form $M(p, s) = G(a_{s1}p_1, \ldots, a_{sJ}p_J)$ where the function M is known or identified, e.g., M could be a conditional mean function estimated by nonparametric regression. We wish to point identify, up to normalization, the vector of coefficients $a_s = (a_{s1}, \ldots, a_{sJ})$ for each value that s can take on.

Note that this is not a single linear index model. Many results exist for identifying coefficients in linear index models, i.e., models that are functions of $a_1p_1 + \ldots + a_jp_j$. But those results are not applicable to this context. Here each p_j appears separately, and generally nonlinearly, in the function G. Still, as we show below, multiple (rather than single) linear index models do form a special case of the models we consider, so our results add to the existing literature on identification of multiple linear index models.

We provide three different assumptions that suffice to point identify the coefficients a_{sj} for $j = 1, \ldots, J$. Each assumption has different strengths and weaknesses, so different ones will be more or less useful depending on context. An attractive feature of these identification results is that they do not impose any monotonicity on the function G.

We then extend these results to show point identification of a general set of collective household consumption models. There is a long literature on the identification and estimation of collective household models of consumption. These are models of 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. Objects of particular interest are resource shares, defined as the fractions of household resources spent on each family member. Virtually all of the identification results in this collective household model literature either point identify specific functional forms, or point identify only a subset of the model's features, or only establishes either set or generic identification rather than point identification.

Generic identification of a model means that the model is usually point identified, but there can exist situations where point identification fails. More formally, generic identification says that in the set of all possible data generating processes that satisfy the model's assumptions, the subset for which point identification fails has measure zero. See McManus (1992) and Lewbel (2019) for more details regarding the formal definition of generic identification.

The well known collective household identification results of Chiappori and Ekeland (2006, 2009) and earlier authors, showing nonparametric identification up to unknown levels for resource shares, are generic identification theorems. As a result, there exist functional forms where point identification fails. For example, their model is not nonparametrically point identified if household members have Cobb-Douglas preferences. Moreover, as is typical for generic identification results,

equation (1) holds where each a_{sj} captures the non-neutral technical efficiency of input j, and/or the quality of input j, by a firm with characteristics s. More generally, the p in $M(p;s)$ could be a vector of both input and output prices in a multiple output production process, with signs of a_{sj} determining which elements are inputs and which are outputs in a firm with characteristics s. A large literature exists on modeling heterogeneity in non-neutral efficiency in both macro, as in Basu and Fernald (1997), and industrial organization, as in Ackerberg, Caves, and Frazer (2015), and Gandhi, Navarro, and Rivers (2020).

3. Multiple linear index models. These can be constructed as a special case of our model. Suppose we add the constraint that all a_{sj} and p_j are strictly positive (this constraint will apply in our empirical application). Then we can equivalently write equation (1) as $M(p;s) = \widetilde{G}(\ln a_{s1} + \ln p_1; \ldots; \ln a_{sJ} + \ln p_J)$. Since s has finite support we can next equivalently replace each $\ln a_{sj}$ with a saturated model $\ln a_j(s) = \beta_j' S$ where S is a vector of binary variables indicating each possible value in the support of s. We then get $M(p;s) = \widetilde{G}\left(\beta_1' S + \ln p_1; \ldots; \beta_j' S + \ln p_J\right)$, which is a multiple linear index structure. Multiple linear index models are popular structures in statistics and econometrics, with estimators including Ichimura and Lee (1991), Horowitz (1998), Xia, Tong, Li, and Zhu (2002), Xia (2008), Donkers and Schafgans (2008), and Ahn, Ichimura, Powell, and Ruud (2018). The restriction that each linear index has one explanatory variable that appears only in that index, with a coefficient of one (corresponding to the $\ln p_j$ terms in $\widetilde{G}$) appears as Assumption 3a in Donkers and Schafgans (2008). They observe this is one way to satisfy some necessary conditions for identification that appeared previously in the literature. Note that, in addition to the constraint that all a_{sj} and p_j in our model be strictly positive, the multiple linear index literature mostly focuses on applications where regressors are continuous, rather than our opposite extreme where only p is continuous.

4. Collective Household Models. The modern literature on Pareto efficient collective household models begins with Becker (1965, 1981) and Chiappori (1988, 1992). An important series of papers in this literature establishes that, from only observing the demand functions of households, one cannot point identify resource shares (a resource share is the fraction of a household's total resources that are spent on the utility of any one household member). However, one can generically identify the marginal effects of policy variables on resource shares. Equivalently, each resource share is only point identified up to an unknown location constant. See, e.g., Browning, Bourguignon, Chiappori, and Lechene (1994), Browning and Chiappori (1998), Vermeulen (2002), and Chiappori

and Ekeland (2006, 2009)[2]. Prominent papers that make use of these identi cation results include Chiappori, Fortin, and Lacroix (2002), and Blundell, Chiappori, and Meghir (2005).

By adding additional assumptions, more recent papers either generically identify the entire model, including the levels of resource shares, e.g., Browning, Chiappori, and Lewbel (2013), or point identify some features of the model (such as resource shares without price e ects), e.g., Lewbel and Pendakur (2008), Bargain and Donni (2012), Dunbar, Lewbel, and Pendakur (2013), and Penglase (2019). Still other papers impose additional parametric restrictions to obtain point identi cation, e.g., Couprie, Peluso, and Trannoy (2010) and Lise and Seitz (2011).

