Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power

Martin Browning, Pierre-André Chiappori and Arthur Lewbel

Oxford University, Columbia University, and Boston College

First version Feb. 2002, final version May 2013

Abstract

How much income would a woman living alone require to attain the same standard of living that she would have if she were married? What percentage of a married couple's expenditures are controlled by the husband? How much money does a couple save on consumption goods by living together versus living apart? We propose and estimate a collective model of household behavior that permits identification and estimation of concepts such as these. We model households in terms of the utility functions of its members, a bargaining or social welfare function, and a consumption technology function. We demonstrate generic nonparametric identification of the model, and hence of a version of adult equivalence scales that we call "indifference scales," as well as consumption economies of scale, the household's resource sharing rule or members' bargaining power, and other related concepts.

1 Introduction

On average, how much income would a woman living alone require to attain the same standard of living that she would have if she were married? What percentage of a married couple's expenditures benefit the husband? How much money does a couple save on consumption

goods by living together versus living apart? The goal of this paper is to propose a collective model of household behavior aimed at answering questions such as these.

A large literature exists on specification and estimation of ordinary demand systems, collective and household bargaining models, and on identification (and lack thereof) of equivalence scales, bargaining power measures, resource shares, consumption economies scale, and other related household welfare measures. Some surveys of this literature include Deaton and Muellbauer (1980), Blundell (1988), Browning (1992), Pollak and Wales (1992), Blundell, Preston, and Walker (1994), Blackorby and Donaldson (1994), Bourguignon and Chiappori (1994), Lewbel (1997), Jorgenson (1997), Slesnick (1998), and Vermeulen (2000). Results from a working paper version of the present paper are discussed in Lewbel (2003a).

We propose a new model of household consumption behavior that has three components, which are separate utility functions over goods for each household member, a consumption technology function that characterizes the jointness or publicness of goods and hence the economies of scale and scope in consumption, and a sharing rule that defines the relative allocation of household resources among the household members. The basic structure of this model is that households purchase a bundle (an n vector of quantities) of goods z. By economies of scale and scope in consumption, that is, by sharing, this z for the household is equivalent to purchasing a bundle of privately consumed goods x where each element of x is typically greater than or equal to the corresponding element of z. The vector of quantities x is then divided up among the household members, and each member i derives utility from consuming their bundle of goods x^i, so the sum of these x^i vectors across household members i is x. The conversion of z to x, which is embodied by the consumption technology function, is essentially an application of the models of Barten (1964) and Gorman (1976) that characterize how demands differ across households of different sizes. The allocation of shares of x to different household members, characterized by a sharing rule, is essentially the collective household model of Chiappori (1988, 1992), Bourguignon and Chiappori (1994) and Vermeulen (2000). Our model combines the features of both approaches, which is what allows us to identify consumption economies of scale, adult equivalence scales, and household bargaining power.

We provide a dual representation of our collective model that facilitates empirical application, and show that the model is generically nonparametrically identified. More precisely, we show that, given the demand functions of an individual man, individual woman, and of the couple living together, all of the underlying structural components of the couple's optimization program can be recovered. In particular, we can identify the level of the sharing rule, even without distributon factors (in contrast to earlier work that could only identify the derivative of the sharing rule with respect to distribution factors). We can also identify the consumption technology function, which summarizes the extent to which each good is shared, in contrast to previous work that required goods to be purely private or purely public.

We then show how the model also overcomes traditional problems regarding nonidentification of equivalence scales, and can be used to address the questions listed above. Our results only require ordinal representations of preferences, and do not depend on any utility cardinalization or interpersonal comparability assumptions. We apply the model to Canadian consumption data on couples and singles. Our identification theorems do not directly address

issues of unobserved preference heterogeneity, so our empirical analysis is based on estimates of mean demand functions, conditioning on observable demographic characteristics.

2   Equivalence Scales and Indifference Scales

Equivalence scales seek to answer the question, "how much money does a household need to spend to be as well off as a single person living alone?" Equivalence scales have many practical applications. They are commonly used for generating poverty lines for households of various compositions given a poverty line for single males. Income inequality measures have been applied to equivalence scaled income rather than observed income to adjust for household composition (see, e.g., Jorgenson 1997). Calculation of appropriate levels of alimony or life insurance also entail comparisons of costs of living for couples versus those of singles.

An equivalence scale is traditionally defined as the expenditures of the household divided by the expenditures of a single person that enjoys the same "standard of living" as the household. Just as a true cost of living price index measures the ratio of costs of attaining the same utility level under different price regimes, equivalence scales are supposed to measure the ratio of costs of attaining the same utility level under different household compositions. Unfortunately, unlike true cost of living indexes, equivalence scales defined in this way can never be identified from revealed preference data (that is, from the observed expenditures of households under different price and income regimes). The reason is that defining a household to have the same utility level as a single individual requires that the utility functions of the household and of the single individual be comparable. We cannot avoid this problem by defining the household and the single to be equally well off when they attain the same indifference curve, analogous to the construction of true cost of living indices, because the household and the single have different preferences and hence do not possess the same indifference curves. Pollak and Wales (1979, 1992) describe these identification problems in detail, while Blundell and Lewbel (1991) prove that only changes in traditional equivalence scales, but not their levels, can be identified by revealed preference. See Lewbel (1997) for a survey of equivalence scale identification issues.

We argue that the source of these identification problems is that the standard equivalence scale question is badly posed, for two reasons. First, by definition any comparison between the preferences of two distinct decision units entails interpersonal utility comparisons. Second, and perhaps more fundamental, the notion of a household utility is flawed. Individuals have utility, not households. What is relevant is not the 'preferences' of a given household, but rather the preferences of the individuals that compose it.

We propose therefore that meaningful comparisons must be undertaken at the individual level, and that the appropriate question to ask is, "how much income would an individual living alone need to attain the same indifference curve over goods that the individual attains as a member of the household?" This latter question avoids issues of interpersonal comparability and does not require us to compare the utility levels of different indifference curves. This question also does not depend on the utility level that is assigned to an indifference curve, i.e., it is unaffected by the fact that a person's utility associated with a particular indifference curve over goods might change as a result of living with a partner. The question only depends

3

on ordinal preferences, and hence is at least in principle answerable from revealed prefer-ence data. Consequently, in sharp contrast with the existing equivalence scale literature, our framework does not assume the existence of a unique household utility function, nor does it require comparability of utility between individuals and collectives (such as the household). Instead, following the basic ideas of the collective approach to household behavior, we as-sume that each individual is characterized by his/her own ordinal utility function, so the only comparisons we make is between the same person's welfare (defined by indifference curves) in different living arrangements.

Define a *equivalent income (or expenditure)* to be the income or total expenditure level y^{i*} required by an individual household member i 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. In our model, this means that when the household spends y to buy a bundle z, household member i consumes a bundle x^i (determined by the consumption technology and sharing rule). Then y^{i*} is the least expenditure required to buy a bundle of goods that lies on member i's same indifference curve as the bundle x^i. Then, instead of a traditional adult equivalence scale, we define individual i's "*indifference scale*" to be $S^i = y^{i*}/y$.

If member i were given the fraction S^i of the household's total expenditures, then by buy-ing goods on the open market individual i could get herself to the same indifference curve (defined in terms of her own utility function) that she attained as a member of the household, taking into account whatever economies of scale in consumption she enjoyed by sharing and joint consumption within the household. Indifference scales depend only on the indifference curves of household members, the resources of the household, and on the degree to which consumption is shared within the household, and so can be identified without any utility car-dinalization or interpersonal comparability assumptions.

To see the usefulness of indifference scales, consider the question of determining an ap-propriate level of life insurance for a spouse. If the couple spends y dollars per year then for the wife to maintain the same standard of living after the husband dies, she will need an insur-ance policy that pays enough to permit spending $S^f y$ dollars per year. Note that this amount of money only compensates the wife enough to reach the same indifference curve over goods that she attained while she was a household member. It does not compensate for any loss of utility due to grief, or for any change in her preferences that might result from the death of her husband. Similarly, in cases of wrongful death, juries are instructed to assess damages both to compensate for the loss in "standard of living," (i.e., S^i) and, separately for "pain and suffer-ing," which would presumably be noneconomic effects (see Lewbel 2003b). Another example is poverty lines. If poverty lines for individuals have been established, then the poverty line for a couple could be defined as the expenditures required for *each member* of the couple to attain his or her own poverty line indifference curve.

Traditional equivalence scales do not properly answer these questions, because they at-tempt to relate the utility of an individual to that of a household, instead of relating the utility of the same individual in two different settings, e.g. living with a husband versus without. This is similar to the distinction Pollak and Wales (1992) make between what they call a welfare comparison versus a situation comparison.

4

3 The Model

This section describes the proposed household model. Let superscripts refer to household members and subscripts refer to goods. Let $U^i(x^i)$ be the direct utility function for a consumer i, consuming the vector of goods $x^i = (x_1^i, ..., x_n^i)$. We consider households consisting of two members, which we will for convenience refer to as the husband ($i = m$) and the wife ($i = f$). For many applications, it may be useful to interpret one of these utility functions as a joint utility function for all but one member of the household, e.g., U^f could be the joint utility function of a wife and her children.

3.1 Household Members

ASSUMPTION A1: Each household member i has a monotonically increasing, continuously twice differentiable and strictly quasi-concave utility function $U^i(x^i)$ over a bundle of n goods x^i.

If member i were to face an n vector of prices p with income level y^i and this utility function, he or she would solve the optimization program

$$\max_{x^i} U^i(x^i) \text{ subject to } p'x^i = y^i. \tag{1}$$

Let $x^i = h^i(p/y^i)$ denote the solution to this individual optimization program, so the vector valued function h^i is the set of Marshallian demand functions corresponding to $U^i(x^i)$. Define V^i by

$$V^i\left(p/y^i\right) = U^i\left[h^i\left(p/y^i\right)\right] \tag{2}$$

so V^i is the indirect utility function corresponding to U^i. The functional form of individual demand functions h^i could be obtained from a functional specification of V^i using Roy's identity.

3.2 The Household Decision Process

Now consider a household consisting of a couple living together, and facing the budget constraint $p'z \leq y$. Following the standard collective approach, our key assumption regarding decision making within the household is that outcomes are Pareto efficient. A standard result of welfare theory (see, e.g., Bourguignon and Chiappori 1994) is that, given ordinality, one can without loss of generality write Pareto efficient decisions as a constrained maximization of the weighted sum $\mu U^f(x^f) + U^m(x^m)$. Here the constraints are the technology constraints that define feasible values of individual consumption vectors x^f and x^m given z, and the budget constraint that defines feasible values of the vector of purchases z. It is important to note that the Pareto weight μ may in general depend on prices, total expenditures, and on a vector s of distribution factors, the latter being defined as variables with no direct impact on preferences, technology or budget constraint, but that may influence the decision process.

We assume the household does not suffer from money illusion, and so write the Pareto weight function as $\mu(p/y, s)$. Possible examples of distribution factors s include individual wages (as in Browning et al., 1994) or non labor income (Thomas 1990), sex ratio on the relevant marriage market and divorce legislation (Chiappori, Fortin and Lacroix 2002), generosity of single parent benefits (Rubalcava and Thomas 2000), spouses' wealth at marriage (Thomas, Contreras and Frankenberg 1997), and the targeting of specific benefits to particular members (Duflo 2000). See also Chiappori and Ekeland (2005) for a general discussion.

ASSUMPTION A2: Given budget and technology constraints and the absence of money illusion, the household makes Pareto efficient decisions, that is, it's choice of x^f and x^m maximizes the weighted sum $\widetilde{U} = \mu(p/y, s)U^f(x^f) + U^m(x^m)$. The Pareto weight function $\mu(p/y, s)$ is positive and twice continuously differentiable in p/y.

All of the functions we define may depend on other variables that we have suppressed for notational simplicity. For example, in our empirical application the functions U^f and U^m also depend on demographic characteristics. We will also often similarly suppress the vector s.

In Assumption A2, $\widetilde{U}$ can be interpreted as a social welfare function for the household, though one in which the relative effects of individual member's utility functions may vary with prices and distribution factors. Alternatively, $\widetilde{U}$ may stem from some specific bargaining model (Nash bargaining for instance), in which distribution factors affect individual threat points. This model also allows for effects such as a wife's utility depending on both her own attained utility level over goods (the value of U^f) and on her spouse's happiness, U^m.

From a bargaining perspective, the Pareto weight $\mu(p/y, s)$ can be seen as a measure of f's influence in the decision process. The larger the value of μ is, the greater is the weight that member f's preferences receive in the resulting household program, and the greater will be the resulting private equivalent quantities x^f versus x^m. One difficulty with using μ as a measure of the weight given to (or the bargaining power of) member f is that the magnitude of μ will depend on the arbitrary cardinalizations of the functions U^f and U^m. Later we will propose an alternative bargaining power measure, the sharing rule, that does not depend upon any cardinalizations.

3.3 The Consumption Technology Function

In our model the household purchases some bundle z, but individual consumptions of the household members add up to some other bundle x, with the difference due to sharing and jointness of consumption. In the technological relation $z = F(x)$, we can interpret z as the inputs and x as the outputs in the household's consumption technology, described by the production function F^{-1}, though what is being 'produced' is additional consumption. This framework is similar to a Becker (1965) type household production model, with the additional restriction that the set of inputs is identical to the set of outputs: instead of using the purchased vector of market goods z to produce different 'commodities' that contribute to utility, the household essentially produces the equivalent of a greater quantity of market goods x via sharing. The unobserved n vectors x^f, x^m, and $x = x^f + x^m$ are private good equivalent

vectors, that is, they are respectively the quantities of transformed goods that are consumed by the female, the male, and in total.

ASSUMPTION A3: Given a purchased bundle of goods z, the feasible set of private good equivalent bundles x^f and x^m are given by the consumption technology function $z = F(x^f + x^m)$.

It will often be convenient to work with a linear consumption technology, which is mathematically identical to Gorman's (1976) linear technology (a special case of which is Barten (1964) scaling), except that we apply it in the context of a collective model.

ASSUMPTION A4: The consumption technology function is linear, so $F(x) = Ax + a$, where A is a nonsingular n by n matrix and a is a n vector.

Most of our theoretical results and all of our empirical work use a linear consumption technology; however, we discuss in Section 4 how our main identification and duality results can be generically extended to arbitrary smooth consumption technologies.

Consider some examples. Let good j be food. Suppose that if an individual or a couple buy a quantity of food z_j, then the total amount of food that the individual or couple can actually consume (that is, get utility from) is $z_j - a_j$, where a_j is waste in food preparation, spoilage, etc.,. If the individuals lived apart, each would waste an amount a_j, so the total amount wasted would be $2a_j$, while living together results in only a waste of a_j. In this simple example the economies of scale to food consumption from living together are a reduction in waste from $2a_j$ to a_j, implying that $x_j = z_j + a_j$, so the consumption technology function for food takes the simple form $z_j = F_j(x) = x_j - a_j$.

