Sharing rule identification for general collective consumption models

Laurens Cherchye[*], Bram De Rock[†], Arthur Lewbel[‡] and Frederic Vermeulen[§]

May, 2012

Abstract

We propose a method to identify bounds (i.e. set identification) on the sharing rule for a general collective household consumption model. Unlike the effects of distribution factors, it is well known that the level of the sharing rule cannot be uniquely identified without strong assumptions on preferences across households of different compositions. Our new results show that, though not point identified without these assumptions, bounds on the sharing rule can still be obtained. We get these bounds by applying revealed preference restrictions implied by the collective model to the household's continuous aggregate demand functions. We obtain informative bounds even if nothing is known about whether each good is public, private, or assignable within the household, though having such information tightens the bounds. An empirical application demonstrates the practical usefulness of our method.

JEL Classification: D11, D12, D13, C14, C30.
Keywords: collective model, consumer demand, revealed preferences, sharing rule, identification, bounds.

[*]Center for Economic Studies, University of Leuven. E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: laurens.cherchye@kuleuven-kortrijk.be. Laurens Cherchye gratefully acknowledges financial support from the Research Fund K.U.Leuven through the grant STRT1/08/004.

[†]ECARES and ECORE, Université Libre de Bruxelles. Avenue F. D. Roosevelt 50, CP 114, B-1050 Brussels, Belgium. E-mail: bderock@ulb.ac.be. Bram De Rock gratefully acknowledges the European Research Council (ERC) for his starting grant.

[‡]Department of Economics, Boston College, 140 Commonwealth Ave., Chestnut Hill, MA, 02467, USA. (617)-552-3678, lewbel@bc.edu, http://www2.bc.edu/~lewbel/.

[§]CentER, Tilburg University. P.O. Box 90153, NL-5000 LE Tilburg, The Netherlands. E-mail: frederic.vermeulen@uvt.nl. Frederic Vermeulen gratefully acknowledges financial support from the Netherlands Organisation for Scientific Research (NWO) through a VIDI grant.

1 Introduction

The collective model has become increasingly popular for analyzing household consumption behavior. Becker (1973, 1981) first considered collective household models, in which the household is characterized as a collection of individuals, each of whom has a well defined objective function, and who interact to generate household level decisions. For consumption, the model assumes that expenditures on each good and service the household buys are the outcome of multi-person decision making, in which each individual household member is characterized by his or her own rational preferences. Following Chiappori (1988, 1992), 'rational' group consumption is defined as any Pareto efficient outcome of a within-group bargaining process. This collective approach contrasts with the conventional unitary approach, which models households as if they were single decision makers.

An intrinsic feature of the collective model is the so-called sharing rule, which governs the within-household distribution of household income. This sharing rule is often interpreted as an indicator of the bargaining power of individual household members. The sharing rule is also useful for recovering information about the economic well being of household members. For example, Lise and Seitz (2011) use sharing rule estimates to recover the population distribution of income across individuals rather than across households, Browning, Chiappori and Lewbel (2006) combine the sharing rule with other information to recover "indifference scales" that measure the welfare implications of changes in household composition, and Dunbar, Lewbel, and Pendakur (2012) use sharing rule estimates to back out rates of child poverty. See, also Browning, Bourguignon, Chiappori and Lechene (1994), Chiappori, Fortin and Lacroix (2002), Blundell, Chiappori and Meghir (2005), Lewbel and Pendakur (2008), Bargain and Donni (2012), and Cherchye, De Rock and Vermeulen (2012a) for various applications of the collective consumption model that make use of the sharing rule concept.

In empirical analyses, the sharing rule is generally not observed. Typically, the only information available is total household expenditures on each good or service the household buys, along with general household characteristics like demographic composition, and information on wages, income, holdings of durables, and wealth measures. Distribution factors are observed household characteristics that affect Pareto weights in a household's model but not the preferences of individual household members. A well known result in this literature is that changes in the sharing rule resulting from changes in distribution factors can be identified given household level demand functions, but the levels of the sharing rules are not themselves identified. See, e.g., Chiappori and Ekeland (2009) for the most general statement and proof of this result.