None of the above results accomplish our goal, which is to provide su cient conditions to semiparametrically point identify (not just generically identify) an entire collective household model, including resource share levels and price e ects.

One large, general class of collective household models in the literature is based on Browning, Chiappori, and Lewbel (2013), which we will hereafter refer to as BCL. All but a handful of the papers cited above can be cast as special cases of BCL. BCL yields demands that can be written as a system of equations, each having a form resembling[3]

$$M(p; s; y) = G(A_s p; \;_s(A_s p)\, y) \qquad (2)$$

where M is quantity demand, G and $_s$ are unknown functions, p is a vector of observed prices, y is observed total expenditures, $_s$ is a resource share function, and the a_{sj} terms on the diagonal of A_s

3 Semiparametric Coefficient Identification

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 (possibly also including some mass points) and let S be a discrete covariate (or vector of covariates). 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. The goal is to identify the unknown vector of coefficients $a_s = (a_{s1}, ..., a_{sJ})$ in the model

$$M(p, s) = G(a_{s1}p_1, ..., a_{sJ}p_J) = G(A_s p) \tag{3}$$

for some unknown function G.

In this section we provide three alternative sets of conditions, each of which suffice for point identification of the vector of coefficients $(a_{s1}, ..., a_{sJ})$ for each value s that S can equal. Each has relative advantages and disadvantages. None, however, require monotonicity of the function G. The following two assumptions are common to all three sets of assumptions.

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 (3) 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. Assumption A2 can be made without loss of generality (as long as a_{tj} is not identically zero), because we can simply 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 $\widetilde{a}_{sj} = a_{sj}/a_{tj}$. Note, however, that the choice of normalization can affect economic interpretation of the function G and the a_{sj} coefficients.[4]

Our first alternative identifying assumption is the following

ASSUMPTION A3: Assume $G(p)$ is continuously differentiable. Let $m_j(p, s) = \partial M(p, s)/\partial p_j$

[4] In our collective household application, the a_{sj} coefficients are measures of how much each good j is shared (consumed jointly by multiple members) in a household of type s. There it will be appropriate to normalize a_{tj} to equal one for singles t (people who live alone), and who therefore cannot be sharing. See Lewbel (2019) for more on the economic implications of scale normalizations.

and let $g_j(p) = \partial G(p)/\partial p_j$. For any J-vector $= (\ _1, \ldots, \ _J)$, define the J-vector valued function $(\ ; p; s)$ as having the elements

$$_j(\ ; p; s) = \frac{m_j(p; s)}{g_j(\ _1 p_1, \ldots, \ _J p_J)} \text{ for } j = 1, \ldots, J$$

For each $s \in \ _s$, assume there exists $a \in \ _p$ such that $A_s p \in \ _p$ and $_j(\ ; p; s)$ is a contraction on a.

Assumption A3, is a high level assumption, which may therefore be hard to verify in practice. However, in the special case of multiple linear index models, Assumption A3 corresponds to uniquely recovering index coefficients from derivatives of M, and so relates to the identification conditions given in Xia (2008) and Donkers and Schafgans (2008).

An alternative to Assumption A3 is Assumption A4, which is more restrictive than A3, but is a much lower level assumption and hence may be simpler to verify in some applications.

ASSUMPTION A4: Assume $_p$ includes a (possibly one sided) neighborhood of zero, and that $G(p)$ is continuously differentiable for all p in that neighborhood of zero. Assume for each $j = 1, \ldots, J$ that $\partial G(p)/\partial p_j$ (or the corresponding one sided derivatives) does not equal zero when $p = 0$.

Assumption A4 exploits how our model simplifies at the point where $p = 0$. This is a method of identification that is also used by Matzkin (2003, 2012) and Lewbel and Pendakur (2017). Applying Assumption A4 when p is prices requires the one sided version of Assumption A4, since prices cannot be negative. In practice, this identification would require some probability of observing arbitrarily low prices (so the support of p contains values in the neighborhood of zero). However, both ordinary consumer demand models and collective household models are linearly homogeneous in prices p and total expenditures y. Therefore, it is only p/y that needs to include a one sided neighborhood of zero, and the presence of very wealthy consumers can insure that some observed values of p/y are very close to zero.

Define the random vector V by $V = (V_1, \ldots, V_J)$ where $V_j = a_{sj} P_j$. Let $_v$ 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 $a_{s1}, \ldots, a_{sJ}$ and the function $G(v)$ are point identified for all $v \in \ _v$ and $s \in \ _s$.

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. See Khan and Tamer (2010) and Lewbel (2019) for details regarding thin set identification.

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 $_p$. 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 a third alternative assumption for obtaining identification. A disadvantage of this identification condition is that it requires a large support assumption on P. However, unlike identification at infinity or other thin set identification arguments, 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 $_j$, define c_j by

$$c_j = \int_0^{\infty} \cdots \int_0^{\infty} {}_j\left[G(p)\right] p_1^{1} \cdots p_{j-1}^{1} p_j p_{j+1}^{1} \cdots p_J^{1} dp_1 \cdots dp_J \tag{4}$$

ASSUMPTION A5: Assume $_p$ is the positive orthant. $G(p)$ is continuous for all $p \in$. All a_{sj} are positive. For each $j \in \{1, \ldots, J\}$, we can find a continuous function $_j$ such that the constant c_j defined by equation (4) exists, is finite, and non-zero.