Economies of scale also arise directly from sharing. For example, let good j be automobile use, measured by distance traveled (or some consumed good that is proportional to distance, perhaps gasoline). If x_j^f and x_j^m are the distances traveled by car by each household member, then the total distance the car travels is $z_j = (x_j^f + x_j^m)/(1 + r)$, where r is the fraction of distance that the couple ride together. This yields a consumption technology function for automobile use of $z_j = F_j(x) = x_j/(1 + r)$. This example is similar to the usual motivation for Barten (1964) scales, but it is operationally more complicated, because Barten scales fail to distinguish the separate utility functions, and hence the separate consumptions, of the household members.

More complicated consumption technologies can arise in a variety of ways. The fraction of time r that the couple share the car could depend on the total usage, resulting in F_j being a nonlinear function of x_j. There could also be economies (or diseconomies) of scope as well as scale in the consumption technology, e.g., the shared travel time percentage r could be related to expenditures on vacations, resulting in $F_j(x)$ being a function of other elements of x in addition to x_j.

The model also allows for possible diseconomies of scale, e.g., diagonal elements of A could be larger than one or elements of a could be negative.

In the collective model literature sharing and jointness of consumption is usually modeled by assuming some goods are purely private and others purely public (see, e.g., Vermeulen

2000), with additional generality obtained by defining some goods to be the sum of public and private components, e.g., solo car travel and joint car travel could be modeled as two separate goods, one private and one public. Recalling that $x_j = x_j^f + x_j^m$, in our model a purely private good has $z_j = x_j = x_j^f + x_j^m$, while a purely public good would have $z_j = x_j/2$, with the additional constraint that $x_j^f = x_j^m$. Our model includes both $z_j = x_j$ and $z_j = x_j/2$ as special cases, and so includes purely private goods and can at least approximate purely public goods, but we do not impose the constraint $x_j^f = x_j^m$. Such constraints can be imposed in our model, but doing so would interfere with the simple duality results we derive that make our model empirically tractable. We later show how our model can be extended to include purely public goods and other, more general consumption technologies.

Goods that are often cited as purely public, such as heating, may actually be consumed in different quantities by different household members, as our model permits. For example a spouse that stays at home consumes more of the household's heat than one that goes to a job in the daytime. Finally, we note that even a linear technology function F includes many types of joint and shared consumption that cannot be captured by models that only consider purely private and purely public goods (such as economies of scope to consumption), and that our model produces household demand functions that nest as special cases the Barten (1964) and Gorman (1976) models, which are widely used in the empirical demand modeling literature.

3.4 The Household's program

ASSUMPTION A5: The household faces the budget constraint $p'z \leq y$.

Here z is the vector of quantities of the n goods the household purchases, p is market prices and y is the household's total expenditures. Assumptions A2, A3, and A5 together say that the household's consumption behavior is determined by the optimization program

$$\max_{x^f, x^m, z} \mu(p/y)U^f(x^f) + U^m(x^m) \text{ subject to } x = x^f + x^m, \ z = F(x), \ p'z = y \quad (3)$$

To save notation we have suppressed distribution factors s that affect the Pareto weight μ and any demographic or other attributes that may affect μ or the utility functions U^f and U^m or the consumption technology F. The solution of the program (3) yields the household's demand functions, which we denote as $z = h(p/y)$, and private good equivalent demand functions $x^i = x^i(p/y)$.

One way to interpret the program (3) is to consider two extreme cases. If all goods were private and there were no shared or joint consumption, then $F(x) = x$, so z would equal $x^f + x^m$ and the program (3) would reduce to $\max_{x^f, z} \mu(p/y)U^f(x^f) + U^m(z - x^f)$ such that $p'z = y$, which is the standard specification of a collective model for purely private goods (see, e.g., Bourguignon and Chiappori 1994 or Vermeulen 2000). At the other extreme, imagine a household that has a consumption technology function F, but assume that the household's utility function for transformed goods just equaled $U^m(x)$, as might happen if the male were a dictator that forced the other household members to consume goods in the same proportion that he does. In that case, the model would reduce to $\max_z U^m F^{-1}(z)$ such that $p'z = y$, which for

linear F (Assumption A4) is equivalent to Gorman's (1976) general linear technology model, a special case of which is Barten (1964) scales (corresponding to A diagonal and a zero). Our general model, program (3), combines the consumption technology logic of the Gorman or Barten framework with the collective model of a household as either a bargaining or a social welfare maximizing group.

3.5 Duality

To prove identification results and to facilitate empirical application of our model, we derive a dual representation of the household's program. To do so, consider the household as an open economy. From the second welfare theorem, any Pareto efficient allocation can be implemented as an equilibrium of this economy, possibly after lump sum transfers between members. We summarize these transfers by the *sharing rule $\eta(p/y)$*, which is defined as the fraction of a suitably constructed measure of household resources consumed by member f. The household's behavior is equivalent to allocating the fraction $\eta(p/y)$ of resources to member f, and the fraction $1 - \eta(p/y)$ to member m. Each member i then maximizes their own utility function U^i given a Lindahl (1919) type shadow price vector π and their own shadow income η^i to calculate their desired private good equivalent consumption vectors x^f and x^m. This concept of a sharing rule is borrowed directly from the collective model literature (see, e.g., Browning and Chiappori 1998 and Vermeulen 2000 for a survey), though our model is richer than these earlier households models because of the inclusion of the consumption technology function. Formally:

PROPOSITION 1: Let Assumptions A1, A2, A3, and A5 hold. There exists a Lindahl shadow price vector $\pi(p/y)$ and a scalar valued sharing rule $\eta(p/y)$, with $0 \leq \eta(p/y) \leq 1$, such that

$$
\begin{aligned}
x^f(p/y) &= h^f\left(\frac{\pi(p/y)}{\eta(p/y)}\right) \\
x^m(p/y) &= h^m\left(\frac{\pi(p/y)}{1 - \eta(p/y)}\right) \\
z &= h(p/y) = Fx^f(p/y) + x^m(p/y)
\end{aligned}
\tag{4}
$$

This and other propositions are proved in the Appendix. Here $\pi(p/y)$ denotes a vector of equilibrium (shadow) prices within the household economy. One feature of this result that will be used later for identification and empirical tractability is that these shadow prices are the same for both household members. This follows from assuming consumption technologies of the form $z = F(x^f + x^m)$. Our conceptual framework can be immediately extended to arbitrary technologies $z = F^*(x^f, x^m)$, but we would then lose this feature of members facing the same shadow prices.

Without loss of generality, shadow prices are scaled so to make total shadow income of the household be $\pi'(x^f + x^m) = 1$. The sharing rule η is then given by $\eta = \pi'x^f$, which is

the fraction of total shadow income consumed by member f. It follows that $\eta(p/y)$ is a direct measure of the weight given to member f in the outcome of the household decision process. The following proposition shows the relationship between the sharing rule η and the Pareto weight μ.

PROPOSITION 2: Let Assumptions A1, A2, A3, and A5 hold, so by Proposition 1, the functions $\eta(p/y)$ and $\pi(p/y)$ exist. Then

$$\mu(p/y) = -\left[\frac{\partial V^m\left(\frac{\pi(p/y)}{1-\eta(p/y)}\right)}{\partial \eta}\right] / \left[\frac{\partial V^f\left(\frac{\pi(p/y)}{\eta(p/y)}\right)}{\partial \eta}\right] \tag{5}$$

Given utility functions U^t and U^m and a technology F, Proposition 1 shows there exists a unique sharing rule function $\eta(p/y)$ corresponding to any Pareto weight function $\mu(p/y)$ and Proposition 2 shows the converse. Our model does not require identification or estimation of μ (since we use η as our measure of resource allocation or bargaining power), however, if one wanted to use Proposition 2 empirically to recover μ, one would need variation in η that keeps π fixed. For that purpose, it suffices that $\eta(p/y)$ depend on some parameter that can be varied holding shadow prices $\pi(p/y)$ fixed. Distribution factors s are obvious candidates. In general, shadow prices will depend in a complicated way on the consumption technology function, on both members' demand functions and on distribution factors. In particular, changing the members' respective bargaining powers will generally affect shadow prices because it modifies the structure of household demand and shadow prices are not independent of this structure. There is, however, an interesting exception to this rule, which is that $\pi(p/y)$ has a simple tractable functional form that only depends on the consumption technology when that technology is linear (Assumption A4), as follows.

PROPOSITION 3: Let Assumptions A1, A2, A3, A4, and A5 hold. Then Propositions 1 and 2 hold with

$$\pi(p/y) = \frac{A'p}{y - a'p}$$

$$z = h(p/y) = Ah^f\left(\frac{A'p}{y-a'p}\frac{1}{\eta(p/y)}\right) + Ah^m\left(\frac{A'p}{y-a'p}\frac{1}{1-\eta(p/y)}\right) + a. \tag{6}$$

4 Identification

Given functional forms for the consumption technology function, the sharing rule, and member demand functions, Proposition 1 shows how the demand functions for households would be constructed. In the case of linear consumption technologies, these household demands have the simple, explicit form given in Proposition 3. Specifically, given linear technologies a simple parametric modeling strategy is to posit functional forms for indirect utility functions of members, and a functional form for the sharing rule η. Roy's identity then gives functional

forms for the household member demand functions h^f and h^m, and the resulting functional form for the demand functions of couples is given by equation (6).

While this result is convenient for empirical work, it would be useful to know if all the features of our model are nonparametrically identified, since it would be undesirable if our estimates of economies of scale, bargaining power, indifference scales, etc.,. were based in part on untestable parametric assumptions. The question is, given the observable demand functions h^m, h^f, and h, can we identify (and later, estimate) the consumption technology function $F(x)$, the shadow prices $\pi(p/y)$, the sharing rule $\eta(p/y)$, and the private good equivalent demand functions $x^f(p/y)$ and $x^m(p/y)$? We show below that these functions are 'generically' nonparametrically identified, meaning that identification will only fail if the utility and technology functions are too simple (for example, a linear F is not identified if demands are of the linear expenditure system form). This identification in turn implies nonparametric identification of objects of interest such as bargaining power, economies of scale, and our indifference scales.

This identification analysis is at the level of individual consumers and couples. That is, we are identifying structural features of a particular couple (their consumption technology, shadow prices, and sharing rule) from their individual and joint demand functions. To bring these results to data, we will later need to deal with heterogeneity across individuals and couples.

Two concepts that are relevant for identification are private goods and assignable goods. Define a good j to be *private*, or privately consumed, if $z_j = x_j$, so private goods have no economies or diseconomies of scale in consumption. Define a good j to be *assignable* if x_j^f and x_j^m are observed, so assignable goods are goods where we know how much is consumed separately by the husband and by the wife. For example, if we observe how much time each household member uses the family car then auto use is assignable (and not necessarily private, since some of that usage is shared when they ride together). Clothing is private if we observe total purchases of clothing and there are no economies or diseconomies of scale or scope in clothing use. Clothing is both private and assignable if it is private and if we separately observe husbands' clothing use and wives' clothing use. Having some goods be private and/or assignable helps identification, by reducing the information required to go from purchases z_j to private good equivalent consumptions x_j^f and x_j^m.

4.1 Generic Identification

ASSUMPTION A6: The household demand function $h(\cdot)$ and the household member demand functions $h^f(\cdot)$ and $h^m(\cdot)$ are identified.

Identification of the household demand function is straightforward, since we would generally estimate $z = h(p/y)$ using ordinary demand data on observed prices p, total expenditures y, and corresponding bundles z purchased by couples. This estimation step would typically require either seeing a given couple in many different price and total expenditure regimes, or would require an assumption that households with common observable demographic characteristics have similar preferences.

11

Identification of member demand functions h^f and h^m is not immediate, but can be obtained in a few different ways. If individual's indifference curves over goods do not change when they marry or otherwise form households, then h^f and h^m could be estimated (and hence identified) using ordinary demand data on observed prices, total expenditures, and quantities purchased by individual men and women living alone, that is, from observing the consumption demands of singles. More generally, h^f and h^m can be identified if preferences do change upon marriage, as long as those changes can be identified, since in that case we can use singles data, couples data, and knowledge of the changes (e.g., estimated parameters that change values after marriage) to recover h^f and h^m. To the extent that preferences vary across individuals, implementing this step would require assumptions that allow us to match single's demand functions to corresponding couples. In our empirical analysis, we accomplish this matching by assuming that, up to additive error terms, preferences vary only as a function of observable demographic characteristics.

Identification could alternatively be obtained if we can directly observe the consumption of goods by each individual within the household under various price and expenditure regimes, since that would provide direct observation of the functions $x^m(p/y)$ and $x^f(p/y)$ along with $h(p/y)$, which could be used to recover h^f and h^m. In other words, h^f and h^m will be identified if goods are assignable.

In practice, identification and associated estimation of h^f and h^m may be obtained by some mix of all of the above, e.g., h^f and h^m may be identified from a combination of estimated demand functions of singles and of couples, parameterization of changes in preferences over goods resulting from marriage, and make use of assignability of some goods, such as separate observation of men's and women's clothing purchases in the household. We will later illustrate these identification methods in the context of an empirical application.

Older demand models such as Barten (1964), Gorman (1976), and Lewbel (1985), and newer specifications such as shape invariance (see, e.g., Pendakur (1999)) are models of how demands vary across households of different sizes, and so exploit data from both singles and couples to jointly identify parameters that are common to both, as well as identifying parameters that characterize the differences between the two. The idea of using singles data to identify some couple's parameters has also been used in some labor supply models, including Barmby and Smith (2001) and Bargain et. al. (2004). An example of using smoothness assumptions to identify features of individual household member consumptions is Chesher (1998).

Ordinary revealed preference theory shows that the utility functions of members U^m and U^f, or equivalently the indirect utility functions V^m and V^f, are identified up to arbitrary monotonic transformation given the member demand functions h^f and h^m. Our goal now is to use the additional information of observed household demand functions h to identify the other features of the household's program.

PROPOSITION 4: Let Assumptions A1, A2, A3, A4, A5, and A6 hold. If the number of goods in the system is $n \geq 3$, then the functions $\pi(p/y)$, $\eta(p/y)$, $x^m(p/y)$ and $x^f(p/y)$, and the technology parameters represented by A and a, are all generically identified.

The Pareto weight $\mu(p/y)$ depends upon the arbitrary cardinalization of member utilities, but if those cardinalizations are fixed and distribution factors are present, then $\mu(p/y)$ may

also be identified using equation (5), given Propositions 2 and 3.

The identification in Proposition 4 is only generic, in the sense that it only shows that we have (many) more equations than unknowns, roughly analogous to satisfying order condition but not necessarily the rank condition for identification in traditional linear simultaneous systems. We show in the Appendix that Proposition 4 can be extended. Generic nonparametric identification holds not just for linear consumption technologies $z = Ax + a$, but also for arbitrary technologies, that is, with $z = F(x)$ for general monotonic vector valued functions F. In other words, our general methodology does not depend on any functional form assumptions to obtain identification. See Chiappori and Ekeland (2004) for more discussion of generic identification in collective household models.