This nonidentification is unfortunate, because many of the uses of sharing rule estimates, such as calculation of poverty lines, indifference scales, and distributions of income and welfare, all depend on the level of the sharing rules. A few different responses to this nonidentification result have been proposed. The commonest response is to ignore the problem, and only report estimates of the impact of distribution factors on sharing rules. A second approach is to try to collect more information on the consumption of individual household members (see, e.g., Cherchye, De Rock and Vermeulen, 2012a), though this method is inherently limited by the difficulty of measuring the fraction of shared goods that are consumed by each individual. A third response is to make additional identifying assumptions. These assumptions take the form of assuming some features of individuals' preferences remain the same across households of different compositions (first proposed in Browning, Chiappori and Lewbel, 2006).

In this paper, we return to the standard Chiappori framework, where all that can be observed is household level demand functions, and no additional assumptions are made. We show that although sharing rules cannot be point identified, the household's demand functions do provide information regarding sharing, and that information can be used to calculate informative bounds on the sharing rules. In short, we show that sharing rules can be usefully set identified, even though they are not in general point identified.

We propose a practical method for calculating upper and lower bounds on the resource shares of each individual in a household, consistent with the collective consumption model. The method allows for the presence of both public and private goods within the household, and does not require the public or private nature of any good to be specified a priori. However, if a subset of goods is known to be private, then we can use that information to tighten the bounds.

Essentially, our method first adopts a revealed preference approach in the tradition of Samuelson (1938) and Houthakker (1950) to characterize the collective model of household consumption. Sharing rule bounds are obtained by combining the information given by a household's demand function with the restrictions implied by having the unobserved demand functions of individual household members satisfy revealed preference theory, and add up to the household's demand functions. In empirical practice, we apply our revealed preference based restrictions to household demand functions that are estimated using nonparametric regression methods, thereby combining standard household demand estimation techniques with less standard revealed preference restrictions that apply to the collective consumption model.

Another paper that combines estimated demand functions with revealed preference restrictions is Blundell, Browning and Crawford (2008). They assume a unitary rather than collective model of consumption behavior and apply revealed preference restrictions to nonparametrically estimated Engel curves to obtain bounds on demand functions that are consistent with the unitary model. In their case, demands as functions of total expenditures are estimated separately in each of a limited number of price regimes, and revealed preference restrictions (assuming the household behaves as a single utility maximizing consumer) are then imposed to bound demands as functions of prices. In contrast, we nonparametrically estimate household demands as functions of both total expenditures and prices, and then impose revealed preference restrictions at the level of individual household members to obtain bounds on the sharing rule. Their use of estimated Engel curves yields tighter bounds than only applying revealed preference restrictions to observed data points. Similarly, our use of estimated demand functions provides more information than applying collective model revealed preference restrictions just to observed data points.

Many papers propose tests or checks of whether household demands are consistent with the Chiappori model of rational, Pareto efficient group consumption, without actually identifying the sharing rule. Browning and Chiappori (1998) provide a differential characterization of the general collective consumption model.[1] They find that household behavior is consistent with this model only if there exists a household pseudo-Slutsky matrix that can be decomposed as the sum of a symmetric negative semi-definite matrix and a matrix of rank 1 (in the case of two household members). Chiappori and Ekeland (2006) show that this condition, together with homogeneity and adding up, is also (locally) sufficient for the existence of individual utility functions and Pareto weights that reproduce the observed household behavior. Working with discrete sets of price and quantity bundles reflecting households' expenditure choices, Cherchye, De Rock and Vermeulen (2007, 2010, 2011) derive revealed preference characterizations of the general version of the collective model that we consider here, in the Afriat (1967), Diewert (1973), and Varian (1982) tradition. In particular, Cherchye, De Rock and Vermeulen (2011) show how information regarding the sharing rule can be recovered using this pure revealed preference characterization, but only with discrete sets of price and quantity bundles. In contrast, the present paper exploits the greater information that is available in continuous demand functions.

