds then again without loss of generality let $j = k$ and we have

$$M(p, s) = \int_0^\infty \left[\eta_s^k(A_s p)\right]^c \left[h_k^k\left(A_s p, \eta_s^k(A_s p) y\right)\right]^c y^{c-1} dy$$

Now do the change of variables $\tau = \eta_s^k(A_s p) y$

$$\begin{aligned} M(p, s) &= \int_0^\infty \left[\eta_s^k(A_s p)\right]^c \left[h_k^k(A_s p, \tau)\right]^c \left[\frac{\tau}{\eta_s^k(A_s p)}\right]^{c-1} \frac{d\tau}{\eta_s^k(A_s p)} \\ &= \int_0^\infty \left[h_k^k(A_s p, \tau)\right]^c \tau^c d\tau = G(A_s p) \end{aligned}$$

where the last equality above defines the function G.

Finally, if Assumption B4 holds then

$$\begin{aligned} M(p, s) &= \int_0^\infty \left(\sum_{k=1}^K \eta_s^k(A_s p) h_j^k\left(A_s p, \eta_s^k(A_s p) y\right)\right) dy \\ &= \sum_{k=1}^K \int_0^\infty \eta_s^k(A_s p) h_j^k\left(A_s p, \eta_s^k(A_s p) y\right) dy \end{aligned}$$

Now do the change of variables $\tau = \eta_s^k(A_s p) y$ in each of the K integrals above.

$$\begin{aligned} M(p, s) &= \sum_{k=1}^K \int_0^\infty \eta_s^k(A_s p) h_j^k(A_s p, \tau) \frac{d\tau}{\eta_s^k(A_s p)} \\ &= \sum_{k=1}^K \int_0^\infty h_j^k(A_s p, \tau) d\tau = G(A_s p) \end{aligned}$$

where the last equality above defines the function G.

PROOF of THEOREM 3:

By Corollary 1, the relative Barten technology parameters a_{sj}/a_{tj} and a_{rj}/a_{tj} are identified for given r, s, and t elements of Ω_s. Let A_{st} be the diagonal matrix that has elements a_{sj}/a_{tj} on the diagonal. Given Assumption C1, define the identified function m by $\eta_s^k(A_{st}p)h_k^k\left(A_{st}p, \eta_s^k(A_{st}p)y\right)$

$$m(p,s) = \omega_j(s, \phi_{st}(p), 0) = \eta_s^k(A_tp)h_j^k\left(A_tp, \eta_s^k(A_tp)0\right). \tag{30}$$

It then follows that relative values of resource shares are identified by

$$\frac{m(p,s)}{m(p,r)} = \frac{\eta_s^k(A_tp)h_j^k(A_tp,0)}{\eta_r^k(A_tp)h_j^k(A_tp,0)} = \frac{\eta_s^k(A_tp)}{\eta_r^k(A_tp)}.$$

Alternatively, given Assumption C2,

$$m(p,s) = \int_0^\infty \left[\eta_s^k(A_tp)\right]^{c_1}\left[h_j^k\left(A_tp, \eta_s^k(A_tp)y\right)\right]^{c_1} y^{c_2}dy$$

Now do the change of variables $\tau = \eta_s^k(A_tp)y$

$$\begin{aligned} m(p,s) &= \int_0^\infty \left[\eta_s^k(A_tp)\right]^{c_1}\left[h_j^k(A_tp,\tau)\right]^{c_1}\left[\frac{\tau}{\eta_s^k(A_tp)}\right]^{c_2}\frac{d\tau}{\eta_s^k(A_tp)} \\ &= \left[\eta_s^k(A_tp)\right]^{c_1-c_2-1}\int_0^\infty \left[h_j^k(A_tp,\tau)\right]^{c_1}\tau^{c_2}d\tau \end{aligned} \tag{31}$$

and it then follows that relative values of resource shares are identified by

$$\left[\frac{m(p,s)}{m(p,r)}\right]^{1/(c_1-c_2-1)} = \frac{\left[\left[\eta_s^k(A_tp)\right]^{c_1-c_2-1}\int_0^\infty\left[h_j^k(A_tp,\tau)\right]^{c_1}\tau^{c_2}d\tau\right]^{1/(c_1-c_2-1)}}{\left[\left[\eta_r^k(A_tp)\right]^{c_1-c_2-1}\int_0^\infty\left[h_j^k(A_tp,\tau)\right]^{c_1}\tau^{c_2}d\tau\right]^{1/(c_1-c_2-1)}} = \frac{\eta_s^k(A_tp)}{\eta_r^k(A_tp)}.$$