Having $_p$ be the positive orthant is the large support assumption. As noted above for Assumption A4, when p is prices we can replace p with p/y, so very low and very high incomes (corresponf147re5T2.99toisofon support also requires that extremes in relative prices of goods be possible.

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 $M(p;t)$ (recalling that by Assumption A2, all a_{tj} equal one). In our empirical application, the a_{sj} coefficients will be sharing parameters that are positive by construction.

Assumption A5 says we can find a continuous function $_j$ that makes the integral given by equation (4) convergent. Note that $G(p)$ is identified by $G(p) = M(p;t)$, so knowing G, the however, that the rate of convergence of the resulting estimator may depend on which identifying assumptions hold.

4 The Collective Household Model of Consumption

We brie y summarize Pareto e cient collective household consumption models here, focusing on the BCL model. Until recently, virtually all collective household models divided goods into two types:

jointness of consumption. For each good j, the household sets x_j/z_j equal to $1/a_{sj}$. Having $a_{sj} = 1$ means good j is not jointly consumed at all (this would be the case if all goods were private, or if the individual lived alone), otherwise the smaller a_{sj} is, the more good j is consumed jointly.

BCL show[5] that the household's demand functions arising from the above optimization have the form

$$\frac{p_j z_j}{y} = \omega_j(p; s; y) = \sum_{k=1}^{K} e_s^k(p; y) h_j^k\left(a_{s1} p_1; \ldots; a_{sJ} p_J; e_s^k(p; y) y\right) \qquad j = 1; \ldots; J \tag{6}$$

The function $\omega_j(p; s; y)$ is the household's budget share demand function for good j. h_j^k is household member k's demand function for good j, based on member j's utility function. $e_s^k(p; y)$ is member k's resource share, that is, the fraction of the household's total budget y

Vermeulen (2015) (see also Bonke and Browning 2011).

With these assumptions, we can write the resulting BCL demand functions as

$$\frac{p_j z_j}{y} = \omega_j(p;s;y) = \sum_{k=1}^{K} \eta_s^k(A_s p) h_j^k\left(A_s p; \eta_s^k(A_s p) y\right). \tag{7}$$

where resource shares now have the simpler form $\eta_s^k(A_s p)$. For each member k who has a private assignable good, we will index that good as good k. The household demand functions of the private assignable good simplify to

$$\frac{p_k z_k}{y} = \omega_k(p;s;y) = \eta_s^k(A_s p) h_k^k\left(A_s p; \eta_s^k(A_s p) y\right) \tag{8}$$

It will be important for some later results to note that utility maximization results in demand functions that are homogeneous of degree zero in p and y (this is known as the absence of money illusion), which means that equation (7) can be equivalently written as

$$\frac{p_j z_j}{y} = \omega_j(p;s;y) = \sum_{k=1}^{K} \eta_s^k\left(A_s \frac{p}{y}\right) h_j^k\left(A_s \frac{p}{y}; \eta_s^k\left(A_s \frac{p}{y}\right)\right). \tag{9}$$

and similarly for equation (8).

5 Identification of the Collective Household Model

We now consider identification of the collective household model given by equations (7) and, for private assignable goods, (8). As with Theorem 1, we present a few alternative sets of identifying assumptions each with relative advantages and disadvantages depending on context.

ASSUMPTION B1: Household budget share demand functions $\omega_j(p;s;y)$ for $j = 1;\ldots;J$ are given by equation (7), which for private assignable goods reduces to equation (8), 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 positive and continuous for each member $k = 1;\ldots;K$, and each $s \in \Omega_s$. The consumption technology constants a_{sj} are bounded and strictly positive for each $s \in \Omega_s$ and each 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 the Barten coefficients a_{sj} be bounded and positive must hold because it is impossible for every member of a household to consume more than the total purchased quantity of a good (even if it is completely shared), and it is impossible to consume negative quantities of goods.

Our first goal is to identify relative values of the Barten constants $a_{s1},...,a_{sJ}$. We cannot immeand has nonzero derivatives with respect to **p**:

$$M(p,s) = \lim_{y \to 0} \frac{\partial \omega_j(p,s,y)/\partial y}{\omega_j(p,s,y)^2} - \frac{\partial \omega_{\bar{j}}(p,s,y)/\partial y}{\omega_{\bar{j}}(p,s,y)^2}$$

ASSUMPTION B4: Assume that Ω_y includes $(0,\infty)$. Assume there exists a private assignable good j. Assume that for all $(p,s) \in \Omega_p \times \Omega_s$ (except possibly on a subset of measure zero), there exists a real constant c such that the function $M(p,s)$ defined by the following equation is finite, and has nonzero derivatives with respect to **p** in Theorem 2 we do not assume a scale normalization, i.e., we do not yet impose Assumption A2. Later we will use data on singles living alone, who therefore cannot share, to properly scale each a_{sj}.

A notable feature of Theorem 2 is that it gets identification from the demand functions of just one or two goods 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. Also, 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.

Another feature of Theorem 2 is that the only constraint it places on the resource share functions $\eta_s^k(p)$ is the minimal regularity given in Assumption B1. In particular, Assumptions B2 to B6 place no additional constraints on the resource share functions, as can be seen by replacing each $\eta_s^k(p)$ with any other suitably bounded regular function $\tilde{\eta}_s^k(p)$ in the proof of Theorem 2.