The rest of this paper is organized as follows. Section 2 introduces the general collective consumption model and the corresponding sharing rule representation. Section 3 discusses revealed preference restrictions. Section 4 considers sharing rule identification. Section 5 introduces extensions to settings where the private consumption of some goods is assignable to individual household members. Section 6 presents an empirical application that demonstrates the practical usefulness of our identification method for various purposes. Section 7 concludes.

2 The collective model and the sharing rule

This section formally presents the collective model, and introduces the sharing rule representation that will be used in the following sections.

2.1 A general collective model

We consider a household with two individuals (1 and 2) who consume a set of goods $N = \{1, ..., |N|\}$.[2] We assume a household demand function $\mathbf{g}$ that defines a quantity bundle $\mathbf{q} \in \mathbb{R}_{+}^{|N|}$ as a function of prices $\mathbf{p} \in \mathbb{R}_{++}^{|N|}$ and income $y \in \mathbb{R}_{++}$, i.e. $\mathbf{q} = \mathbf{g}(\mathbf{p}, y)$

[1] The term 'differential' refers to the fact that the characterization is obtained by differentiating the functional specifications of the fundamentals of the model (e.g. the utility functions or demands of the household members) as in the calculation of Slutsky matrices.

[2] This choice is for expositional convenience. All results can be generalized towards households with any number of household members.

for

$$\mathbf{g} : (\mathbf{p}, y) \rightarrow \mathbf{g}(\mathbf{p}, y).$$

Throughout, we will focus on the 'general' version of the collective model discussed by Browning and Chiappori (1998) and Chiappori and Ekeland (2006, 2009). This model accounts for public consumption and private consumption of any good:

$$\mathbf{q} = \mathbf{q}^1 + \mathbf{q}^2 + \mathbf{q}^H \text{ with } \mathbf{q}^c \in \mathbb{R}_+^{|N|} \ (c = 1, 2, H),$$

with $\mathbf{q}^1$ and $\mathbf{q}^2$ the privately consumed quantities of individuals 1 and 2, and $\mathbf{q}^H$ the publicly consumed quantities. Next, it considers general (concave) utility functions U^1 and U^2 that also account for externalities for the privately consumed goods:

$$U^1\left(\mathbf{q}^1, \mathbf{q}^2, \mathbf{q}^H\right) \text{ and } U^2\left(\mathbf{q}^1, \mathbf{q}^2, \mathbf{q}^H\right).$$

Finally, the model assumes a Pareto efficient intrahousehold allocation, i.e. there exists a Pareto weight $\mu \in \mathbb{R}_{++}$ such that $\mathbf{q} = \mathbf{q}^1 + \mathbf{q}^2 + \mathbf{q}^H = \mathbf{g}(\mathbf{p}, y)$ for

$$\begin{aligned}\left(\mathbf{q}^1, \mathbf{q}^2, \mathbf{q}^H\right) = \arg \max_{\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^H} [U^1\left(\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^H\right) + \mu U^2\left(\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^H\right) \text{ s.t.} \\ \mathbf{p}'\left(\mathbf{x}^1 + \mathbf{x}^2 + \mathbf{x}^H\right) \leq y, \ \mathbf{x}^c \in \mathbb{R}_+^{|N|} \ (c = 1, 2, H)].\end{aligned} \tag{1}$$

The Pareto weight μ represents the 'bargaining power' of member 2 relative to member 1. Generally, μ can vary with prices, household income and other exogenous variables that affect household decisions but not the preferences or the household budget (i.e. so-called extra-environmental parameters in the terminology of McElroy, 1990, or distribution factors in the terminology of Browning, Bourguignon, Chiappori, and Lechene, 1994, and Bourguignon, Browning and Chiappori, 2009).