The sense in which these identification results are generic is that identification could fail to hold, but only when the true functional forms for demand equations are too simple, thereby causing too many of these equations (equation (35) in the appendix for $t = 1, ..., T$) to be linearly dependent. The model will be identified as long as the true functional forms for either demands or technologies are sufficiently nonlinear, since nonlinearities in functional form will tend to eliminate linear dependencies across equations.

We illustrate this point in the appendix and in the next section. In particular, in the appendix we first show that identification fails with a linear technology in the linear expenditure system, because in that case one cannot separate linear demand function coefficients from linear technology coefficients. However, we also show in that case that adding a single nonlinear term to just one demand equation for just one household member then suffices to identify the entire model. This shows that even the slightest departures from linearity suffice for identification.

Another example we provide in the appendix is linear preferences with a nonlinear consumption technology. In this example, the model is identified unless almost all the parameters in the technology equal zero. The next section provides a third example, proving identification for the popular Almost Ideal and Quadratic Almost Ideal demand systems, as long as at least one the demand functions of husbands and wives differ by an income coefficient.

To summarize: a linear technology is not fully identified if members' demand functions are also linear. However, having either a nonlinear technology or at least one nonlinear member demand function will generally suffice for identification. Exceptions are difficult to construct and require extreme restrictions, as in the nonlinear technology example where nonidentification requires many parameters to be exactly zero, or in the example of the next section where nonidentification requires husbands and wives to have identical income coefficients for every good they consume.

4.2 The case of QUAIDS individual demands

For our empirical application, we use the following convenient method of constructing functional forms for estimation. First, choose ordinary indirect utility functions for members m and f, and let h^m and h^f be the corresponding ordinary Marshallian demand functions. Assume a linear consumption technology F. Next choose a functional form for the sharing rule η, which could simply be a constant, or a function of measures of bargaining power such

as relative wages of the household members or other distribution parameters. Proposition 3 then provides the resulting functional form for the household demand function $h(p/y)$, and ensures that a corresponding household program exists that rationalizes the choice of functions h^m, h^f, and η. To satisfy Assumption A6 we assume initially that members have the same demand functions h^m and h^f as single men and single women living alone, respectively (we later consider some very simple parameterizations of preference change resulting from marriage).

In our empirical application we assume singles have preferences given by the Integrable QUAIDS demand system of Banks, Blundell and Lewbel (1997). For $i = f$ or m, let $w^i = \omega^i(p/y^i)$ denote the n-vector of member i's budget shares w_k^i ($k = 1, ..., n$) when living as a single, facing prices p and having total expenditure level y^i. The QUAIDS demand system we estimate takes the vector form

$$\omega^i(p/y^i) = \alpha^i + \Gamma^i \ln p + \beta^i \left[\ln\left(y^i\right) - c^i(p) \right] + \frac{\lambda^i}{b^i(p)} \left[\ln\left(y^i\right) - c^i(p) \right]^2 \qquad (7)$$

where $c^i(p)$ and $b^i(p)$ are price indices defined as

$$c^i(p) = \delta^i + (\ln p)'\alpha^i + \frac{1}{2}(\ln p)'\Gamma^i \ln p \qquad (8)$$

$$\ln b^i(p) = (\ln p)'\beta^i. \qquad (9)$$

Here α^i, β^i and λ^i are n-vectors of parameters, Γ^i is an $n \times n$ matrix of parameters and δ^i is a scalar parameter which we take to equal zero, based on the insensitivity reported in Banks, Blundell, and Lewbel (1997). Adding up implies that $e'\alpha^i = 1$ and $e'\beta^i = e'\lambda^i = \Gamma^i e = 0$ where e is an n-vector of ones. Homogeneity implies that $\Gamma^{i\prime}e = 0$ and Slutsky symmetry is equivalent to Γ^i being symmetric. The above restrictions yield the integrable QUAIDS demand system, which has the indirect utility function:

$$V^i\left(\frac{p}{y^i}\right) = \left[\left(\frac{\ln\left(y^i\right) - c^i(p)}{b^i(p)} \right)^{-1} + \lambda^{i\prime} \ln(p) \right]^{-1} \qquad (10)$$

for $i = f$ and $i = m$. The singles demand functions $\omega^i(p/y^i)$ in equation (7) are obtained by applying Roy's identity to equation (10). Deaton and Muellbauer's (1980) Almost Ideal Demand System (AIDS) is the special case of the integrable QUAIDS in which $\lambda^i = 0$. Each element of $h_k^i(p/y^i)$ of $h^i(p/y^i)$ for $i = m$ and $i = f$, that is, member $i's$ quantity demand function for good j is given by $h_k^i(p/y^i) = \omega_k^i(p/y^i)y^i/p_k$.

PROPOSITION 5: Let Assumptions A1, A2, A3, A4, A5, and A6 hold. Assume members have demand functions given by the integrable AIDS or QUAIDS. Assume $\beta^f \neq \beta^m$ and each element of β^f, β^m, and the diagonal of A is nonzero. Then the functions $F(x)$, $\eta(p/y)$, $x^m(p/y)$ and $x^f(p/y)$ are identified.

Proposition 5 confirms that the QUAIDS model for member's preferences, with a linear household technology, is sufficiently nonlinear to permit identification of all the components

of the couple's model. The assumption regarding nonzero and unequal elements in Proposition 5 can be relaxed. See the proof in the Appendix for details.

That proof shows that not only are the relevant household model functions are identified, but that they are massively overidentified. In fact, given member demands, most of the parameters comprising these functions can be identified just from observing the household's demand function for one good. This suggests that there is considerable scope for maintaining identification even if one adds many additional parameters to characterize changes in tastes that result from marriage.

4.3 Multiple Households

All of our results so far have been expressed in terms of a given couple with members m and f. Our Identification results, specifically Assumption A6, depend on being able to observe the demand functions $h(\cdot)$, $h^f(\cdot)$, and $h^m(\cdot)$ of the couple and of singles f and m. In practice these demand functions will need to be estimated, which will generally require combining data across multiple consumers.

This raises the issue that consumers can have varying preferences. To apply our identification theorems to estimated data, it will be convenient to assume that single's preferences vary by observable characteristics d^f and d^m. In our later empirical application these will include age and education level among other demographic variables. As long as these characteristics are observable both for singles and for each member of a couple, we can rewrite the demand functions as $h(\cdot, d^f, d^m)$, $h^f(\cdot, d^f)$, and $h^m(\cdot, d^m)$, and thereby relate the demand functions of couples with the proper associated singles demand functions.

Note that we take these characteristics, and the demand functions that depend on them, as given. Some, like education level, are themselves choices, but these choices are assumed to have been made prior to our analysis. Similarly, the choice of which type of person to marry is taken as given. Our results should therefore all be interpreted as conditional upon the outcomes of choice characteristics and marriage markets.

Conditioning on observable characteristics in this way still leaves the potential problem of unobserved preference heterogeneity. There are a few ways one can bridge the gap between individual level identification results and data drawn from populations of consumers with possibly unobserved heterogeneity of tastes. Lewbel (2001) provides sufficient conditions to insure that the conditional mean demand functions of consumers with heterogenous tastes equal the demand functions of a single utility maximizing individual. We could assume those conditions hold separately for men and for women (after conditioning on d^f and d^m), and thereby interpret our estimated results as applying to conditional mean demand functions. Note Lewbel (2001) specifically considers the case of QUAIDS preferences, as in our empirical example. More simply, as we discuss later in Section 7.3, in our empirical application we treat unobserved heterogeneity as ordinary additive error terms, and apply the QUAIDS based identification to estimated conditional mean functions.

Before turning to empirical results, the next sections discuss various applications and extensions of our results. We will for clarity continue to express these results in terms of a given couple, but all of these results can be immediately extended to deal with households of varying

15

(observable) types by inclusion of characteristics vectors d^f and d^m.

5   Applications

Here we summarize some potential uses for our model of household consumption behavior. These uses are in addition to the standard applications of demand models, such as evaluation of price and income elasticities.

5.1   Resource Allocation Across Members and Bargaining Power

The sharing rule $\eta(p/y)$ provides a direct measure of the allocation of household resources among the household members, and hence may also be interpreted as a measure of relative bargaining power after taking altruism into account. If all goods were private, with no economies of scale or scope in consumption (i.e., if $x = z$) then η would exactly equal the share of total expenditures y that were used to purchase the bundle x^f consumed by member f. Browning and Chiappori (1998), using household data alone, show that the relative bargaining power measures can only be identified up to an arbitrary location. By contrast, Propositions 4 and 5 show that in our model, by combining demand functions of households and the demand functions of individual members, the sharing rule η is completely identified.

We argue that the Pareto weight μ is a less tractable measure of bargaining power than η, because μ depends on the unobservable cardinalizations of the utility functions of individual household members, while η is recoverable just from observable demand functions. However, if these cardinalizations are known, then $\mu(p/y)$ can be calculated and used with our model.

5.2   Economies of Scale in Consumption

Previous attempts to measure economies of scale in consumption have required very restrictive assumptions regarding the preferences of household members (see, e.g., Nelson 1988). In contrast, given estimates of the private good equivalents vector $x = x^f + x^m$ from our model, we may calculate $y/p'x$, which is a measure of the overall economies of scale from living together, since y is what the household spends to buy z and $p'x$ is the cost of buying the private good equivalents of z. This economy of scale measure takes the form

$$\frac{y}{p'x} = \frac{y}{p'\left(h^f\pi/\eta + h^m\pi/(1-\eta)\right)}.$$

This measure does *not* directly provide an estimate of adult equivalence scales or indifference scales. The reason is that, in general, the shadow prices used within the household are not proportional to market prices (this is exactly Barten's intuition). It follows that individuals living alone would *not* buy the bundles x^f and x^m, which were optimal given shadow prices. They would in general reach a higher level of utility by re-optimizing using market instead of shadow prices.

We may also calculate the corresponding economies of scale in consumption for each good k separately as z_k/x_k. This can also be interpreted as a measure of the publicness or

privateness of each good. For a purely private good with no jointness of consumption, z_k/x_k equals one, while goods that are mostly shared will have z_k/x_k close to one half. When the consumption technology has the Barten form, z_k/x_k equals the Barten scale for good k.

5.3 Indifference Scales

Given each household member i's private good equivalents x^i, we define member i's collective based equivalent income, y^{i*}, as the minimum expenditures required to buy a vector of goods that is on the same member i indifference curve as x^i. The ratio $S^i = y^{i*}/y$ is then what we call member i's *indifference scale,* so S^i is the fraction of the household's income that member i would need to buy a bundle of privately consumed goods at market prices that put her on the same indifference curve over goods that she attained as a member of the household.

Like the above described economies of scale and bargaining power measures, these indifference scales do not depend on any utility cardinalizations or assumptions of interpersonally comparable utility.

Recall that $V^f(p/y^f)$ is the indirect utility function of member f, and the private good equivalent vector consumed by member f in the household is $x^f(p/y) = h^f\left[\pi(p/y)/\eta(p/y)\right]$, where h^f is the Marshallian demand function obtained from $V^f(p/y^f)$ by Roy's identity. The indifference scale $S^f(p/y, \eta)$ is defined as the solution to the equation

$$V^f\left(\frac{p/y}{S^f(p/y, \eta)}\right) = V^f\left(\frac{\pi(p/y)}{\eta}\right) \tag{11}$$

This definition of $S^f(p/y, \eta)$ only depends upon an ordinal representation of utility, i.e., it does not depend on the chosen cardinalization for member f's utility function. Specifically, and in marked contrast to traditional adult equivalence scales, replacing V^f with any monotonic transformation of V^f in equation (11) leaves S^f unchanged. The expression for member m defining S^m, is the same as equation (11), replacing f with m and replacing η with $1 - \eta$.

As discussed earlier, these indifference scales $S^i(p/y, \eta)$ could be used for poverty, life insurance, and wrongful death, calculations. For example, in the case of a wrongful death or insurance calculation, a woman (or a mother and her children, if V^f is defined as the joint utility function of a mother and her children) would need income $S^i(p/y, \eta(p/y))y$ to attain the same standard of living without her husband that she (or they) attained while in the household with the husband present and total expenditures y. This would take into account the share of household resources $\eta(p/y)$ that she consumed, and would be exactly sufficient to compensate for the loss of economies of scale and scope from shared consumption, but would not compensate her for the husband's consumption, or for grieving, loss of companionship, or other components of utility that are assumed to be separable from consumption of goods. It would also not compensate for any change in preferences over goods that might occur as a result of the death.

Another interesting indifference scale to construct is $S^f(p/y, \eta)/S^m(p/y, 1-\eta)$, the ratio of how much income a woman needs when living alone to the income a man needs when living alone to make each as well off as they would be in a household. Even if men and women

had identical preferences, this ratio need not equal one, because if the sharing rule η is bigger than one half, then the wife receives more than half of the household's resources, and hence would need more income when living alone to attain the same standard of living. To separate bargaining effects from other considerations, one might instead calculate $S^i(p/y, 1/2)$, which is the indifference scale assuming equal sharing of resources. For example, the ratio $S^f(p/y, 1/2)/S^m(p/y, 1/2)$ might better match the intuition of an equivalence scale comparing women to men.

In other applications, one might want to consider the roles of equivalent incomes and the sharing rule jointly. For example, given poverty lines for singles, one might define the corresponding poverty line for the couple as the minimum y such that, by choosing η optimally, each member i of the couple would have an equivalent income $S^i y$ equal to his or her poverty line as a single.

This discussion illustrates an important feature of our model, which is the ability to separately evaluate and measure the roles of individual household member preferences, of intra-household allocation and control of resources, and of economies of scale from sharing and joint consumption, on indifference scales and other welfare calculations.

6 Additional Results

6.1 Barten and Gorman Scales

Gorman's (1976) general linear technology model assumes that household demands are given by

$$z = Ah\left(\frac{A'p}{y - a'p}\right) + a \tag{12}$$

Barten (1964) scaling (also known as demographic scaling) is the special case of Gorman's model in which $a = 0$ and A is a diagonal matrix; see also Muellbauer (1977). Demographic translation is Gorman's linear technology with A equal to the identity matrix and a non zero, and what Pollak and Wales (1992) call the Gorman and reverse Gorman forms have both A diagonal and a nonzero. These are all standard models for incorporating demographic variation (such as the difference between couples and individuals) into demand systems.

The motivation for these models is identical to the motivation for our linear technology F, but they fail to account for the structure of the household's program. Even if the household members have identical preferences ($h^f = h^m = h$) and identical private equivalent incomes ($\eta = 1/2$), comparison of equations (6) and (12) shows that household demand functions will still not actually be given by the Gorman or Barten model. In fact, comparison of these models shows that household demands will take the form of Gorman's linear technology, or some special case of Gorman such as Barten, only if demands are linear in prices (i.e., the linear expenditure system), or if η is zero or one, corresponding to one household member consuming all the goods (or more realistically, one member using their own preferences to dictate the entire household's purchases). Gorman's (1976) famous comment, "If I have a wife and child, a penny bun costs threepence," rationalizes the Barten scale model but implicitly assumes a dictator imposing his taste for buns on his family members.