To illustrate some of the above alternative identifying assumptions, consider the general case of private assignable demand functions that are polynomials in y. More precisely, let good k be assignable to member j, and assume the function h_j^k is an arbitrary L'th order polynomial in y, so

$$\omega_j(p, s, y) = \eta_s^j(A_s p) \sum_{\ell=0}^{L} a_{\ell j}(A_s p)\, \eta_s^j(A_s p)^{\ell} y^{\ell}$$

for some functions

Given identi cation of the Barten technology, our next goal is identi cation of the relative values of the resource share functions$^{k}_{s}$. De ne the vector $_{st}(p)$ to be the vector of elements $_{stj}(p_j)$ de ned by

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

for some t 2 $_{s}$ chosen by the econometrician.

ASSUMPTION C1: Assume that $_{y}$ includes a one sided neighborhood of zero, that there exists a private good j that is assignable to some household member k, and for that good j the budget share function $!_j(s; {}_{st}(p); 0)$ is nite and nonzero for all $(p; s)$ 2 $_{p}$ $_{s}$.

ASSUMPTION C2: Assume that $_{y}$ includes (0;1), that there exists a private good j that is assignable to some household member k, and for that good $p; s$ $_{p}$ $_{s}$

in Theorem 3 uses just the demand functions of at most two goods 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 using multiple goods that the household consumes. Another limitation of Theorem 3 relative to the earlier generic identification literature (albeit a restriction with a great deal of theoretical precedent and empirical support, as discussed earlier) is our assumed restriction that the resource share function not depend on y.

Identification of relative values of resource shares does not suffice to answer some questions of economic significance. In particular, 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.

ASSUMPTION D1: For each household member $k = 1, \ldots, K - 1$ assume there exists a private assignable good, which without loss of generality denote as good k. Assume that we observe singles of member type k living alone, and that the demand functions for these assignable goods, the functions h_k^k, are the same whether a member of type k is in a collective household or not.

To identify the levels of resource shares, BCL assume that we can observe singles of every household member type $k = 1, \ldots, K$, and that their demand functions for all goods remain the same whether inside or outside a collective household. Assumption D1 considerably weakens the BCL assumptions, by only requiring that we observe singles of $K - 1$ member types, and that only one good for each type needs to have a demand function that doesn't change when in a collective household.[7] However, Assumption D1 is stronger than BCL in one sense, which is that it requires existence of some private assignable goods.

THEOREM 4: Let the Assumptions of Corollary 1 hold for all $s \in \Omega_s$, and let Assumption D1 hold. Let either Assumption C1, C2, or B6 hold. Then the entire model is identified.

What we mean by the entire model being identified in Theorem 4 is that all the Barten scales

[7]Note that when we say the demand function doesn't change, we only mean the functions h_k^k (which are derived from individual k's utility function) stay the same. Actual consumption quantities as functions of prices and total expenditures will differ, because within the collective household, each function h_k^k is evaluated at shadow prices and a shadow budget, rather than market prices and the single's actual budget.

a_{sj}, all the resource share functions $\eta_s^k(p)$, and all the demand functions $h_j^k(p; y)$ are identified.

In our application, we have $K = 3$: men, women, and children. So for Theorem 4 we need $K - 1 = 2$ of these three to have an identifiable private assignable good. In our case we observe men's clothes and women's clothes, which have demand functions that we identify from single men

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

$$\ln[b^{hk}(p)] = (\ln p)' \;\; {}^{hk}; \qquad (11)$$

$$c^{hk}(p) = c_0^{hk} + (\ln p)' \;\; {}^{hk} + \frac{1}{2}(\ln p) \;\; {}^{k\prime} \ln p; \qquad (12)$$

^{hk}, ^{hk}, and ^{k} are J-vectors of preference parameters, ^{k} is a J J matrix of preference parameters ${}^{k}_{jj'}$ having rank J 1, 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 de nition, budget shares must add up to one, i.e., $1_J' ! ^{hk} = 1$ for all $p=y$, where 1_J is a J-vector of ones. This, in turn, implies that $1_J' \; {}^{hk} = 1$, $1_J' \; {}^{hk} = 0$, $1_J' \; {}^{k} = 0$, and ${}^{k\prime} 1_J = 0_J$, where 0_J is a J-vector of zeros. Slutsky symmetry requires that ^{k} be a symmetric matrix.

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

$${}^{hk} = {}^{k}_{0} + \sum_{m=1}^{M} {}^{k}_{m} d^{hk}_{m;} \qquad (13)$$

$${}^{hk} = {}^{k}_{0} + \sum_{m=1}^{M} {}^{k}_{m} d^{hk}_{m;} \; ; \qquad (14)$$

where $d^{hk}_{m;}$ and $d^{hk}_{m;}$ are observed demographic characteristics, and M and M M and

$$\eta^m = \frac{\exp(\gamma^{m\prime}s)}{1+\exp(\gamma^{f\prime}s)+\exp(\gamma^{m\prime}s)}; \tag{16}$$

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

$$\eta^f = \frac{\exp(\gamma^{f\prime}s)}{1+\exp(\gamma^{f\prime}s)}; \tag{17}$$

and the husband's share is $1-\eta^f$. This is a commonly used functional form for imposing the constraint that resource shares are positive and sum to one.