2.2 Sharing rule representation

We now focus on a sharing rule representation of demand behavior that is consistent with the model introduced above.[3] Intuitively, this representation implies a decentralized interpretation of Pareto efficient household consumption behavior as defined in (1). Specifically, it represents Pareto efficient household behavior *as if* it is the outcome of a two-step allocation procedure. In the first step, the so-called *sharing rule* distributes the aggregate household income across the group members, which defines individual income shares y^1 and y^2 (so that $y = y^1 + y^2$). In the second step, each individual maximizes her/his utility subject to the resulting income share and accounting for her/his Lindahl prices associated with privately and publicly consumed quantities;

[3]See Chiappori (1988, 1992) for a detailed discussion on the equivalence between the characterizations of Pareto efficient consumption behavior in (1) and (2). Chiappori concentrated on a simplified setting with privately consumed quantities without externalities. However, extending his argument to our setting is fairly straightforward.

the Lindahl prices then represent the individual's marginal willingness to pay for the different quantities. Of course, we are not assuming that households explicitly use the sharing rule. The two-step representation simply states that the outcome of the group allocation process can be characterized in this way.

To formalize the sharing rule representation, for each individual m $(m = 1, 2)$ we consider an individual demand function $\mathbf{g}^m$ that defines a quantity bundle $\widetilde{\mathbf{q}}^m \in \mathbb{R}_+^{|N|}$ as a function of individual prices $\mathbf{p}^{m,1} \in \mathbb{R}_+^{|N|}$, $\mathbf{p}^{m,2} \in \mathbb{R}_+^{|N|}$, $\mathbf{p}^{m,H} \in \mathbb{R}_+^{|N|}$ and individual income $y^m \in \mathbb{R}_{++}$. Specifically, $\widetilde{\mathbf{q}}^m = \mathbf{g}^m(\mathbf{p}^{m,1}, \mathbf{p}^{m,2}, \mathbf{p}^{m,H}, y^m)$ for

$$\mathbf{g}^m : \left(\mathbf{p}^{m,1}, \mathbf{p}^{m,2}, \mathbf{p}^{m,H}, y^m\right) \rightarrow \mathbf{g}^m \left(\mathbf{p}^{m,1}, \mathbf{p}^{m,2}, \mathbf{p}^{m,H}, y^m\right),$$

such that $\widetilde{\mathbf{q}}^m = \widetilde{\mathbf{q}}^{m,1} + \widetilde{\mathbf{q}}^{m,2} + \widetilde{\mathbf{q}}^{m,H}$ with

$$\begin{aligned}(\widetilde{\mathbf{q}}^{m,1}, \widetilde{\mathbf{q}}^{m,2}, \widetilde{\mathbf{q}}^{m,H}) = \arg \max_{\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^H} [U^m \left(\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^H\right) \text{ s.t.} \\ \left(\mathbf{p}^{m,1}\right)' \mathbf{x}^1 + \left(\mathbf{p}^{m,2}\right)' \mathbf{x}^2 + \left(\mathbf{p}^{m,H}\right)' \mathbf{x}^H \leq y^m, \mathbf{x}^c \in \mathbb{R}_+^{|N|} \ (c = 1, 2, H)].\end{aligned} \tag{2}$$

Consider $\mathbf{q}^1$, $\mathbf{q}^2$ and $\mathbf{q}^H$ (with $\mathbf{q} = \mathbf{q}^1 + \mathbf{q}^2 + \mathbf{q}^H$) that solve (1). The same quantities solve (2) (and thus $\mathbf{q} = \widetilde{\mathbf{q}}^m$) if

$$\mathbf{p}^{m,c} = U^m_{\mathbf{q}^c} / \lambda^m, \tag{3}$$

for $U^m_{\mathbf{q}^c}$ the gradient of the function U^m defined at $\mathbf{q}^c$ $(c = 1, 2, H)$, λ^1 the Lagrange multiplier (in the optimum) for the budget constraint in (1), and $\lambda^2 = \lambda^1/\mu$. In words, each vector $\mathbf{p}^{m,c} \in \mathbb{R}_+^{|N|}$ represents individual m's marginal willingness to pay for $\mathbf{q}^c$.