Gorman (1976) allows A to be rectangular rather than just square, as in Becker (1965) type models where the outputs are composites that generate utility. Our theoretical framework can be immediately extended to include this case, but then it becomes considerably more difficult to obtain identification because we would not observe singles consuming these composites when they arise due to economies of scale and scope in consumption.

6.2 Other Consumption Technologies

We first show how our framework could be extended to allow for the existence of pure public goods within the household. For notational simplicity, assume for now that there is only one purely public and one purely private good, denoted by X and x respectively (with market prices P and p respectively), and that the household technology is separable across goods. Pareto efficiency implies that the vector (x^f, x^m, X) solves the program (with obvious notations):

$$\max_{x^f, x^m, z, Z} \mu(p/y, P/y)U^f(x^f, X) + U^m(x^m, X)$$

$$\text{subject to } z = f(x^f + x^m), \ Z = F(X), \ pz + PZ = y$$

An equivalent formulation is

$$\max_{x^f, x^m, z, Z} \mu(p/y, P/y)U^f(x^f, X^f) + U^m(x^m, X^m) \tag{13}$$

$$\text{subject to } z = f(x^f + x^m), \ Z = F(\max(X^f, X^m)), \ pz + PZ = y$$

Here, X^i can be interpreted as the quantity of the public good desired by member i. In principle, we allow X^f and X^m to differ, though the solution to this program will impose $X^f = X^m$. Equation (13) is formally equivalent to equation (3), except that F is now a function of the pair of individual consumptions, not of their sum only. In this model the demands of member f depend on the vector $(\pi/\eta, \pi^i/\eta)$, and similarly for member m. Each member i's (Lindahl) decentralizing price π^i for the public good has $\pi^i > 0$.

If the technologies for the public and private goods are both linear, so $f(x) = Ax + a$ and $F(X) = BX + b$, then the shadow prices satisfy

$$\pi = \frac{Ap}{y - ap - bP}, \qquad \pi^f + \pi^m = \frac{BP}{y - ap - bP}$$

so the sum of individual prices for the public good plays the same role as the price of a private commodity. A similar analysis could be used to incorporate externalities within the household, whether positive or negative. Note, however, that with negative externalities individual prices π^i for the externality could be negative.

Returning to the general case of n goods, a more general program that encompasses both models (13) and (3) as special cases is

$$\max_{x^f, x^m, z} \mu(p/y)U^f(x^f) + U^m(x^m) \text{ subject to } z = F(x^f, x^m), \ p'z = y \tag{14}$$

which allows for completely general consumption technology functions $F(x^f, x^m)$. Like the other programs, the general household model (14) can be decentralized, but now the complete vector of Lindahl shadow prices will be different for members m and f, being functions of the separate derivatives of F (assuming these derivatives exist) with respect to x^m and x^f. In contrast, our program (3) assumption of an additive technology $z = F(x^f + x^m)$ makes the derivatives of F with respect to x^m and x^f equal, resulting in shadow prices $\pi(p/y)$ that are the same for m and f. The problem with nonadditive technology models is that our nonparametric identification result does not hold for them. In particular, it may be the case that several different technology functions $F(x^f, x^m)$ could generate the same observable demand functions. Identification would then need to rely either on detailed functional form assumptions or on the availability of additional information, such as assignability of goods.

Suppose that our model, which assumes an additive consumption technology $z = F(x^f + x^m)$, were applied to data derived from a more general technology $z = F(x^f, x^m)$. How would this misspecification bias the results? Our additive technology implies that both individuals face the same shadow prices π. This will provide a good approximation to a more general technology if agents have similar marginal valuations for the goods, and hence similar shadow prices.

In the particular example of a purely public good, our additive model will provide a good approximation to actual behavior if the household members have a similar willingness to pay for the public good. For example, the shadow price of purely public good should be one half the market price, leading in the Barten technology $\pi_k = A_k p_k/y$ to a coefficient of $A_k = 1/2$. If additional economies of scale are present, the coefficient may become smaller than one half; conversely, diseconomies of scale (or partly private use) would generate Barten coefficients between .5 and 1.

6.3 Preference Changes

Our model assumes that the demand functions of household members, h^f and h^m, are identified. The simplest way to satisfy this assumption is to assume that individual's preferences for goods do not change when they marry. This does not mean that individuals consume the same bundles as singles in couples (because the model of a couple permits economies of scale and scope in consumption) nor does it rule out individuals getting utility from marriage, rather, all it implies is that the indifference curves of single men or women living alone are the same as the indifference curves associated with the utility functions U^m and U^f in the household model. If this assumption holds, then the functions h^m and h^f can be estimated directly by observing the consumption behavior of single men and single women.

Instead of assuming tastes do not change, we could also attain identification by estimating the demand functions for singles, parameterize how those preferences change as a result of marriage, and use couple's data to estimate these preference change parameters along with the other features of the household models, that is, the consumption technology and sharing rule. For example, suppose that the utility function of individual i living alone is ordinally represented by $U^i(L^{-1}z^i)$ for some matrix L. We could write this more generally as utility given by $T(U^i(L^{-1}z^i), L)$, where T is monotonic in its first element. The utility function of

individual i in the household is ordinally represented by $U^i(z^i)$, so the matrix L embodies the changes in taste for goods that result from marriage. Then the demand functions for a single i living alone will be

$$z^i = \widetilde{h}^i(p/y^i) = L^{-1}h^i\left(L^{-1\prime}p/y^i\right)$$

and therefore, given a linear consumption technology, the demand functions of the couple, expressed in terms of the demand functions $\widetilde{h}^i$ for the members, will be

$$z = h(p/y) = AL\widetilde{h}^f\left(\frac{L'A'p}{y - a'p}\frac{1}{\eta(p/y)}\right) + AL\widetilde{h}^m\left(\frac{L'A'p}{y - a'p}\frac{1}{1 - \eta(p/y)}\right) + a \quad (15)$$

In particular, if A and L are diagonal and $a = 0$, so both the technology and preference change have a Barten form, then for each good k, A_k is a technological economies of scale parameter that should lie between 1/2 and 1 (assuming no diseconomies of scale), while L_k captures common taste changes over the private good equivalent $x_k = (z_k/A_k)$. If $L_k = 1$ then tastes for good k do not change with marriage. If $L_k > 1$ then married individuals like good k less than when they are single.

In this example, taste changes and economies of scale affect couple's demands in exactly the same way. If we had estimated the model assuming no taste change, we would obtain Barten technology estimates of $L_k A_k$ instead of just A_k, and so would end up attributing taste changes to economies of scale. Additional information would then be required to separately identify L and A. For example, any good k that is known to be purely private would have $A_k = 1$, so an estimated Barten scale other than one for that good would equal L_k. Similarly, the estimated Barten scale for any good k that is known to have $L_k = 1$ (no taste change) would equal A_k. Any good k that is assignable (meaning that x_k is directly observed) would allow direct estimation of h_k^i, which could then be used in the household's model to estimate some or all of the parameters of A and compared to the singles demand functions $\widetilde{h}_k^i$ to estimate some or all of the taste change parameters L. All prices appear in the demand functions for each good, so it may be possible to identify many of the technology and taste change parameters from observations of $\widetilde{h}_k^i$ and h_k^i for a small number of goods k that are private or assignable or do not undergo a change in taste upon marriage.

In our empirical application we estimate a Barten technology and apply the model with two different interpretations of the result. One is to assume no taste change, interpreting the estimated parameters as equalling each A_k. We also apply the model assuming the estimated Barten parameters equal $L_k A_k$, where $L_k = 1$ if A_k is less than one, otherwise we assume $A_k = 1$ and the estimated parameter corresponds to taste change L_k. We also experiment with exploiting assignability of clothing.

In our indifference scale and other welfare calculations, we do not attempt to compensate for taste changes, since that, like traditional adult equivalence scales, would require untestable cardinalization assumptions, e.g., equivalence scales based on equating $U^i(L^{-1}x^i)$ to $U^i(x^i)$ would differ from equivalence scales based on equating the observationally equivalent representations of preferences $T\left(U^i(L^{-1}z^i), L\right)$ to $T\left(U^i(z^i), I\right)$ for monotonic transformations T. This is precisely the problem we avoid with our indifference scale calculations, i.e., this problem does not arise for technology parameters A instead of L, because preferences do not

depend upon A. As a result, using our indifference scales, compensation is purely for re-sources one received as a household member, and the loss of economies of scale and scope in consumption. We believe this more closely matches the usual notion of maintaining a standard of living than any attempt to also compensate for changes in tastes.

Note that separating L from A will not be important if the dollar effect of a change in tastes for goods is small, implying that L_k is close to one for most goods. For example, joint consumption of heating is likely to have a much larger effect on measured cost savings of living together than on any change in tastes for heat. Separating L from A is largely irrelevant for some relevant calculations such as demand elasticities, or the interpretation of the sharing rule η as a measure of bargaining power.

In our empirical application we also considered other models of taste change, such as allowing some of the intercept parameters α^i to differ for singles living alone versus those in couples, but attempts to estimate such models were abandoned when the resulting models, which are identified through functional form, failed to converge numerically.

6.4   Externalities

The analysis of positive externalities within a household is similar to that of jointly consumed goods, while negative externalities affect shadow prices in the opposite direction. Suppose, e.g., that a commodity consumed by one member has a negative impact on the other person. Since the second person's marginal willingness to pay is negative, the first person's individual shadow price must exceed the household price. In the absence of economies of scale or scope for this commodity, the estimated shadow price would be larger than the market price, leading to a Barten coefficient greater than one. More generally, if a commodity (say, tobacco) is consumed by both members and each individual's consumption has a negative impact on the other person's welfare, then each individual price may exceed the household price (because at the optimum each member will have a positive marginal willingness to pay to *get rid of* the other person's consumption). Overall, in the absence of preference changes, an estimated Barten coefficient larger than one would be suggestive of the presence of negative consumption externalities for the commodity under consideration.

The differences between externalities versus changes in preferences would be hard if not impossible to distinguish empirically. For example, if we found that married couples consume proportionally more restaurant meals than either single men or single women, this could be due either to externalities (each member's marginal utility of restaurant dining might be en-hanced by the presence of their spouse) or by a simple change in preferences (members liking restaurant dining better when married). The two explanations have different implications, e.g., under the externality story only dinners taken with one's spouse are increased in value, in con-trast to the preference change explanation. In practice, available data would hardly allow the estimation of such subtle distinctions.

6.5   Labor Supply

Our model assumes consumption is separable from leisure. There are some advantages but also some complications that would arise from relaxing this assumption and bringing leisure into the model. As a practical matter, few household data sets simultaneously provide consumption allocation, time use, and wages.[1] If such data were available, and both spouses work, then our model could be immediately applied by defining husband and wife's leisure as additional goods, using their wages as prices, and redefining y accordingly to include total hours times wages. The considerable variation in wages that exists, and the fact that leisure is an assignable good, would both be helpful in improving the precision with which resource shares and other features of our model would be identified and estimated with this data.

Relative wages, which are often used as a distribution factor in collective household models, would with labor supply no longer be usable in that role. However, one advantage of our model is precisely that we do not require distribution factors.

A complication that would arise regarding labor supply is when one or both spouses are unemployed. In this case our model would need to be modified both to account for the fact that leisure demand for the unemployed spouse hits a maximum (analogous to the effect of rationing in standard demand system models), and that the price of leisure for that household member becomes unobserved. A version of our framework that aggregates up to just total consumption and leisure, and ignores Barten scales, but incorporates these complications associated with some spouses not working, is Lise and Seitz (2011).

7   Empirical Application

7.1   Data

We use Canadian FAMEX data from 1974, 1978, 1982, 1984, 1986, 1990 and 1992 on annual expenditures, incomes, labour supply and demographics for individual ('economic') households. Prices are taken from Statistics Canada. Composite commodity prices are constructed as the weighted geometric mean of the component prices with budget shares averaged across the strata (couples, single males and single females) for weights. The (non-durable) goods we model are: food at home, restaurant expenditures, clothing, vices (alcohol and tobacco), transport, services and recreation.

We sample single females, single males and couples with no one else present in the household, with women and men aged less than or equal to 42 and 45, respectively. All agents are in full year full time employment with non-negative net and gross incomes and all couples own a car (to avoid modeling the demand patterns of the 5% of couples in the original sample that did not own a car). Table 1 presents summary statistics for our data.

---

[1]A notable exception is the Dutch data used by Cherchye *et al* (2012).

	Single	Single	Couples	
	females	males	Husband	Wife
Sample size	1393	1574	1610	
Net income*	26, 137	30, 890	56, 578	
Total expenditure*	11, 855	13, 645	23, 175	
Wife's share of gross income	-	-	0.43	
bs(food at home)	0.18	0.16	0.18	
bs(restaurant)	0.11	0.15	0.11	
bs(clothing)	0.16	0.09	0.14	
bs(vices)	0.07	0.12	0.08	
bs(transport)	0.21	0.25	0.25	
bs(services)	0.17	0.10	0.12	
bs(recreation)	0.11	0.13	0.11	
car	0.64	0.78	1	
home owner	0.13	0.23	0.58	
city dweller	0.85	0.81	0.83	
age	29.9	31.1	30.2	28.2
higher education	0.20	0.25	0.21	0.19
Francophone	0.19	0.17	0.21	0.18
Allophone	0.08	0.08	0.08	0.07
White collar	0.43	0.40	0.39	0.37
Notes. * mean for 1992 (when prices are unity).				

Table 1: Descriptive statistics

7.2 Budget Shares for Singles

Our model starts with a utility derived functional form for the budget shares of singles. For this we use the QUAIDS model described in equations (7), (8), (9), and (10). We also estimate a (unitary) QUAIDS for couples to provide an initial comparison with singles and with our collective household model. To define the system for a consumer of type i ($=$ female, male or a couple) we have the vector budget share system:

$$\omega^i \left(p^h / y^h \right) = \alpha^i + \Gamma^i \ln p^h + \beta^i \left[\ln \left(y^h \right) - c^i \left(p^h \right) \right] + \frac{\lambda^i}{b^i \left(p^h \right)} \left[\ln \left(y^h \right) - c^i \left(p^h \right) \right]^2 \quad (16)$$

where $c^i(p)$ and $b^i(p)$ are price indices defined as

$$c^i \left(p^h \right) = \delta^i + (\ln p^h)' \alpha^i + \frac{1}{2} (\ln p^h)' \Gamma^i \ln p^h \quad (17)$$

$$\ln b^i \left(p^h \right) = (\ln p^h)' \beta^i. \quad (18)$$

Our model allows for substantial observable preference heterogeneity by letting the α and β parameters in the QUAIDS depend on demographic characteristics. This dependence takes the form of

$$\alpha_k^i = \alpha_{k0}^i + \sum_{m=1}^{M_\alpha} \alpha_{km}^i d_m \quad (19)$$

$$\beta_k^i = \beta_{k0}^i + \sum_{m=1}^{M_\beta} \beta_{km}^i d_m \quad (20)$$

where $M_\beta = 2$ for singles (dummies for owning a car and being a home owner) and $M_\beta = 1$ for couples (a home owner dummy). For the intercept demographics we have $M_\alpha = 18$ for couples and $M_\alpha = 13$ for singles. These demographics include four regional dummies; a house ownership dummy and a dummy for living in a city. For singles we also have a car ownership dummy. As well we include individual specific demographics: age and age squared, a dummy for having more than high school education, dummies for being French speaking or neither English nor French speaking and an occupation dummy for being in a white collar job. Although these demographics are highly correlated within couples we need to include both sets so that we can allow for the dependence of couples' budget shares on both sets of individual characteristics. We have 24 and 28 parameters per good/equation for singles and couples respectively.