In the collective household literature, variables s that affect resource shares are called "distribution factors." See, e.g., Browning, Bourguignon, Chiappori, and Lechene (1994), Browning and Chiappori (1998). In our model, these s variables also affect the Barten parameters a_{sj}. Lewbel and Pendakur (2019) call variables that affect both resource shares and sharing, "cooperation factors." The vector s in our application consists of the difference in age between the wife and husband, the difference in log income between the wife and husband,[8] 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 Barten consumption technology, we obtain the following expression for the budget shares of couples with one to four children:

$$\omega^h_j(p;s^h;y^h) = \eta^{hf}_s\,\omega^{hf}_j\left(\frac{A_s p}{\eta^{hf}_s y}\right) + \eta^{hm}_s\,\omega^{hm}_j\left(\frac{A_s p}{\eta^{hm}_s y}\right) + (1-\eta^{hf}_s-\eta^{hm}_s)\,\omega^{hc}_j\left(\frac{A_s p}{(1-\eta^{hf}_s-\eta^{hm}_s)\,y}\right). \tag{18}$$

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

We next require one of Assumptions B2, B3, B4, B5, or B6 to hold. In our demand model, Assumption B6 holds with L=2. Alternatively, it can be directly verified that Assumption B3 holds as well. Either suffices for Theorem 2.

Next consider Corollary 1. This entails identification of relative values of A_s from $M(p;s)$. This is most readily satisfied with $M(p;s) = c^{hk}(A_s p)$. Applying Assumption A3 we get that $\partial M(p;s) = \partial p_j$ at $p = 1$ is $\sum_{j'=1}^{J} {}^{k}_{jj'} a_{sj}$ for $j = 1; \ldots; J$, from which we can recover the relative values of the a_{sj}. Note that the matrix of parameters ${}^{k}_{jj'}$ is identified from variation in p.

Finally, consider Theorems 3 and 4. Assumption C1 is in some ways a mild restriction, since it only requires that budget shares, which should lie between zero and one, stay well bahaved even when y goes to zero. However, some popular functional forms, including our QUAIDS model, violate this assumption, because it's a polynomial in $\ln y$. The demand functions here do not satisfy either Assumption C1 or C2, and so Theorem 3 identifying relative resource shares does not apply. However, in this case we do not need Theorem 3, because we satisfy the assumptions of Theorem 4, 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) single women and men are restricted to be between 22 to 65 years old; (2) couples with children aged 15 or over are excluded (since adult clothing purchases could be consumed by older children); (3) households with members as students are excluded; (4) for married couples, households with members over 50 are excluded; (5) observations where expenditures on four or more of the six goods is zero are excluded; and (6) to mitigate the possible effects of outliers, we trim the 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 276 single women, 357 single men, and 1068 married couples having from zero to four children.

Price data comes from the 2015 based Consumer Price Index (CPI) from e-sTat, the portal site of official statistics of Japan. The detailed construction of price indexes for each aggregate good is reported in Appendix B of the Supplemental Appendix.

7.2 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. 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).[10] We assume that U^{hk} are uncorrelated across households. Adding up requires $1_J' U^{hk} = 0$, which implies that nonzero correlations must exist among the elements of each U^{hk}, that is, within households across goods j. Budget share demand equations are estimated using GMM, allowing for arbitrary correlations in the errors across goods.

[10] Additive errors can either be rationalized as measurement errors in budget shares, or by imposing restrictions on preference heterogeneity as in Lewbel (2001).

Let $u^{jhk}\ {}^{k} = U^{jhk}$ denote $!^{jhk}$ minus the j'th element of the right hand side of equation (10), where ^{k} is the vector of all the parameters in that equation. Note that $u^{jhk}\ {}^{k}$ is implicitly a function of $!^{jhk}$ and of all the regressors in the model. The moments used for GMM estimation take the form $E\ \ u^{jhk}\ {}^{k}\ {}^{hk} = 0$, with ^{hk} being a vector of covariates as de ned 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 $E\ \ u^{jhk}\ {}^{k}\ {}^{hk} = 0$ for $j = 1; :::; J\ \ 1$. Letting $u^{hk}\ {}^{k}$ be the $J\ \ 1$ vector of elements $u^{jhk}\ {}^{k}$ for $j = 1,...,J\ \ 1$, we equivalently write these moments as $E\ \ I_{J\ 1}\ \ {}^{hk}\ u^{hk}\ {}^{k} = 0$.

The set of covariates ^{hk} (for single households h) consists of region dummies, age, log relative prices, log real total expenditure (de ned 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

thereforey2s1 11.358 9.09rrT54.33684 T4rshipE

99th percentile. We shift the plots for couples with 0-4 children to the left in these gures to make them comparable to the singles plots. We nd 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 1 in Appendix D of the Supplemental Appendix.

7.3 The Joint Model

Unlike singles, who have budget share equations for six goods, couples have budget shares $h^h_j(p, s^h; y^h)$ 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 consist of all the QUAIDS parameters of budget shares ω^{hf}, ω^{hm}, and ω^{hc}, the Barten scales A_s, and the parameters of the sharing rules $^{hf}_s$ and $^{hm}_s$. We jointly estimate all the parameters of the model using data from both singles and couples.

We have 150 preference parameters (517 - 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. 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:

$$\min(v^c(\ )'W^c v^c(\ ) + v^f(\ )'W^f v^f(\ ) + v^m(\ )'W^m v^m(\ )); \qquad (20)$$

where is the full parameter vector of the joint model and the instrument matrices are de ned as in equation (19).