Importantly, Pareto efficiency implies $\mathbf{p}^{1,c} + \mathbf{p}^{2,c} = \mathbf{p}$ $(c = 1, 2, H)$ by construction. Hence, we can interpret $\mathbf{p}^{m,c}$ as the Lindahl price vector for individual m associated with $\mathbf{q}^c$. Thus, under these prices the maximization program (2) corresponds to the second step of the two-step procedure described above (for given y^1 and y^2 defined in the first step).

The sharing rule is the crucial concept underlying the characterization in (2). In the literature on the collective consumption model, this sharing rule is often interpreted as an indicator for the bargaining power of the individual group members: a higher relative income share of member m (y^m/y) is then regarded as an indication of increased bargaining power for that member.[4] The sharing rule concept is particularly useful in applications, because it is independent of cardinal representations of preferences (in contrast to the bargaining weight μ).

3 Revealed preferences

From an empirical point of view, we typically only observe the household demand function $\mathbf{g}$ and income y and not the individual demand functions $\mathbf{g}^m$ or income shares

[4]See, for example, Browning, Chiappori and Lewbel (2006) for a detailed discussion on the relation between income shares y^m in (2) and the bargaining weight μ in (1).

y^m $(m = 1, 2)$. Our primary goal is then to identify the sharing rule that defines the individual incomes y^1 and y^2 given this observed data. In Section 4, we will introduce our method of identifying bounds on the sharing rule, starting from a revealed preference characterization of the above defined collective consumption model. We provide this revealed preference characterization in the current section.

3.1 Basic concepts

As a preliminary step, we define some concepts that pertain to an individual demand function $\mathbf{g}^m$. For given $\mathbf{p}^{m,1}, \mathbf{p}^{m,2}, \mathbf{p}^{m,H}$ and y^m, we define the budget set

$$B\left(\mathbf{p}^{m,1}, \mathbf{p}^{m,2}, \mathbf{p}^{m,H}, y^m\right) = \{\mathbf{x} \in \mathbb{R}_+^{|N|} | \mathbf{x} = \mathbf{x}^1 + \mathbf{x}^2 + \mathbf{x}^H, \\ \left(\mathbf{p}^{m,1}\right)' \mathbf{x}^1 + \left(\mathbf{p}^{m,2}\right)' \mathbf{x}^2 + \left(\mathbf{p}^{m,H}\right)' \mathbf{x}^H \leq y^m\}.$$

We can now define the following concept.

Definition 1 (direct revealed preference) *Let $\mathbf{g}^m$ be an individual demand function. The direct revealed preference relation associated with $\mathbf{g}^m$ is defined by: for all $\mathbf{x}$, $\mathbf{z} \in \mathbb{R}_+^{|N|}$: $\mathbf{x} R_o^{\mathbf{g}^m}$ $\mathbf{z}$ if there exist $\mathbf{p}^{m,c} \in \mathbb{R}_+^{|N|}$ $(c = 1, 2, H)$ and $y^m \in \mathbb{R}_+$ such that $\mathbf{x} = \mathbf{g}^m\left(\mathbf{p}^{m,1}, \mathbf{p}^{m,2}, \mathbf{p}^{m,H}, y^m\right)$ and $\mathbf{z} \in B(\mathbf{p}^{m,1}, \mathbf{p}^{m,2}, \mathbf{p}^{m,H}, y^m)$ with $\mathbf{x} \neq \mathbf{z}$.*

In words, quantity bundle $\mathbf{x}$ is revealed preferred to another bundle $\mathbf{z}$ if $\mathbf{z}$ belonged to the budget set $B(\mathbf{p}^{m,1}, \mathbf{p}^{m,2}, \mathbf{p}^{m,H}, y^m)$ under which $\mathbf{x}$ was chosen (i.e. $\mathbf{g}^m(\mathbf{p}^{m,1}, \mathbf{p}^{m,2}, \mathbf{p}^{m,H}, y^m)$). At this point, it is important to distinguish the revealed preference concept in Definition 1 from the more standard one that applies to the unitary consumption model. Specifically, in our collective setting revealed preferences are defined at the level of an individual household member m, while in a unitary context they are defined at the level of the aggregate household. Correspondingly, we consider preferences that pertain to decomposed quantity bundles, which are evaluated at individual prices ($\mathbf{p}^{m,1}$, $\mathbf{p}^{m,2}$ and $\mathbf{p}^{m,H}$ for individual m). The associated quantity decomposition appears from the definition of the budget set $B\left(\mathbf{p}^{m,1}, \mathbf{p}^{m,2}, \mathbf{p}^{m,H}, y^m\right)$.