7.3 The stochastic structure

We allow for two sources of stochastic variation. The first is due to unobserved preference heterogeneity over and above that due to observable preference heterogeneity based on demographic characteristics such as age that are discussed in the previous section. In principle, for individuals i all of the parameters α^i, β^i and λ^i could be subject to latent heterogeneity. In practice, we cannot identify unobserved heterogeneity in β^i and λ^i so we only allow the α^i

terms to be heterogeneous. This introduces heterogeneity in a nonlinear fashion since each of the individual good specific α_k^i parameters enter all equations through the $c^i(p)$ function. In practice, we ignore the effect that filters through the price index[2] and simply include additive errors u^{ih} in equation (16):

$$\omega^i\left(p^h/y^h\right) = \alpha^i + \Gamma^i \ln p^h + \beta^i \left[\ln\left(y^h\right) - c^i\left(p^h\right)\right] + \frac{\lambda^i}{b^i\left(p^h\right)} \left[\ln\left(y^h\right) - c^i\left(p^h\right)\right]^2 + u^{ih} \quad (21)$$

Our estimator is GMM as described below, which allows the errors u^{ih} to have arbitrary heteroskedasticity of unknown form, and so the random variation in u could have a distribution that depends on observables. For example, our estimator, would cover the case where all of the α_{km}^i elements in equation (19) were replaced with random coefficients. However, even if we interpret a component of u^{ih} as unobserved preference heterogeneity in the parameter vector $\alpha^i + u^{ih}$, our structural analysis would ignore the impact of this preference variation due to u. So, e.g., a given couple with demographic characteristics vector d is assumed to be comprised of individuals having the male and female components of d, irrespective of the value of u.

The second source of stochastic variation, and hence a component of u^{ih} in our model, is due to measurement error in the budget shares. Suppose, for example, that the agent buys an expensive winter coat in the survey year. This will induce a high budget share for clothing *and* higher than normal total expenditure. Consequently y^h will be correlated with the clothing element of u^{ih}. We follow the convention in the demand literature and use log net income as the principal instrument for log total expenditure. The full set of instruments for singles are the demographics and log relative prices (the included variables) plus log real net income (defined as the log of nominal net income divided by a Stone price index computed for our seven nondurables goods) and its square, the product of log real net income with the car and home ownership dummies and absolute log prices. This gives 25 instruments for singles. For instruments for couples we use the instruments used for singles (without, of course, the car dummy variables), with individual specific values for husbands and wives, where appropriate. We also include logs of the individual gross incomes and the ratio of the wife's gross income to total gross income (we discuss these variables below). This gives 32 instruments for couples.

We estimate the joint system by GMM. We assume that the error terms are uncorrelated across households but are correlated across goods within households. To allow for adding-up we drop one demand equation and let ω_h now be the $(n-1)$ vector of budget shares for the first $n-1$ goods consumed by household h. Denote the vector of all parameter values by θ and let $\widehat{\omega}_h(\theta)$ be the predicted budget shares for household h. The error vector for household h is thus given by $\hat{u}_h(\theta) = \omega_h - \widehat{\omega}_h(\theta)$. Let the numbers of single females by H_f. Denote the $(1 \times g_f)$ vector of instruments single female h by z_h^f. The vector of moment conditions is given by the $(n-1)\, g_f \times 1$ vector:

$$\nu^f\left(\theta\right) = \sum_{h=1}^{H_f} \hat{u}_h'\left(I_{n-1} \otimes z_h^f\right) \quad (22)$$

---

[2]Lecocq and Robin (2006) provide evidence that this (very convenient) simplification is innocuous.

and similarly for single males and couples. Define the weighting matrix for single females by:

$$W_f = \left(\sum_{h=1}^{H_f} \left(I_{n-1} \otimes z_h^f \right)' \tilde{u}^f \tilde{u}^{f'} \left(I_{n-1} \otimes z_h^f \right) \right)^{-1} \tag{23}$$

where the residuals $\tilde{u}^c$ are taken from a first stage GMM with an identity weighting matrix. The GMM criterion for single females is:

$$\min_\theta \left\{ v^f(\theta)' W_f v^f(\theta) \right\} \tag{24}$$

and similarly for single males and couples.

The over-identifying restriction tests for the singles are implicit tests of the homogeneity and symmetry restrictions as well as the exclusion restrictions to take account of the endogeneity of total expenditure. The OI test statistics $\chi^2(30)$ for the system without symmetry imposed are 31.2 and 40.5 for single females and single males respectively; for couples the $\chi^2(39)$ statistic for the system without symmetry is 69.2. Thus we reject the OI restriction tests for couples but not for singles.

To illustrate the differences in demands of single men, single women, and couples, Figure 1 presents fitted demand (Engel curve) plots for our seven goods, calculated for agents at the mean level of demographic variables, who live in Ontario in 1992 (the year and region for which prices are set to unity), with total expenditures ranging from the first to the ninth decile. We shift the plots for couples to the left in these figures to make them comparable to the singles plots. We find that food at home is a necessity for all and that transport and services are necessities for singles. Restaurants and recreation are luxuries for all strata and vices are a luxury for singles. Clothing is a necessity for low income single men but a luxury for high income single men (the QUAIDS quadratic term is significant). Vices are a necessity and transport is a luxury for low income couples. In both Table 1 and in these figures, the demands of couples more closely resemble the demands of single women than those of single men.

The question now is whether our model of a couple as a combination of two singles can capture these effects.

7.4 The Joint Model

For our empirical application, we assume singles demands have the integrable QUAIDS functional form described in equations (7), (8), (9), (10), (19) and (20). The joint model therefore includes one set of α^i, β^i, Γ^i and λ^i parameters (including demographic components) for men, $i = m$, and another complete set of these parameters for women, $i = f$. Let $w_k^f = \omega_k^f(p/y^f)$ and $w_k^m = \omega_k^m(p/y^m)$ denote the QUAIDS budget share of consumption of good k for single women and single men, respectively.

For couples we assume a Barten type technology function defined as

$$z_k = A_k x_k \tag{25}$$

27

for each good k, and so is equivalent to the general linear technology $z = Ax + a$ when the matrix A is diagonal and overheads a are zero. The shadow prices for this technology are

$$\pi_k = \frac{A_k p_k}{y} \tag{26}$$

where the couple faces prices p and has total expenditure level y. As we shall see, this technology function is particularly convenient for budget share models.

Our parameterization for η is:

$$\eta = \frac{\exp\left(s'\delta\right)}{1 + \exp\left(s'\delta\right)} \tag{27}$$

where s is a household specific vector of distribution factors and δ is a vector of parameters. This logistic form bounds the sharing rule between zero and one. The distribution factors we consider are four region dummies; a home-ownership dummy; a dummy for living in a city; the difference in age between husband and wife; the wife's share in total gross income, and log household total expenditure (deflated by a Stone price index).[3] This gives a total of 9 candidate distribution factors.

With the technology function (25) and corresponding shadow prices (26) and the sharing rule (27), equation (6) yields the following simple expression for the couple's budget shares:

$$\omega_k\left(p/y\right) = \eta\omega_k^f\left(\frac{\pi}{\eta}\right) + (1 - \eta)\omega_k^m\left(\frac{\pi}{1 - \eta}\right) \tag{28}$$

(where π is a vector of prices given by (26)). This equation shows that, with a Barten type technology, the budget shares of the couple are equal to a weighted average of the budget shares of its members, with weights given by the income sharing rule η and $1 - \eta$. Thus the parameter η represents both the fraction of resources controlled by the wife and the extent to which the household's demands resemble her demands, evaluated at shadow prices.

The parameters of the couples model consist of all the QUAIDS parameters of both the single's budget shares, ω^f and ω^m; the Barten scales A_k and the parameters of the sharing rule η. There are two options for estimation. The first option is to use a two-step procedure in which we first estimate the preference parameters using singles and then plug these into (28) to estimate the Barten and sharing rule parameters. This has the virtue of simplicity in estimation (two conventional QUAIDS systems and then one small joint system) and has a transparent link with our framework. The alternative option is the one-step estimator in which we estimate the preference parameters jointly with the technology and sharing parameters. Although this latter gives a highly nonlinear system, we found that the two step procedure gave a much worse fit than the one-step (the details are given below) and we consequently focus on the latter.

We have 258 preference parameters ($6 * 24 - 15 = 129$ symmetry constrained QUAIDS parameters for each of men and women). We also have 7 Barten scale parameters and 10

---

[3]Since there is nothing in the theory to rule out that prices enter the sharing rule, we also tried including the six log relative prices . They were never strongly significant.

distribution factors (the 9 listed above plus the constant); this gives a total of 275 parameters. We have 492 instruments (for each of the 6 goods there are 25 instruments for each single strata and 32 for couples), giving a maximum degrees of freedom of 217 for the most general model. The weighting matrices for single females and single males, W_f and W_m respectively, are taken from the QUAIDS in the previous subsection; see (23).The weighting matrix for couples, W_c, is derived using an iterated GMM for the full system, starting with an initial identity weighting matrix. The GMM criterion is:

$$\min_{\theta} \left\{ v^c (\theta)' W_c v^c (\theta) + v^f (\theta)' W_f v^f (\theta) + v^m (\theta)' W_m v^m (\theta) \right\} \qquad (29)$$

where θ is the full parameter vector and the instrument matrices are defined as in (22).

7.5 Parameter Estimates and Analysis

The main results for our preferred model are displayed in Table 2, with standard errors in parentheses. This model, which has two distribution factors (a home ownership dummy and log deflated total expenditure) is the result of a general to specific search which began with nine candidate distribution factors.[4] We present standard errors but note that to test for restrictions in nonlinear models, Wald tests are not invariant to the parameterisation of hypotheses (see Gregory and Veall (1985)). For testing, we therefore use tests that compare the fit, as measured by the GMM criterion of equation (29), of the constrained and unconstrained models and take the difference as a χ^2 statistic with degrees of freedom equal to the number of restrictions.

The model estimated with the two step procedure described in the previous subsection (using the estimates of the budget shares of single) had a considerably poorer fit than the one-step estimation procedure (an equation (29) criterion value of 544.5), so our preferred model uses the one-step estimation procedure. Still, in our model the test of over-identifying restrictions rejects. This is a common finding in demand system estimation, where utility maximization gives rise to many strong over-identifying restrictions.

In our preferred model the Barten scales are restricted to be between 0.5 and 1, corresponding to the extremes of purely public and purely private goods, respectively. The consequent loss of fit from imposing these economic constraints in any of our models is very small, so any formal test for the inequality restriction would not reject. The resulting estimate for clothing has a Barten scale constrained at 1 and the scale for transportation is constrained at 0.5 while the scales for the five other goods are unconstrained. The final two columns in Table 2 present tests for setting each Barten scales to 0.5 (testing if each good is wholly public) and 1 (testing if each good is purely private) respectively. For each goods other than clothing, transport and vices, we reject that the good is wholly private or public (that is, $A_k = 0.5$ or 1). The estimate of vices is very imprecise and we cannot reject that it is purely private or purely public.[5]

---

[4]Attempts to add additional extensions, such as fixed costs to the consumption technology, did not yield a significantly better fit.

[5]This imprecision in the vices equation is a common finding in demand studies; essentially the latent heterogeneity dominates any other effect.

7.6 The sharing rule

The mean value of the sharing is 0.63 with minimum and maximum values of 0.48 and 0.75 respectively. Table 3 gives the values of the sharing rule η for different sets of characteristics. The benchmark household has a home-owning husband and wife with the same gross income and with *median* total expenditure for Ontario in 1992 (the province/year in which prices are normalised to unity). Table 3 presents the estimates of the level and variations in the sharing rule.

At our benchmark values, the wife's share is 0.58. If this was purely for private goods then this would be rather high but it also reflects the effects of joint or public consumption. This value reflects that the household demand functions look more like single women's demand functions than single men's demand functions. From rows 2 and 3 we see that the wife's share is increasing in total expenditure so that in poorer households the wife's share is lower.

Row 4 of the table shows that the home-ownership status variable has a very strong impact with renting households apparently having a much higher share for the wife. As we saw above, the pattern of home ownership between singles and couples is very different so this result is probably due at least in part to a selection effect rather than a genuine difference in intrahousehold allocations. One of the candidate distribution factors our specification search excluded is the wife's share of income; this was not found to be significant. This is very different from the results reported in Browning, Bourguignon, Chiappori and Lechene (1994), which examines conventional sharing rules for private goods using similar data on couples. The latter paper did not include the home ownership dummy in the sharing rule and included real household income instead of total expenditures. Nonetheless the qualitative differences are striking and suggest that allowing for household technology may have a major impact on our conclusions regarding intrahousehold allocations.

7.7 Barten scales and Economies of scale

We now consider estimates of the private good equivalent consumption for married men and married women and the resulting economies of scale. Given estimates of the budget share systems for singles, $w_k^i = \omega_k^i(p/y^i)$, and for households (equation 28), the private good equivalent quantities for each household member are given by

$$
\begin{aligned}
x_k^f &= \frac{\eta \omega_k^f \left(\pi/\eta\right)}{\pi_k} = \frac{\omega_k^f}{A_k}\eta y \\
x_k^m &= \frac{(1-\eta)\omega_k^m \left(\pi/(1-\eta)\right)}{\pi_k} = \frac{\omega_k^m}{A_k}(1-\eta)y
\end{aligned}
\tag{30}
$$

The equivalent incomes for each are given by:

$$
x^f = \sum_k x_k^f = \eta y \sum_k \frac{\omega_k^f}{A_k}
\tag{31}
$$

$$
x^m = \sum_k x_k^m = (1-\eta)\, y \sum_k \frac{\omega_k^m}{A_k}
\tag{32}
$$

Sharing rule				
Intercept	0.56	(0.28)		
Homeowner	−0.46	(0.16)		
Log total expenditure	0.28	(0.29)		
Barten scales				
			$\chi^2(1)$ for $A_k =$	
			0.5	1
			Public	Private
Food at home	0.77	(0.08)	17.2	7.3
Restaurant	0.66	(0.08)	4.3	12.8
Clothing	1	(0.34)	7.4	0
Vices	0.65	(0.47)	0.1	1.4
Transport	0.5	(0.09)	0	13.1
Services	0.75	(0.12)	7.2	2.2
Recreation	0.74	(0.16)	3.9	4.8
df	224			
χ^2	326.1			
Standard errors shown in brackets				

Table 2: Results

Household characteristics	Wife's share η
Benchmark	0.58
First quartile total expenditure	0.56
Third quartile total expenditure	0.60
Renters	0.69

Table 3: Sharing rule implications

Note the equivalent incomes can differ even if there is equal sharing ($\eta = 1$) since the two partners may have different budget shares and reap different rewards from the publicness of goods. As an extreme example, if she only consumes private goods then $x^f = \eta y$ and he only consumes public goods then $x^m = 2(1 - \eta)y$. With equal sharing, $\eta = 0.5$, this gives him twice as much equivalent income as her.