PROOF of THEOREM 4:

Let $k = 1$ be a household member that satisfies Assumption D1, and let the private assignable good for member $k = 1$ be the good $j = 1$. We have that each a_{sj}/a_{tj} and each $\eta_s^k(A_tp)/\eta_r^k(A_tp)$ is identified from Corollary 1 and Theorem 2. Let household type $t \in \Omega_s$ be a household that only contains member $k = 1$. Then for that t, $a_{tj} = 1$ for $j = 1, ..., J$ and A_t is the J by J, identity matrix. Substituting those values into a_{sj}/a_{tj} and $\eta_s^k(A_tp)/\eta_r^k(A_tp)$ shows that each a_{sj} and each $\eta_s^k(p)/\eta_r^k(p)$ is identified. In addition, If Assumption C1 holds then for each assignable k, equation (30) simplifies to $m(p,s) = \eta_s^k(p)h_k^k(p,0)$, which we can solve for, and thereby identify, $\eta_s^k(p)$. Alternatively, if Assumption C2 holds then equation (31) simplifies to

$$m(p,s) = \left[\eta_s^k(p)\right]^{c_1-c_2-1}\int_0^\infty \left[h_k^k(p,\tau)\right]^{c_1}\tau^{c_2}d\tau$$

which again we can solve for, and thereby identify, $\eta_s^k(p)$. Finally, if we have identified $\eta_s^k(p)$ for $k = 1, ..., K-1$, then we can identify $\eta_s^K(p)$ by $\eta_s^K(p) = 1 - \sum_{k=1}^{K-1} \eta_s^k(p)$.