We can now define the Weak Axiom of Revealed Preference (WARP; after Samuelson, 1938). Like before, in the present context this WARP concept is defined for an individual m.

Definition 2 (WARP) *Let $\mathbf{g}^m$ be an individual demand function. This function $\mathbf{g}^m$ satisfies WARP if the relation $R_o^{\mathbf{g}^m}$ is asymmetric.*

Our sharing rule identification method will exploit the empirical implications of WARP in the context of the collective household consumption model. At this point, it is worth indicating that Houthakker (1950) actually presented a strengthened version of WARP, i.e. the Strong Axiom of Revealed Preference (SARP). Essentially, SARP extends WARP by also exploiting transitivity of preferences. By construction, a demand function satisfies SARP only if it satisfies WARP. In view of our following discussion, an interesting question is whether the converse is also true, i.e. a demand function satisfies SARP if it satisfies WARP. It can be verified that this does not hold in general (see Gale, 1960, and Kihlstrom, Mas-Colell and Sonnenschein, 1976). However, Uzawa (1960, 1971) derived minimal regularity conditions for the function $\mathbf{g}^m$ that make the WARP condition equivalent to the SARP condition. See also Bossert (1993) for discussion.

In conclusion, our focus on WARP means that we do not consider implications of transitive preferences captured by SARP. Our motivation is that it is difficult to operationalize transitivity when implementing the procedure that we outline in Sections 4 and 5. The implication is that we do not fully exploit all the information that could be used and that, in principle, even tighter bounds might be obtained when transitivity can be accounted for. However, the loss of information is zero for individual demand functions $\mathbf{g}^m$ that satisfy Uzawa's regularity conditions mentioned above.

3.2 Characterizing the collective model

So far, we have used individual demand functions $\mathbf{g}^m$, which are typically not known in empirical analysis (i.e. we only know the household demand function $\mathbf{g}$). Therefore, to develop an approach for sharing rule identification that can be used in practice, we consider the concept of 'admissible' individual demand functions. Essentially, this concept captures all possible specifications of the (unknown) individual demand functions that are consistent with the (known) household demand function. More formally, we use the following definition.

Definition 3 (admissible individual demands) *Let* $\mathbf{g}$ *be a household demand function. For this function* $\mathbf{g}$*, the individual demand functions* $\mathbf{g}^1$ *and* $\mathbf{g}^2$ *are admissible individual demand functions if, for all* $\mathbf{p}$ *and* y*,*

$$\mathbf{g}(\mathbf{p}, y) = \mathbf{g}^1\left(\mathbf{p}^{1,1}, \mathbf{p}^{1,2}, \mathbf{p}^{1,H}, y^1\right) = \mathbf{g}^2\left(\mathbf{p}^{2,1}, \mathbf{p}^{2,2}, \mathbf{p}^{2,H}, y^2\right),$$

for some $\mathbf{p}^{m,c}\,(m = 1, 2;\ c = 1, 2, H)$ *and* y^m *such that*

$$y^1 + y^2 = y \text{ and } \mathbf{p}^{1,c} + \mathbf{p}^{2,c} = \mathbf{p}, \text{ with } y^m \in \mathbb{R}_{++} \text{ and } \mathbf{p}^{m,c} \in \mathbb{R}_{+}^{|N|}.$$

Next, $Q(\mathbf{g})$ *represents the collection of all admissible individual demands* $\mathbf{g}^1$ *and* $\mathbf{g}^2$*, i.e.*

$$Q(\mathbf{g}) = \left\{\left(\mathbf{g}^1, \mathbf{g}^2\right) | \mathbf{g}^1 \text{ and } \mathbf{g}^2 \text{ are admissible individual demand functions}\right\}.$$

We can now define the condition for a collective rationalization that we will use. Basically, this condition states that (for the given function $\mathbf{g}$) there must exist at least one specification of admissible individual demand functions that solves (2).