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

$$R = \frac{p'(x^f + x^m)}{y} - 1 = \frac{p'(x^f + x^m - z)}{p'z}$$

If all good are private (respectively, public) then $R = 0$ (respectively, $R = 1$).

In Table 4 we present estimates of member's private good equivalent budget shares; equivalent incomes and the scale measure R. To disentangle the effects of Barten scaling and sharing we present values given $\eta = 0.5$ (equal sharing), and for the benchmark estimate from the previous subsection, $\eta = 0.58$. In the table, the first set of estimates (labelled "EBS") give the equivalent budget shares for different goods for each member at unit market prices; these are defined as $100\omega_k^f(\pi/\eta)$ and $100\omega_k^m(\pi/(1-\eta))$ evaluated at shadow prices $\pi_k = A_k$ for each good k. Each column of this table therefore shows, in percent terms, how each member f and m allocates their share of household resources to the various goods. The differences in budget shares between the m and f columns reflect the differences in the budget shares for single men and women seen in Table 1. With equal sharing, budget shares are similar for food at home but women have higher budget shares for clothing and personal services and lower shares for restaurant, vices, transport and recreation. The predicted budget shares with a share of 0.58 for women are similar to those with equal sharing except for food at home and clothing. The changes here reflect the chnages in relative shares and that food at home is a necessity and clothing is a luxury.

Wife's share, η	0.50		0.58	
EBS	f	m	f	m
Food at home	16.8	16.3	14.7	18.7
Restaurant	7.7	15.3	8.3	14.5
Clothing	15.3	4.5	16.8	3.9
Vices	6.8	12.8	7.0	12.4
Transport	24.3	29.7	23.5	30.6
Services	17.6	6.9	17.9	7.2
Recreation	11.3	14.5	11.7	12.6
Her equivalent income	1.73		2.00	
His equivalent income	1.84		1.54	
Scale economy, R	0.52		0.51	

Table 4: Implications of estimates

The next two rows of Table 1 show equivalent incomes (the actual total household income is 2.34). Even when we have equal sharing the equivalent incomes are different with husbands

Wife's share η	0.50	0.58
Indifference scale for women, S^f	0.72	0.83
Indifference scale for men, S^m	0.80	0.66

Table 5: Adult equivalence scales

having a higher value. This reflects the fact that wives have a higher budget share for clothing which is a private good and a lower budget share for transport and vices which are largely public. Note that we cannot infer from this that husbands are better off with equal sharing since that would require interpersonal comparisons. If we take the preferred value for sharing, $\eta = 0.58$, then wives have a higher equivalent income. The sum of the two equivalent incomes is similar across the η values.

The overall scale economy measure, R, is given in the last row of Table 4. In the equal sharing scenario the estimate of $R = 0.52$, implies that it would cost the couple 52% more to buy the (private equivalent) goods they consumed if there had been no shared or joint consumption. The estimate with the preferred value of $\eta = 0.58$ gives a very similar value. We conclude that the economies of scale are largely independent of the sharing rule.

7.8   Welfare comparisons and Indifference scales

The finding that two singles require about 50% more total expenditure than a couple to buy the same equivalent consumption levels $x^f + x^m$ represents an upper bound on the total amount needed if they actually physically lived apart (while maintaining the same preferences). This is because differences between market prices and household shadow prices mean that singles living alone can re-allocate and more cheaply attain the same indifference curves that x^f and x^m lie on. That is what our indifference scales $S^i(p/y, \eta)$ measure.

In Table 5 we present indifference scales and also the (technological) scale economy parameter, for comparison purposes. With equal sharing we have $S^f + S^m = 1.52$, that is, 52% more income is required in total. This is the same as the value without re-allocation, so that the gains each can get by re-budgeting are very modest. With unequal sharing ($\eta = 0.58$) the wife requires 83% of joint income to be as well when she is single. In this case, the two require 49% more income jointly. Once again the gains from re-allocating are modest.

7.9   Main Empirical results

A summary of our main empirical results is as follows.

In all specifications we considered, tests of over-identifying restrictions are rejected, suggesting that richer models of preference heterogeneity may be needed in the future.

At the mean of the data the wife's share of income is 0.63. The only variable that has a large impact on the wife's share is whether the couple are home-owners. For a benchmark home-owning household the share is 0.58. These estimates are rather high compared to other (later) studies based on our methodology, such as Lewbel and Pendakur (2008). Correspondingly, A woman needs 83% of joint income to be as well off single as she was in a couple whereas a

man needs only 66% of joint income. Living alone the partners need 49% more income than when single for each to be as well off. A couple would need 51% more income to buy the same amount of consumption if all goods were private.

The previous literature on collective household models assumed all goods are either purely private or purely public, while we find that most goods are neither. Our results are that clothing is purely private, transport is purely public, and vices could be either. Food at home, restaurants, personal services and recreation are all partly private and partly public.

8   Conclusions

We model households in terms of the utility functions of its members, a bargaining or social welfare function, and a consumption technology function. By employing a collective model and a separate consumption technology, we combine data from singles and couples to identify a version of equivalence scales that we call indifference scales, as well as consumption economies of scale, individual members' control of resources within the household, and other related concepts, without assumptions regarding cardinalizations or interpersonal comparability of preferences. Unlike traditional adult equivalence scales that attempt to compare the utility of an individual to that of a household, our indifference scales directly compare the cost of living for an individual within a household to the cost of living of the same individual living alone.

We also provide duality results that facilitate the empirical application of our model. These include the use of indirect utility functions to describe individual household members' preferences, an income sharing rule to model the outcome of the household bargaining or resource allocation process, and shadow prices that embody economies of scale and scope arising from joint consumption. We demonstrate generic nonparametric identification of our model.

We empirically apply our model to Canadian consumption data on single men, single women and married couples. Among other empirical results, we find that couples save the equivalent of about one third of their total expenditures through shared and joint consumption of goods, that wives on average control between a half and two thirds of household resources, and that singles need to spend between one half to three fourths as much as couples to attain the same standard of living by themselves as they attain as members of a two person household.

Given our framework, one useful area for further work would be the development of more detailed alternative identifying assumptions. Without preference changes or externalities within the household, our model is very much over identified, in fact, in our QUAIDS specification many of the parameters of the entire consumption technology could be separately identified from the couple's demands for a single good. This overidentification might be exploited by relaxing the model in ways that incorporate and identify possible externalities (among individuals within the household) or by allowing for more general models of individual's preference changes that result from marriage or household formation in general. Direct observation of some elements of the private good equivalent consumption vector x would also aid identification, by distinguishing changes in preferences from pure economies of scale or scope in consumption. Example of such data would be direct measures of the quantity of food wasted (discarded) by singles and couples, or measures of the fraction of time couples drive

together versus alone. Exploiting assignability of some goods also helps theoretical identification, though we did not find empirical improvements from assuming assignability of clothing.

Alternative identifying conditions that minimize reliance on the consumption data of singles would also be valuable for extending the model to directly include the utility functions of children, which for now must be assumed to be joint with the utility of an adult household member. Another useful direction for future research would be the development of empirically tractable duality and identification results for more general, nonlinear consumption technology functions. It would also be valuable to introduce unobserved preference variation in a more structural way than in our current application.

The generality of our estimated model in the number of goods and functional forms of its components resulted in many numerical and convergence difficulties. It would be useful to investigate simplifications that preserve the main features of the model. For example, the sharing rule can in principle be identified and estimated just from combining data on singles and couples for a single good.

An interesting extension of our model would be the introduction of time allocation as another good. In particular, our consumption technology function could directly embody the extent to which couples achieve economies of scale in time use by combining chores (such as one member cooking meals for both), and economies of scope such as saving both time and gasoline by commuting together. More generally, the consumption technology function could be merged with home production functions, as in rural households that grow crops for personal consumption.

While our particular identifying and functional form assumptions should be open to debate and improvement, we believe that the general model we have provided is an appropriate framework for estimating and analyzing adult equivalence scales, consumption economies of scale, household members' bargaining power, and other concepts relating to household preferences, consumption, and demand behavior. Some authors that have begun applying and extending our framework include Lewbel and Pendakur (2008), Lise and Seitz (2011), Donni and Bargain (2011), and Dunbar, Lewbel and Pendakur (2012). In particular, the latter two papers provide ways to include children, having their own utility functions and resource shares, into our framework.

It could be interesting to extend the model to other aspects of collective decision making, such as household formation and dissolution, fertility, the allocation of time, durables consumption and savings decisions.

9 Appendix

9.1 Proofs of Propositions

PROOF OF PROPOSITION 1: The Lagrangian for the household's program can be written as $\mu(p/y)U^f(x^f) + U^m(x^m) + \lambda \left[1 - (p/y)'F(x^f + x^m)\right]$ which gives first order conditions

$$\mu(p/y)\frac{\partial U^f(x^f)}{\partial x_k^f} = \frac{\lambda}{y}\sum_{j=1}^n p_j \frac{\partial F^j(x^f + x^m)}{\partial x_k^f}, \quad \frac{\partial U^m(x^m)}{\partial x_k^m} = \frac{\lambda}{y}\sum_{j=1}^n p_j \frac{\partial F^j(x^m + x^m)}{\partial x_k^m}$$

Define π to be the vector of elements π_k, $k = 1, ..., n$ and define η by

$$\pi_k = \frac{\sum_{j=1}^{n} (p_j/y) \, \partial F^j(x)/\partial x_k}{\sum_{i=1}^{n} x_i \sum_{j=1}^{n} (p_j/y) \, \partial F^j(x)/\partial x_i} \tag{33}$$

$$\eta = \sum_{i=1}^{n} x_i^f \pi_i \tag{34}$$

evaluated at the optimizing values $x^f = x^f(p/y)$, $x^m = x^m(p/y)$, and $x = x^f(p/y) + x^m(p/y)$. By construction $1 - \eta = \sum_{i=1}^{n} x_i^m \pi_i$. The above first order conditions can then be rewritten as

$$\mu(p/y) \frac{\partial U^f(x^f)}{\partial x_k^f} = \kappa \pi_k$$

$$\frac{\partial U^m(x^m)}{\partial x_k^m} = \kappa \pi_k$$

where $\kappa = (\lambda/y) \sum_{i=1}^{n} x_i \sum_{j=1}^{n} p_j \partial F^j(x)/\partial x_i$. Then

$$\frac{\partial U^f(x^f)/\partial x_k^f}{\sum_{i=1}^{n} x_i^f \partial U^f(x^f)/\partial x_i^f} = \frac{\pi_k}{\sum_{i=1}^{n} x_i^f \pi_i} = \frac{\pi_k}{\eta}$$

$$\frac{\partial U^m(x^m)/\partial x_k^m}{\sum_{i=1}^{n} x_i^m \partial U^m(x^m)/\partial x_i^m} = \frac{\pi_k}{\sum_{i=1}^{n} x_i^m \pi_i} = \frac{\pi_k}{1 - \eta}$$

and these are identical to the first order conditions of the programs

$$\max_{x^f} U^f(x^f) \text{ subject to } \pi' x^f = \eta$$

$$\max_{x^m} U^m(x^m) \text{ subject to } \pi' x^m = 1 - \eta$$

which define the functions h^f and h^m.

PROOF OF PROPOSITION 2: From the proof of Proposition 1 we have that a first order condition of the household's program is

$$\mu(p/y) \frac{\partial U^f(x^f)}{\partial x_k^f} = \kappa \pi_k$$

The first order condition of member f's optimization $V^f(\pi/\eta) = \max\{U^f(x^f) \mid \pi' x^f \leq \eta\}$ gives $\partial U^f(x^f)/\partial x_k^f = \lambda^f \pi_k$ and using the envelope theorem we obtain

$$\mu(p/y) \frac{\partial V^f(\pi/\eta)}{\partial \eta} = \lambda_0.$$

36

Apply the same logic to member m gives

$$\frac{\partial V^m(\pi/(1-\eta))}{\partial(1-\eta)} = \lambda_0$$

The proposition follows from equating these two expression for λ_0 and solving for $\mu(p/y)$.

PROOF OF PROPOSITION 3: Let A_{jk} denote the $j,k'th$ element of A and a_j denote the $j'th$ element of a. Then Equation (33) with $F(x) = Ax + a$ is

$$\pi_k = \frac{\sum_{j=1}^n p_j A_{jk}}{\sum_{i=1}^n x_i \sum_{j=1}^n p_j A_{ji}} = \frac{\sum_{j=1}^n p_j A_{jk}}{\sum_{i=1}^n x_i \sum_{j=1}^n p_j A_{ji}}$$

so $\pi = A'p/(p'Ax)$ and $p'Ax = p'A(A^{-1}(z-a)) = p'(z-a) = y - p'a$. Given π, equation (6) the follows from equation (4) with $z = A(x^f + x^m) + a$. Having $\pi = A'p/(y - a'p)$ means that π is not a function of U^f, U^m or μ, so any changes in those function that change the resulting optimal x^f for a given p/y can change $\eta = x^{f'}\pi$ while holding π fixed.

PROOF OF PROPOSITION 4: Take a finite set of price vectors $p^1, ..., p^T$ where $T \geq n+10$. These may be interpreted as prices in different time periods. For any p^t and any y, let $z^t = h(p^t/y)$ denote the corresponding household demands. From Proposition 3 we have

$$z^t = Ah^f\left(\frac{A'p^t}{y - a'p^t}\frac{1}{\eta(p^t/y)}\right) + Ah^m\left(\frac{A'p^t}{y - a'p^t}\frac{1}{1 - \eta(p^t/y)}\right) + a, \quad t = 1, ..., T \quad (35)$$

This provides, for each t, $(n-1)$ independent equations, hence a total of $(n-1)T$ equations. The unknowns are the matrix A, the vector a and the scalars $\eta^t = \eta(p^t/y), t = 1, ..., T$; hence a total of $n^2 + n + T$ unknowns. If $n \geq 3$, then with $T \geq n + 10$, the number of equations exceeds the number of unknowns. This shows that, generically on the demand functions h^f and h^m and on the sharing rule η, the equations are linearly independent. With linear independence, the system is generically (over) identified. Given identification of A and a, we then have identification of $F(x) = Ax + a$ and $\pi(p/y) = A'p/(y - a'p)$. The sharing rule $\eta(p/y)$ is then identified for all p/y from equation (6) and identification of x^m and x^f follows from equation (4).