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	1,180	822	375	706	1,364	395
Number of unique households	357	276	192	282	456	138
Household income (thousand yen)	3460.66	.	7520.87	5910.11	6408.41	6400.97
Total expenditures (January, thousand yen)	121.20	113.33	181.33	175.14	191.00	202.10
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	0.08	0.07
Budget share (communication)	0.11	0.11	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	-	-	0.05	0.02	0.02	0.01
Children clothing&shoes share	-	-	0.00	0.03	0.04	0.04
Female age	-	47.09	38.29	37.79	38.37	38.26
Female unemployed	-	0.11	0.10	0.23	0.22	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	-	39.20	39.12	39.89	39.29
Male unemployed	0.07	-	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	-	-	-	6.80	9.72	11.41
Child 2 age	-	-	-	-	6.50	8.68
Child 3 age	-	-	-	-	-	5.34
Child 4 age	-	-	-	-	-	-
Child average age	-	-	-	6.79	8.11	8.33
Home ownership						

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

	Wife	Husband

Table 4: Sharing Rule Implications

Household Characteristics	Wife's resource share
	All households
Benchmark	0.21
Wife with some college education	0.45
Husband with some college education	0.32
Home owner	0.19

Notes: The benchmark households (row 1) are ones in which neither the wife nor the husband has college education and are renters with median total expenditure. Row 2 shows the wife's resource share in households that are similar to the benchmark households but in which the wife has college education. Row 3 shows the wife's resource share in households that are similar to the benchmark households but in which the husband has college education. Row 4 shows the wife's resource share in households that are similar to the benchmark households but are home owners.

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
Indi erence scale for women	0.67	0.40	0.32	0.23
Indi erence scale for men	0.66	0.33	0.34	0.36
Indi erence 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.

(which constitutes over three fourths of all household consumption in the JPSC).

As another check on our estimates, we do our own comparison (in Appendix C of the Supplemental Appendix) to self-reports of individual private consumption in the JPSC. Overall, our estimates are comparable to the JPSC reports, however, by failing to allocate shared goods, we find that the JPSC appears to underestimate the relative contribution of wives vs. husbands to children's resources.

Estimates of Barten scales are reported in Panel B of Table 2. Clothing and shoes are our private assignable goods, so their Barten scales equal one. We find that food and communication are highly the private good equivalent quantities for each household member k for each good j are given by

$$x_j^{hk} = \frac{\omega_j^{hk}}{a_{sj}} \eta_s^{hk} y^h \tag{21}$$

and relative economies of scale to consumption R are defined as

$$R = \frac{\sum_j p_j \sum_k x_j^{hk}}{y^h} - 1 = \frac{\sum_j p_j \left(\left(\sum_k x_j^{hk}\right) - z_j^h\right)}{\sum_j p_j z_j^h}. \tag{22}$$

BCL define a member's indifference scale to be the cost (as a fraction of y), at market prices, of the cheapest bundle of goods that gets member k 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. Let $\tilde{V}^k$ denote the QUAIDS indirect utility function of member k. The indifference scale IS^{hk} for each member k is defined as the solution to

$$\tilde{V}^k\left(\frac{p/y}{IS^{hk}}\right) = \tilde{V}^k\left(\frac{A_s p/y}{\eta_s^{hk}}\right). \tag{23}$$

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, indifference scales for children still provide a measure of the savings in costs of children that households attain by sharing consumption, and it is meaningful to 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 families 33% to 36% more to buy the (private equivalent) goods they consumed if there had been no shared or joint consumption.

8 Conclusions

We provide theorems for point identifying a general class of semiparametric models that are applicable to a variety of applications, including continuous consumer demand, production functions, and multiple index models. We then extend these results to show point identification for a large class of collective household models, which previously had only been shown to be generically identified. Moreover, we do so in a model that allows goods to be partly shared, including identifying the demand functions and resource shares of children.

We apply our model to Japanese data consisting of single men, single women, and married couples with zero to four children. Our findings have important policy implications for the analysis of individual welfare, particularly children's welfare, in multi-person households. For example, one potential application of our identification and resulting estimates could be to calculate appropriate levels of compensation for children, to maintain their standard of living, if parents separate or a parent dies. Also, since we identify (ordinally) the utility functions of children and their parents, the framework can be used to evaluate the impact of welfare programs (e.g., taxes or subsidies) on the individual welfare of mothers, fathers, and children.

References

Ackerberg, D.A., K. Caves, and G. Frazer (2015), Identification Properties of Recent Production Function Estimators. Econometrica, 83, 2411-2451

Ahn, H., H. Ichimura, J. L. Powell, and P. A. Ruud (2018) Simple Estimators for Invertible Index Models, Journal of Business & Economic Statistics, 36(1), 1-10.

Ai, C. and Chen, X. (2003), Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions. Econometrica, 71: 1795-1843.

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.

Basu, S. and J. G. Fernald, 1997, Returns to Scale in U.S. Production: Estimates and Implications Journal of Political Economy 105(2), 249-283.

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.

Bonke, J. and M. Browning, (2011), "Spending on Children: Direct Survey Evidence," The Economic Journal, 121(554), F123{F143.

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). E cient intra-household allocations: A general characterization and empirical tests. Econometrica, 1241-1278.