8 APPENDIX B: Figures and Tables

Figure B1: *QUAIDS fits for singles and couples with 0-4 children*

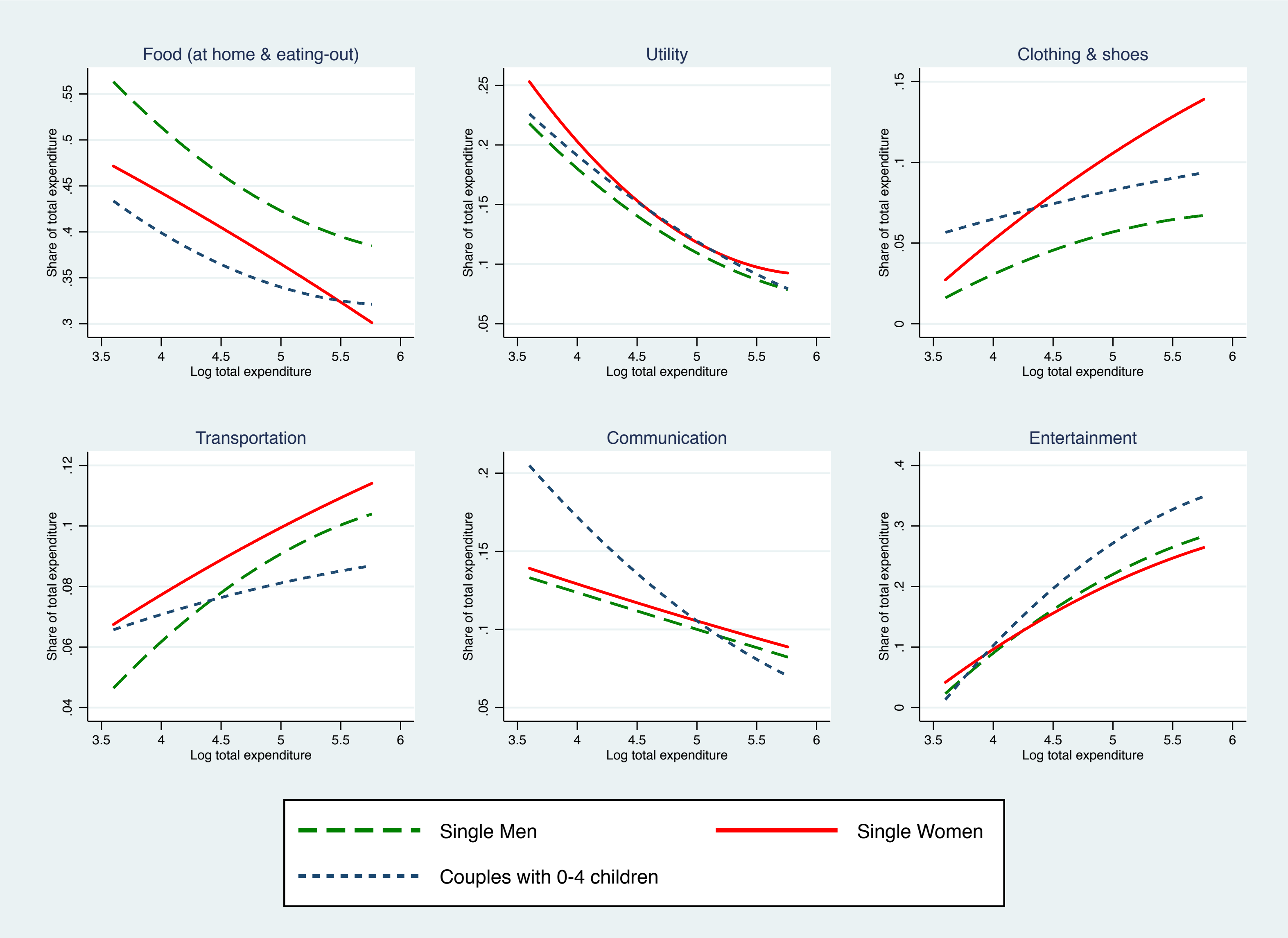

Table B1: *Elasticities Estimates of Single Men and Women*

Budget Elasticities

	Single women	Single men
Food	0.74	0.81
Clothing	1.45	1.20
Communication	0.78	0.76
Entertainment	1.45	1.53
Transportation	1.13	1.24
Utility	0.54	0.43

Uncompensated Price Elasticities (single women)

	Food	Clothing	Communication	Entertainment	Transportation	Utility
Food	-1.01	0.21	-0.59	0.95	-0.04	-0.23
Clothing	0.71	-1.69	0.91	-6.26	4.36	0.83
Communication	-2.57	0.97	-0.37	1.61	0.56	-1.05
Entertainment	2.77	-4.01	0.98	-3.07	-0.27	3.45
Transportation	-0.30	5.48	0.77	-0.53	-5.67	4.08
Utility	-1.37	1.03	-0.92	4.51	2.25	-5.14

Compensated Price Elasticities/Slutsky Matrix (single women)

	Food	Clothing	Communication	Entertainment	Transportation	Utility
Food	-0.72	0.29	-0.51	1.08	0.03	-0.13
Clothing	1.35	-1.44	1.11	-5.92	4.56	1.06
Communication	-2.33	1.05	-0.26	1.74	0.64	-0.96
Entertainment	3.50	-3.78	1.21	-2.69	-0.06	3.72
Transportation	0.18	5.63	0.91	-0.28	-5.48	4.25
Utility	-1.25	1.07	-0.89	4.58	2.29	-5.05

Uncompensated Price Elasticities (single men)

	Food	Clothing	Communication	Entertainment	Transportation	Utility
Food	-1.29	-0.31	-0.42	1.60	0.18	-0.53
Clothing	-2.44	-0.42	-0.10	1.21	-0.21	0.13
Communication	-1.85	-0.25	-1.67	2.89	-0.02	0.46
Entertainment	3.69	-1.92	1.56	-4.30	-0.04	1.18
Transportation	0.99	5.38	-0.05	-0.21	-3.82	-1.27
Utility	-2.11	0.60	0.16	-1.27	2.67	-0.45

Compensated Price Elasticities/Slutsky Matrix (single men)

	Food	Clothing	Communication	Entertainment	Transportation	Utility
Food	-0.93	-0.25	-0.34	1.76	0.26	-0.44
Clothing	-1.87	-0.28	0.05	1.51	-0.07	0.29
Communication	-1.59	-0.19	-1.57	3.03	0.05	0.53
Entertainment	4.50	-1.76	1.78	-3.89	0.15	1.41
Transportation	1.60	5.49	0.11	0.11	-3.63	-1.10
Utility	-2.01	0.62	0.18	-1.22	2.69	-0.39