PROOF OF PROPOSITION 5: The AIDS quantity demand for good j consumed by member i is

$$h_j^i(p/y^i) = \frac{y^i}{p_j}\left(\alpha_j^i + \Gamma_j^i \ln p + \beta_j^i\left[\ln(y^i) - c^i(p)\right]\right)$$

where

$$c^i(p) = (\ln p)'\alpha^i + \frac{1}{2}(\ln p)'\Gamma^i \ln p$$

$$\ln b^i(p) = (\ln p)'\beta^i$$

37

α^i and β^i are n-vectors of parameters, Γ^i is an $n \times n$ matrix of parameters, $e'\alpha^i = 1$, $e'\beta^i = \Gamma^i e = 0$ where e is an n-vector of ones, and Γ^i is symmetric.

Linear technology implies that the shadow price for good j is

$$\pi_j = \frac{A_{\cdot j} p}{y - a'p} = \frac{\sum_k A_{kj} p_k}{y - a'p}$$

where $A_{\cdot j}$ is the transpose of the j'th column of A and A_{kj} is the k'th element of $A_{\cdot j}$.

With members having AID demand functions and linear consumption technology, the couple's quantity demand for good k is

$$h_k(p/y) = \sum_j A_{kj} \left[\begin{array}{c} \frac{\eta}{\pi_j} \left(\alpha_j^f + \Gamma_j^f \ln \pi + \beta_j^f \left[\ln(\eta) - c^f(\pi) \right] \right) + \\ \frac{1-\eta}{\pi_j} \left(\alpha_j^m + \Gamma_j^m \ln \pi + \beta_j^m \left[\ln(1-\eta) - c^m(\pi) \right] \right) \end{array} \right] + a_k$$

$$h_k(p/y) = (y - a'p) \sum_j A_{kj} \left[\begin{array}{c} \eta \left(\frac{1}{A_{\cdot j} p} \right) \left(\alpha_j^f + \Gamma_j^f \ln \left(\frac{Ap}{y-a'p} \right) + \beta_j^f \left[\ln(\eta) - c^f \left(\frac{Ap}{y-a'p} \right) \right] \right) \\ + (1-\eta) \left(\frac{1}{A_{\cdot j} p} \right) \left(\begin{array}{c} \alpha_j^m + \Gamma_j^m \ln \left(\frac{Ap}{y-a'p} \right) \\ + \beta_j^m \left[\ln(1-\eta) - c^m \left(\frac{Ap}{y-a'p} \right) \right] \end{array} \right) \end{array} \right] + a_k$$

Given $h_k(p/y)$ for each good k, the constants a_k are identified as the intercept terms, and we may identify the summation by $h_k(p/y) - a_k/(y - a'p)$. Now use

$$c^i(\pi) = c^i \left(\frac{Ap}{y - a'p} \right) = (\ln Ap)'\alpha^i - (\ln y - a'p) + \frac{1}{2}(\ln(Ap))'\Gamma^i \ln(Ap)$$

to write this summation term as

$$\sum_j \frac{A_{jk}}{A_{\cdot j} p} \left[\begin{array}{c} \eta \left(\begin{array}{c} \alpha_j^f + \Gamma_j^f \ln(Ap) + \\ \beta_j^f \left[\ln(\eta) - \left((\ln Ap)'\alpha^f - (\ln y - a'p) + \frac{1}{2}(\ln(Ap))'\Gamma^f \ln(Ap) \right) \right] \end{array} \right) \\ + (1-\eta) \left(\begin{array}{c} \alpha_j^m + \Gamma_j^m \ln(Ap) + \\ \beta_j^m \left[\ln(1-\eta) - \left((\ln Ap)'\alpha^m - (\ln y - a'p) + \frac{1}{2}(\ln(Ap))'\Gamma^m \ln(Ap) \right) \right] \end{array} \right) \end{array} \right]$$

For each good k, the coefficient of $\ln y$ in this expression is

$$-\sum_j \frac{\eta \beta_j^f + (1-\eta)\beta_j^m}{\sum_\ell (A_{\ell j}/A_{kj}) p_\ell}$$

where the summation is now over all goods j for which A_{kj} is not equal to zero. Variation in y therefore provides identification of the above expression for each good k. Variation of p in the above expression for each k identifies

$$\frac{\eta \beta_j^f + (1-\eta)\beta_j^m}{A_{\ell j}/A_{kj}}$$

38

for triplets of goods j, k, ℓ having $A_{kj} \neq 0$ and $A_{\ell j} \neq 0$. The above expression with $\ell = k$ identifies η, since β_j^m and β_j^f are identified from singles, and therefore $A_{\ell j}/A_{kj}$ is also identified. Define $\overline{A}_{\ell j} = A_{\ell j}/A_{jj}$ and $d_j = A_{jj}$. Identification of $A_{\ell j}/A_{kj}$ implies identification of $\overline{A}_{\ell j}$ for all j, ℓ. What remains is to identify each d_j.

The term $A_{kj}/A_{.j}\,p = 1/\sum_\ell (\overline{A}_{\ell j}/\overline{A}_{kj})p_\ell$ is identified for each j, which implies based on the large summation term above that

$$\eta \left(\alpha_j^f + \Gamma_j^f \ln(Ap) + \beta_j^f \left[\ln(\eta) - \left((\ln Ap)'\alpha^f + \tfrac{1}{2}(\ln(Ap))'\Gamma^f \ln(Ap) \right) \right] \right)$$
$$+ (1 - \eta) \left(\alpha_j^m + \Gamma_j^m \ln(Ap) + \beta_j^m \left[\ln(1 - \eta) - \left((\ln Ap)'\alpha^m + \tfrac{1}{2}(\ln(Ap))'\Gamma^m \ln(Ap) \right) \right] \right)$$

is identified for each good j. The component of this expression that is quadratic in $\ln(Ap)$ is

$$\ln(Ap)' \left(-\eta\beta_j^f \frac{1}{2}\Gamma^f - (1 - \eta)\beta_j^m \frac{1}{2}\Gamma^m \right) \ln(Ap)$$
$$= \sum_k \sum_\ell \ln\left(\sum_j d_j \overline{A}_{kj}p_j \right) \ln\left(\sum_j d_j \overline{A}_{\ell j}p_j \right) \varphi_{k\ell}$$

where $\varphi_{k\ell}$ and $\overline{A}_{kj}$ are already identified for all j, k, ℓ. Variation in p in this expression then identifies d_j, which completes the identification.

Note that the technology and sharing rule are substantially overidentified. Given the demand functions of members, the couple's demands for each good k separately identify η, a (from the $y - a'p$ term) and $A_{\ell j}/A_{kj}$ for all j, k, ℓ having $A_{kj} \neq 0$. Therefore, observing couple's demands for all goods, not just one, greatly overidentifies these terms. In particular, if the $k'th$ row of A has all nonzero elements, then (given the demands of singles) the technology and the sharing rule can be completely identified just by observing couple's demands for the $k'th$ good. In addition to the overidentification resulting from observing the couple's demands for multiple goods, only the terms involving the intercept, the coefficient of $\ln y$, and the quadratic in $\ln(Ap)$ were used for identification, so all the other terms in the couple's demands, such as the linear in $\ln(Ap)$ term, provide additional overidentifying restrictions. These various overidentifying assumptions may be used to relax the assumptions regarding unequal and nonzero parameters in the statement of Proposition 5.

Similarly, the above proof applies when single's have QUAIDS demands, since in that case the only difference in the model is the addition of more terms (in particular, a quadratic in $\ln y$) which provide more overidentifying restrictions.

9.2   Nonparametric Identification

9.2.1   'Generic' Identification: an illustration

Here we illustrate the generic identification of Proposition 4 by showing that identification can fail when individual demands belong to the linear expenditures system (LES) class, but that even a minimal departure from linearity suffices to restores (over)identification.

Begin with a simple form for the wife's preferences, namely Cobb Douglas utility:

$$U^f = \sum_i \alpha_i \log\left(x_i^f\right)$$

The husband's preference are also Cobb Douglas except for a 'small', linear deviation:

$$U^m = \sum_i \beta_i \log\left(x_i^m\right) + \varepsilon\gamma_1 x_1^m$$

where $\varepsilon > 0$ is 'small'. Here, $\varepsilon = 0$ would give LES individual demands for both spouses, whereas for $\varepsilon > 0$ the husband's utility has a Cobb Douglas form for all goods except good 1. Finally, let technology take the linear form:

$$z_i = A_i\left(x_i^m + x_i^f + a_i\right)$$

where the A's and the a's are parameters to be identified.

The household's program is:

$$\max \sum_i \beta_i \log\left(x_i^m\right) + \varepsilon\gamma_1 x_1^m + \mu \sum_i \alpha_i \log\left(x_i^f\right)$$

$$\text{under } \sum_i p_i z_i = \sum_i p_i A_i\left(x_i^m + x_i^f + a_i\right) = y$$

and first order conditions give:

$$
\begin{aligned}
\mu\alpha_i &= \lambda A_i p_i x_i^f \\
\beta_i &= \lambda A_i p_i x_i^m, i \geq 2 \\
\beta_1 + \varepsilon\gamma_1 x_1^m &= \lambda A_1 p_1 x_1^m
\end{aligned}
$$

where μ is the wife's Pareto weight and λ denotes the Lagrange multiplier of the budget constraint. Adding up gives:

$$Y \equiv y - \sum_i p_i A_i a_i = \frac{1 + \mu + \varepsilon\gamma_1 \frac{\beta_1}{\lambda A_1 p_1 - \varepsilon\gamma_1}}{\lambda}$$

therefore

$$\lambda = \frac{\left(A_1 p_1 + Y\varepsilon\gamma_1 + \mu A_1 p_1 + K\right)}{2Y A_1 p_1}$$

where

$$K = \sqrt{A_1^2 p_1^2 + \mu^2 A_1^2 p_1^2 + 2\mu A_1^2 p_1^2 + Y^2\varepsilon^2\gamma_1^2 - 2Y\varepsilon\gamma_1 A_1 p_1 - 2Y\mu\varepsilon\gamma_1 A_1 p_1 + 4Y\varepsilon\beta_1\gamma_1 A_1 p_1}.$$

Note the equation actually has two solutions in λ, but only one is non zero when $\varepsilon = 0$.

It follows that:

$$x_i^f = \frac{\mu \alpha_i}{\lambda A_i p_i} = \frac{2A_1 p_1 \mu \alpha_i Y}{A_i p_i \left(A_1 p_1 + Y \varepsilon \gamma_1 + \mu A_1 p_1 + K\right)}$$

$$x_i^m = \frac{\beta_i}{\lambda A_i p_i} = \frac{2Y A_1 p_1 \beta_i}{A_i p_i \left(A_1 p_1 + Y \varepsilon \gamma_1 + \mu A_1 p_1 + K\right)}, \quad i \geq 2$$

$$x_1^m = \frac{\beta_1}{\lambda A_1 p_1 - \varepsilon \gamma_1} = \frac{2Y A_1 p_1 \beta_1}{\left(A_1 p_1 + Y \varepsilon \gamma_1 + \mu A_1 p_1 + K\right) A_1 p_1 - 2Y A_1 p_1 \varepsilon \gamma_1}$$

and finally

$$x_i = \frac{2A_1 p_1 \left(\mu \alpha_i + \beta_i\right) Y}{A_i p_i \left(A_1 p_1 + Y \varepsilon \gamma_1 + \mu A_1 p_1 + K\right)}, \quad i \geq 2$$

$$x_1 = \frac{2A_1 p_1 \mu \alpha_i Y}{A_i p_i \left(A_1 p_1 + Y \varepsilon \gamma_1 + \mu A_1 p_1 + K\right)} + \frac{2Y A_1 p_1 \beta_1}{\left(A_1 p_1 + Y \varepsilon \gamma_1 + \mu A_1 p_1 + K\right) A_1 p_1 - 2Y A_1 p_1 \varepsilon \gamma_1}$$

Now if $\varepsilon = 0$, $\lambda = (\mu + 1)/Y$ and

$$x_i = \frac{\beta_i + \mu \alpha_i}{A_i p_i (\mu + 1)} Y = \frac{\beta_i + \mu \alpha_i}{A_i p_i (\mu + 1)} \left(y - \sum_j p_j A_j a_j\right)$$

then the technology is not identified. Indeed, $p_i x_i$ is linear in (y, p) with coefficients equal to:

$$\frac{\partial (p_i x_i)}{\partial y} = \frac{\beta_i + \mu \alpha_i}{A_i (\mu + 1)}, \qquad \frac{\partial (p_i x_i)}{\partial p_j} = -A_j a_j \frac{\beta_i + \mu \alpha_i}{A_i (\mu + 1)}$$

and this system (in A, a, μ) has more unknown than equations. However, if $\varepsilon > 0$ then a host of additional, nonlinear terms appear. For instance

$$\frac{\partial (p_i x_i)}{\partial y} = \frac{2A_1 p_1 \left(\mu \alpha_i + \beta_i\right)}{A_i} \left(\frac{\begin{array}{c}\left(A_1 p_1 + Y \varepsilon \gamma_1 + \mu A_1 p_1 + K\right) \\ -Y \left(\varepsilon \gamma_1 + \frac{1}{2K} \left(2Y \varepsilon^2 \gamma_1^2 - 2\varepsilon \gamma_1 A_1 p_1 (\mu - 2\beta_1 + 1)\right)\right)\end{array}}{\left(A_1 p_1 + Y \varepsilon \gamma_1 + \mu A_1 p_1 + K\right)^2} \right).$$

therefore second derivatives like $\frac{\partial^2 (p_i x_i)}{\partial y^2}$ and $\frac{\partial^2 (p_i x_i)}{\partial y \partial p_j}$ are non zero, and are actually highly nonlinear functions of all variables. These corresponding, additional equations pin down the coefficients, and so suffice to over identify the model.

9.2.2 'Generic' Identification: generalization

Here we provide a heuristic argument for why generic nonparametric identification should hold not just for linear consumption technologies $z = Ax + a$ as in Proposition 4, but also for arbitrary technologies, that is, with $z = F(x)$ for general monotonic vector valued functions

F. In other words, our general methodology does not depend on functional form assumptions to obtain identification.

Assume that the consumption technology function F is an unknown smooth function. Let $\rho = p/y$ and let Ω_F be the space of increasing C^2 functions with range and domain equal to the positive orthant. Define a mapping $\mathcal{T}$ by the following procedure. Given any element $F \in \Omega_F$, treat F as if it were the true technology function by defining the corresponding shadow price vector $\overline{\pi}(\rho) = DF(x)'.\rho / [x'DF(x)'.\rho]$ evaluated at $x = F^{-1}h(\rho)$. Now let $\overline{x}(\rho, \eta) = h^m \overline{\pi}(\rho)/(1 - \eta) + h^f \overline{\pi}(\rho)/\eta$, and for every ρ in the positive orthant define $\overline{\eta}(\rho)$ by $\overline{\eta}(\rho) = \arg\min_{\eta^* \in 0,1} \max|\overline{x}(\rho, \eta^*) - F^{-1}h(\rho)|$. Finally, define the function $\widetilde{F}$ by $\widetilde{F}[\overline{x}(\rho, \overline{\eta}(\rho))] = h(\rho)$. This procedure defines a mapping $\widetilde{F} = \mathcal{T}(F)$.