Browning, M., Chiappori, P. A., and Lewbel, A. Philde01(har-327.01ntiou21.00nd)mpzationntiou21.7(ec

d(el,51001.el,51003(Chiapp)-27.001(ori,51081.998(P)81.004(.47425.00A.ri,51081.990(and)7425.00Megh

andcPorutcwning, Journal ofolitical Economy

Lewbel,51827.001.el,51928.000(and)-380(P)82.00ommasiel,51827.00D.nd PPc and Famild

withBaredof Frenoionm

Camb(1996)G44.99[(with)4126.90cens1(ong.nd)-38j(Journa.)]TJ T* 7(of)-307(Econometrs),)-323(1)105(2),

ke,M.,e,and35827.00V(P)82.69ernoucad(e,)6-27.00F(G.)5780.99(Pw)277001(e35827.00labpp)326.993r8.

withand36348.992(in)27.001(th(Jou)0.99s(.u)0.99er)80.99sehold

Cherchye, L., De Rock, B., Lewbel, A., and Vermeulen, F. (2015). Sharing rule identi cation 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 Econometric 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 E cient Group Behavior: Identi cation 1. Econometrica, 77(3), 763-799.

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.

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.

Donkers, B., and M. Schafgans, (2008). Speci cation and estimation of semiparametric multiple-index models. Econometric Theory, 24(6), 6heorywheuji(2008)icedr(6),)9-5676.005-.004(ran)1.19005-.00

Horowitz, J. (1998) Semiparametric Methods in Econometrics, NewYork:Springer

Ichimura, H., and Lee, L. F. (1991), \Semiparametric Least Squares Estimation of Multiple Index Models: Single Equation Estimation," in Nonparametric and Semiparametric Methods in Econometrics and Statistics: Proceedings ofthe Fifth International Symposium in Economic Theory and Econometrics,eds. W. A. Barnett, J. Powell, and G. Tauchen, Cambridge, U.K.: Cambridge University Press

Jorgenson, D., L.J. Lau, and T.M. Stoker (1982) \The Transcendental Logarithmic Model of Aggregate Consumer Behavior." In Advances in Econometrics , edited by RL Basmann and G Rhodes, 97-238. Greenwich: JAI Press.

Khan, S. and Tamer, E. (2010), "Irregular Identi cation, Support Conditions, and Inverse Weight Estimation," Econometrica, 78, 2021-2042.

Lewbel, A. (1991), "The Rank of Demand Systems: Theory and Nonparametric Estimation," Econometrica, 59(3) 711-730.

Lewbel, A. (1998), "Semiparametric latent variable model estimation with endogenous or mismeasured regressors," Econometrica, 66, 105-121.

Lewbel A. (2001), "Demand Systems with and without Errors." American Economic Review, 91(3), 611-618.

Lewbel, A. (2019), "The Identi cation Zoo: Meanings of Identi cation in dels: SingleA.267.001(G)
I

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 identi cation. Handbook of econometrics, 6, 5307-5368.

Matzkin, R. L. (2012). Identi cation in nonparametric limited dependent variable models with simultaneity and unobserved heterogeneity. Journal of Econometrics, 166(1), 106-115.

McManus, D. A. (1992), \How common is identi cation in parametric models?" Journal of Econometrics, 53(1), 5-23.

Menon, M., Pendakur, K., and Perali, F. (2012). On the expenditure-dependence of children's resource shares. Economics Letters, 117(3), 739-742.

Penglase, J. (2019), "Consumption Inequality Among Children:Evidence from Child Fostering in Malawi," Unpublished manuscript, Boston College.

Pollak, R., and Wales, T. (1981). Demographic Variables in Demand Analysis. Econometrica, 49(6), 1533-1551.

Prais, Sigbert J., and Hendrik S. Houtakker, 1955, The analysis of household budgets,Cambridge.

A Appendix

A.1 Appendix A: Proofs

PROOF of LEMMA 1: The function $G(p)$ is identified for all $p \in {}_p$ 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 $p \in {}_p$ 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

$$_j(\;;\;p;s) = a_{sj}\frac{g_j(a_{s1}p_1;:::a_{sJ}p_J)}{g_j(\;_1p_1;:::\;_Jp_J)} \text{ for } j = 1;:::;J \tag{24}$$

Since this mapping is a contraction, by the Banach fixed point theorem there exists is a unique such that $\;=\;(\;;\;p;s)$. This unique is identified, because the value of the function $(\;;\;p;s)$ is identified. But by equation (24), $a_s = \;(a_s;p;s)$, and therefore the unique identified 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 $m_j(p;s) = @M(p;s)=@_jp$ and let $g_j(p) = @G(p)=@_jp$ (these can be one sided derivatives). These functions are identified by construction given that $M(p;s)$ and $G(p)$ are identified. Then, it follows from equation (24) that a_s is identified by $a_{sj} = \;_j(0;0;s) = \lim_{p\,!\,0} m(p;s)=g(p)$ (where, e.g., this limit is from above if $p > 0$).

Finally, given identification of each a_s, the function $G(m$ G

for each good j,

$$
\begin{aligned}
C_{sj} &= \int_0^\infty \cdots \int_0^\infty {}_j[G(a_{s1}p_1,\ldots,a_{sJ}p_J)]\, p_1^{-1}\cdots p_{j-1}^{-1}p_{j+1}^{-1}\cdots p_J^{-1}\, dp_1\cdots dp_J \\
&= \int_0^\infty \cdots \int_0^\infty [G({}_1,\ldots {}_J)]\, \frac{a_{s1}}{{}_1}\cdots\frac{a_{s,j-1}}{{}_{j-1}}\frac{a_{s,j+1}}{{}_{j+1}}\cdots\frac{a_{sJ}}{{}_J}\,\frac{d{}_1}{a_{s1}}\cdots\frac{d{}_J}{a_{sJ}} \\
&= \int_0^\infty \cdots \int_0^\infty [G({}_1,\ldots {}_J)]\, \frac{1}{{}_1}\cdots\frac{1}{{}_{j-1}}\frac{1}{{}_{j+1}}\cdots\frac{1}{{}_J}\, d{}_1\cdots d{}_J\,\frac{1}{a_{sj}} = \frac{c_j}{a_{sj}}
\end{aligned}
$$

so a_{sj} is identified for each $s \in {}_s$ and $j \in \{1,\ldots,J\}$ 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 $t \in {}_s$ where the function $G(p)$ in these proofs corresponds to $G(a_{t1}p_1,\ldots,a_{tJ}p_J)$

j = k and we have

$$M\ (p;s) = \int_0^1 \ {}^{k}_{s}(A_s p)\ ^{c}\ h^k_k\ \ A_s p;\ {}^{k}_{s}(A_s p)y\ \ ^{c} y^{c\ 1} dy$$

Now do the change of variables $= \ {}^{k}_{s}(A_s p)y$

$$M\ (p;s) = \int_0^1 \ {}^{k}_{s}(A_s p)\ ^{c}\ h^k_k\,(A_s p;\ )\ ^{c}\ \ \frac{}{{}^{k}_{s}(A_s p)}^{c\ 1}\ \frac{d}{{}^{k}_{s}(A_s p)}$$
$$= \int_0^1 h^k_k\,(A_s p;\ )\ ^{c}\ {}^{c}d\ \ = G\,(A_s p)$$

where the last equality above de nes the function G.

Now, if Assumption B5 holds then

$$M\ (p;s) = \int_0^1 \ \sum_{k=1}^{K} {}^{k}_{s}(A_s p) h^k_j\ \ A_s p;\ {}^{k}_{s}(A_s p)y\ \ !\ dy$$
$$= \sum_{k=1}^{K} \int_0^1 {}^{k}_{s}(A_s p) h^k_j\ \ A_s p;\ {}^{k}_{s}(A_s p)y\ \ dy$$

Next do the change of variables $= \ {}^{k}_{s}(A_s p)y$ in each of the K integrals above.

$$M\ (p;s) = \sum_{k=1}^{K} \int_0^1 {}^{k}_{s}(A_s p) h^k_j\,(A_s p;\ )\ \frac{d}{{}^{k}_{s}(A_s p)}$$
$$= \sum_{k=1}^{K} \int_0^1 h^k_j\,(A_s p;\ )\,d\ \ = G\,(A_s p)$$

where the last equality above de nes the function G.

Finally, consider the case where B6 holds. If $h^k_j\,(p;y) = \sum_{`=0}^{L} \ {}^{k}_{j`}\,(p)\,(\ln y)^{`}$ for the private assignable k = j, then

$$!_j\ (p;s;y) = \ {}^{j}_{s}(A_s p)\ \sum_{`=0}^{L} \ {}^{k}_{j`}\ (A_s p)\ \ \ln y + \ln\ \ {}^{k}_{s}(A_s p)\ \ ^{`}$$

Therefore $(p;s) = (\ {}^{j}_{s}(A_s p) j\ \ {}^{k}_{jL}\ (A_s p)\, j)\ ^{1}$ (using the fact that resource shares are positive), so with $M\ (p;s) = \ \ (p;s)\,!_j\ (p;s;\ \ (p;s))$ we get

$$M\ (p;s) = \frac{\sum_{`=0}^{L} \ {}^{k}_{j`}\ (A_s p)\ \ \ \ \ln j\ \ {}^{k}_{jL}\ (A_s p)\, j\ \ ^{`}}{j\ \ {}^{k}_{jL}\ (A_s p)\, j}$$

which is a function of just terms of the form $^{k}_{j}(A_s p)$, and so defines 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, s, and we also identify all relative resource shares ${}^k_s(A_r p) = {}^k_t(A_r p)$ by Theorem 3. For singles of type $k = 1, \ldots, K - 1$, resource shares ${}^k_t$ must equal one, so taking $r = t$ we identify ${}^k_s(p) = {}^k_t(p) = {}^k_s(p)$.

Alternatively, if Assumption B6 holds, then $h^k_j(p; y) = \sum_{\ell=0}^{L} {}^k_{j\ell}(p)(\ln(y))^\ell$, so the ${}^k_{k\ell}$ functions are known. This, along with all a_{sj} being known for $k = 1, \ldots, K - 1$ means that resource shares ${}^k_s$ can be recovered from equation (8).

Finally resource shares sum to one, so given the resource share functions ${}^k_s$ for all household types s and members $k = 1, \ldots, K - 1$, we identify the resource share functions for the last household type K by ${}^K_s(p) = 1 - \sum_{k=1}^{K-1} {}^k_s(p)$.

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.[2] 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 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 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 re ects the price for food faced by individual households than the total food index provided by the CPI.

0.2 APPENDIX C: External Validation of Model Predictions

The estimated resource shares are unobserved, and may su er from measurement error or estimation error due to possible model misspeci cation (see, e.g. Calvi et al. 2019). To verify our results, we compare our estimated resource shares to individual priv996(vHated)-408(u996(results,)-343)-408(67.99TD |

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.

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.

Table 1: Elastcites Estmates 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

Figure 1: QUAIDS Fits for Singles and Couples with 0-4 Children