The true technology function, say $F^*(x)$ is a fixed point of this mapping, and the true sharing rule $\eta(\rho)$ equals $\overline{\eta}(\rho)$ evaluated using $F = F^*$. A sufficient condition to ensure the identification of η given identification of F is that there exist an element of the vector $\overline{x}(\rho, \eta)$ that is strictly monotonic in η (and a primitive sufficient condition for this is the existence of a good that is only consumed by one household member, who has a demand function for that good that is strictly monotonic in total expenditure).

$\mathcal{T}$ might not be a contraction mapping, and so may not have a unique fixed point. However, for suitably regular demand functions (those for which the tangent application to $\mathcal{T}$ is Fredholm) each fixed point will be locally unique as a consequence of Smale's (1973) generalized transversality theorem. This provides local generic identification. Ordinary regularity conditions such as monotonicity and domain constraints further restrict the set of feasible solution functions.

Global identification results if only one fixed point is present or if, among all fixed points, only one satisfies the required regularity conditions. This describes formally what we mean by 'generic' identification. A similar case of generic identification appears in Chiappori and Ekeland (1999). Informally, our construction of the mapping $\mathcal{T}$ shows that a sufficient number of demand functions are identified to permit recovery of F and η, so these functions will be identified as long as the demand functions are 'not too simple'. Nonlinear technologies help to identify the model, and so should permit identification even in cases where proposition 4, which assumes a linear technology, fails to provide identification.

This general argument can be illustrated bythe following, simple example. Let us assume there are at least three commodities ($n \geq 3$), and let us take the simplest possible individual preferences, i.e. Cobb Douglas:

$$U^f = \sum_i \alpha_i \log\left(x_i^f\right), U^m = \sum_i \beta_i \log\left(x_i^m\right)$$

Assume the technology takes the (non linear) form:

$$z_i = \left(x_i^m + x_i^f + \gamma_i\right)^{\delta_i} \tag{36}$$

where the δ's and the γ's are parameters to be identified. Then one can show that, with at least three commodities, these parameters are generically identified, even though preferences are Cobb Douglas. However, there exists an exception: if

$$\gamma_1 = ... = \gamma_n = 0$$

(obviously a non generic case), then two parameters (say, δ_1 and δ_2) can be arbitrarily chosen and then the others will be exactly determined. In other words, we do not have full identification in this special case where all the γ_i parameters are zero.

References

[1] BANKS, J., R. BLUNDELL, AND A. LEWBEL, (1997), "Quadratic Engel Curves and Consumer Demand," with James Banks and Richard Blundell, *Review of Economics and Statistics*, 79, 527-539.

[2] BARMBY, T. AND N. SMITH, (2001), "Household labour supply in Britain and Denmark: some interpretation using a model of Pareto optimal behaviour," *Applied Economics*, 33, 1109-1116.

[3] BARGAIN, O., M. BEBLO, D. BENINGER, R. BLUNDELL, R. CARRASCO, M.-C. CHIURI, F. LAISNEY, V. LECHENE, E. LONGOBARDI, M. MYCK, N. MOREAU, J. RUIZ-CASTILLO, AND F. VERMEULEN (2004), "Collective models of household labour supply with non-convex budget sets and non-participation: a calibration approach," *Review of the Economics of the Household*, forthcoming.

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

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

[6] BLACKORBY, C. AND D. DONALDSON (1994), "Measuring Costs of Children: A Theoretical Framework" in The Measurement of Household Behavior, R. Blundell, I. Preston, and I. Walker, Eds., 70-85, Cambridge: Cambridge University Press

[7] BLUNDELL, R. (1988), "Consumer Behavior: Theory and Empirical Evidence - A Survey," *Economic Journal*, 98, 16-65.

[8] BLUNDELL, R. AND A. LEWBEL (1991), "The Information Content of Equivalence Scales," *Journal of Econometrics*, 50, 49-68.

[9] BLUNDELL, R, I. PRESTON, AND I. WALKER (EDS) (1994), The Measurement of Household Welfare, Cambridge: Cambridge University Press.

[10] BOURGUIGNON, F. AND P.-A. CHIAPPORI (1994), "The Collective Approach to Household Behavior" in The Measurement of Household Behavior, R. Blundell, I. Preston, and I. Walker, Eds., 70-85, Cambridge: Cambridge University Press

[11] BROWNING, M. (1992), "Children and Household Economic Behavior," *Journal of Economic Literature*, 30, 1434-1475.

[12] BROWNING, M., F. BOURGUIGNON, P.A. CHIAPPORI, AND V. LECHENE (1994), "Incomes and Outcomes: a Structural Model of Intrahousehold Allocation," *Journal of Political Economy*, 102, 1067-1096.

[13] BROWNING, M., AND P.A. CHIAPPORI, (1998), "Efficient Intra-household Allocations: A General Characterization and Empirical Tests," *Econometrica*, 66, 1241-78.

[14] CHERCHYE, L., B. DE ROCK AND F. VERMEULEN (2012), 'Married with Children: A Collective Labor Supply Model with Detailed Time Use and Intrahousehold Expenditure Information', *American Economic Review*, forthcoming.

[15] CHESHER, A., (1998) "Individual Demands from Household Aggregates: Time and Age Variation in the Quality of Diet," *Journal of Applied Econometrics*, 13, 505-524.

[16] CHIAPPORI, P.A. AND I. EKELAND (1999), "Aggregation and Market Demand: An Exterior Differential Calculus Viewpoint," *Econometrica*, 67, 1435-1457.

[17] CHIAPPORI, P.A. AND I. EKELAND (2005), "Characterizing Group Behavior", unpublished manuscript.

[18] CHIAPPORI, P.-A., FORTIN, B. AND G. LACROIX (2002), "Marriage Market, Divorce Legislation and Household Labor Supply", *Journal of Political Economy*, 110, 37-72

[19] DEATON, A. AND J. MUELLBAUER, (1980a), Economics and Consumer Behavior, Cambridge: Cambridge University Press.

[20] DONNI, O. AND O. BARGAIN, (2011), "Extending the Collective Approach to Children: A Rothbarth-Type Measure of Child Costs Consistent with Scale Economies and Parents Bargaining," THEMA (THéorie Economique, Modélisation et Applications) Working Papers 2011-06, Université de Cergy-Pontoise.

[21] DUFLO, E. (2000), "Child outcomes in Africa. Child health and household resources in South Africa: Evidenced from the old age pension program," *American Economic Association Papers and Proceedings*, 90, 393-398.

[22] DUNBAR, G., A. LEWBEL, AND K. PENDAKUR, (2012), "Children's Resources in Collective Households: Identification, Estimation and an Application to Child Poverty in Malawi," American Economic Review, forthcoming.

[23] GORMAN, T. (1976), "Tricks With Utility Functions," In Essays in Economic Analysis: Proceedings of the 1975 AUTE Conference, Sheffield, M. J. Artis and A. R. Nobay, eds., Cambridge: Cambridge University Press.

[24] GREGORY, A. AND M. VEALL (1985), "Formulating Wald Tests of Nonlinear Restrictions", *Econometrica*, 53, 1465-68.

[25] JORGENSON, D. W. (1997), Welfare Volume 2: Measuring Social Welfare, Cambridge, Massachusetts: MIT Press.

[26] LECOCQ, S. AND J. M. ROBIN (2006), "Estimating Demand Response with Panel Data." *Empirical Economics*, 31, 1043-60.

[27] LEWBEL, A. (1985), "A Unified Approach to Incorporating Demographic or Other Effects into Demand Systems," *Review of Economic Studies,* 52, 1-18.

[28] LEWBEL, A. (1989), "Household Equivalence Scales and Welfare Comparisons," *Journal of Public Economics*, 39, 377-391.

[29] LEWBEL, A. (1997), "Consumer Demand Systems and Household Equivalence Scales," Handbook of Applied Econometrics, Volume II: Microeconomics, M. H. Pesaran and P. Schmidt, eds., Oxford: Blackwell Publishers Ltd.

[30] LEWBEL, A. (2001), "Demand Systems With and Without Errors," *American Economic Review*, 91, 611-618.

[31] LEWBEL, A. (2003a), "Equivalence Scales Based on Collective Household Models," in Household Behavior, Equivalence Scales, Welfare and Poverty; Camilo Dagum and Guido Ferrari, eds., Physica Verlag: New York, pp 1-9.

[32] LEWBEL, A., (2003b), "Calculating Compensation in Cases of Wrongful Death," *Journal of Econometrics*, 113, 115-128.

[33] LEWBEL, A. AND K. PENDAKUR, (2008) "Estimation of Collective Household Models With Engel Curves," *Journal of Econometrics*, 147, 350-358.

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

[35] LISE, J. AND S. SEITZ, (2011), "Consumption Inequality and Intra-Household Allocations," Review of Economic Studies, 78, 328-355.

[36] MUELLBAUER, J. (1977), "Testing the Barten Model of Household Consumption Effects and the Cost of Children," *Economic Journal*, 87, 460-487.

[37] NELSON, J. A. (1988), "Household Economies of Scale in Consumption: Theory and Evidence," .*Econometrica,* 56 1301-1314.

[38] PENDAKUR, K. (1999), "Estimates and Tests of Base-Independent Equivalence Scales", *Journal of Econometrics*, 88, 1-40.

[39] POLLAK, R. A. AND T. J. WALES (1979), "Welfare Comparisons and Equivalence Scales," *American Economic Review*, 69, 216-221.

[40] POLLAK, R. A. AND T. J. WALES (1992), Demand System Specification and Estimation, Oxford: Oxford University Press.

[41] RUBALCAVA, L., AND D. THOMAS (2000), "Family Bargaining and Welfare", Mimeo RAND, UCLA.

[42] SLESNICK, D. (1998), "Empirical Approaches to the Measurement of Welfare" *Journal of Economic Literature*, 36, 2108-2165.

[43] SMALE, S. (1973), "An infinite dimensional version of Sard's theorem," *American Journal of Mathematics*, 87, 213-221.

[44] THOMAS, D. (1990), "Intra-Household Resource Allocation: An Inferential Approach", J*ournal of Human Resources*, 25, 635-664.

[45] THOMAS D, CONTRERAS D, FRANKENBERG E. (1997), "Distribution of Power Within the Household and Child Health." Santa Monica, CA: RAND.

[46] VERMEULEN, F. (2000), "Collective Household Models: Principles and Main Results," Center for Economic Studies, University Leuven, Discussion Paper.
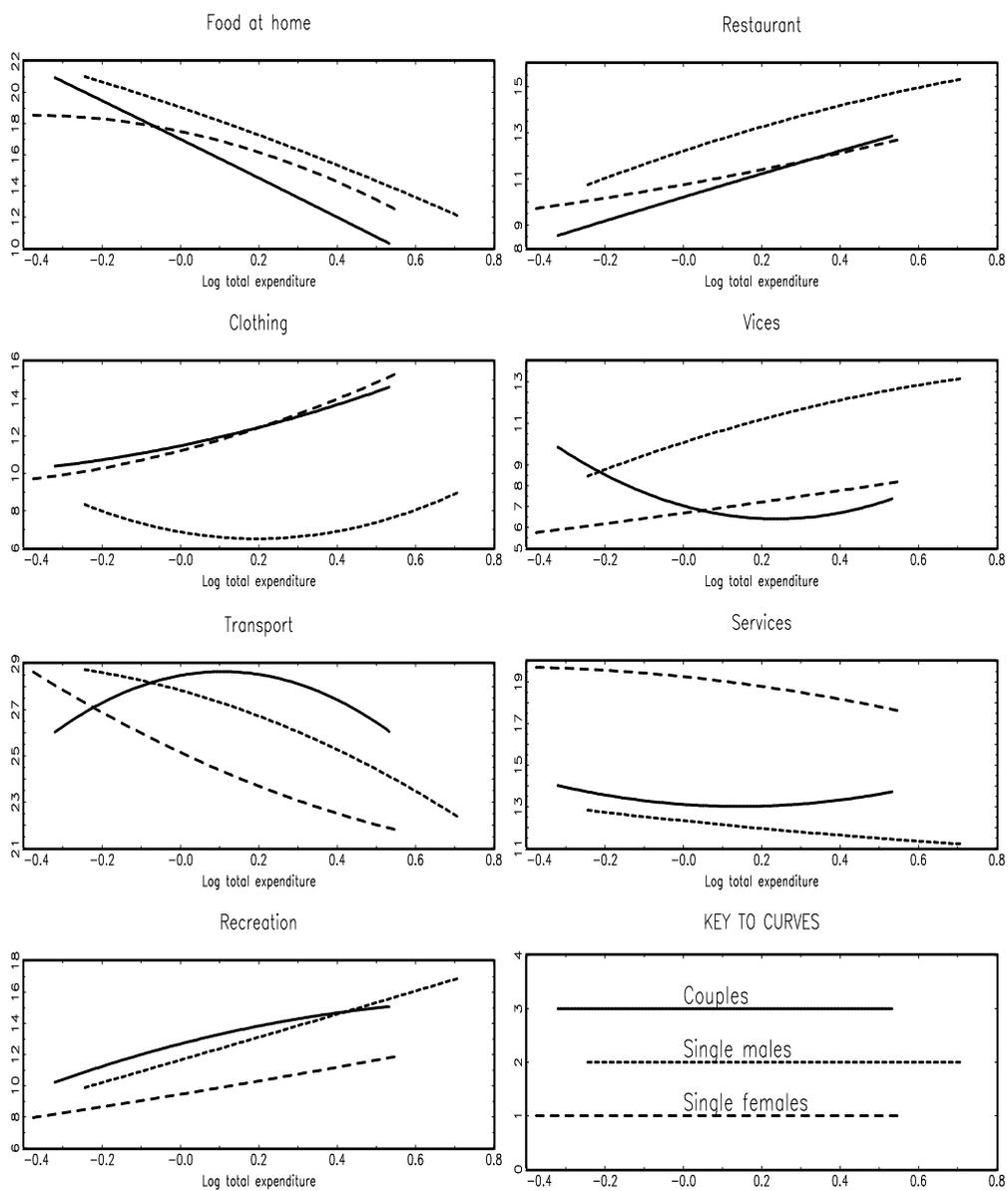
Figure 1: QAIDS fits for singles and couples.