Definition 4 (collective rationalization) *Let* $\mathbf{g}$ *be a household demand function. A pair of utility functions* U^1 *and* U^2 *provides a collective rationalization of* $\mathbf{g}$ *if there exist admissible individual demand functions* $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$ *such that, for each* m,

$$\mathbf{g}^m \left(\mathbf{p}^{m,1}, \mathbf{p}^{m,2}, \mathbf{p}^{m,H}, y^m\right) = \mathbf{q}^1 + \mathbf{q}^2 + \mathbf{q}^H$$

for

$$(\mathbf{q}^1, \mathbf{q}^2, \mathbf{q}^H) = \arg \max_{\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^H} [U^m \left(\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^H\right) \;\; s.t.$$
$$\left(\mathbf{p}^{m,1}\right)' \mathbf{x}^1 + \left(\mathbf{p}^{m,2}\right)' \mathbf{x}^2 + \left(\mathbf{p}^{m,H}\right)' \mathbf{x}^H \leq y^m].$$

We have the following result, which establishes a revealed preference characterization of the collective consumption model under consideration. (Appendix 1 contains the proofs of our main results.)

Proposition 1 *Consider a household demand function* $\mathbf{g}$. *There exists a pair of utility functions* U^1 *and* U^2 *that provides a collective rationalization of* $\mathbf{g}$ *only if there exist admissible individual demand functions* $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$ *that both satisfy WARP.*

In the next section, we will show that this WARP condition provides a useful starting point to develop a practical method for sharing rule identification.

4 Sharing rule identification

Consider a household demand function $\mathbf{g}$. We focus on identifying the sharing rule for prices $\mathbf{p}_E$ and household income y_E, for which we observe $\mathbf{g}(\mathbf{p}_E, y_E)$. The (set) identification question asks for bounds on the individual incomes (y_E^1 and y_E^2) that are consistent with a collective rationalization of the observed household demand $\mathbf{g}$. Our procedure will start from the characterization given in Proposition 1. Essentially, it defines lower bounds y_E^{l1} and y_E^{l2} and upper bounds y_E^{u1} and y_E^{u2} so that

$$y_E^{l1} < y_E^1 < y_E^{u1} \text{ and } y_E^{l2} < y_E^2 < y_E^{u2}. \tag{4}$$

These bounds will be independent of the specification of the admissible individual demand functions.

4.1 Identification in theory

To sketch the basic idea of our approach, we first suppose the individual demand functions $\mathbf{g}^1$ and $\mathbf{g}^2$ are given (in addition to the household demand $\mathbf{g}$). We remark that in this case we would also know the individual prices $\mathbf{p}^{m,c}$ ($m = 1, 2;\ c = 1, 2, H$) and income y^m that imply $\mathbf{g}(\mathbf{p}, y) = \mathbf{g}^m(\mathbf{p}^{m,1}, \mathbf{p}^{m,2}, \mathbf{p}^{m,c}, y^m)$ for each $\mathbf{p}$ and y. See Definition 3.

Let $\mathbf{q}_E = \mathbf{g}(\mathbf{p}_E, y_E) = \mathbf{g}^m(\mathbf{p}_E^{m,1},\ \mathbf{p}_E^{m,2},\ \mathbf{p}_E^{m,c},\ y_E)$. Then, $\mathbf{g}^1$ and $\mathbf{g}^2$ are consistent with a collective rationalization of $\mathbf{g}$ only if

$$\begin{aligned} y_E^1 &< \min_{\mathbf{x}_1^1,\mathbf{x}_1^2,\mathbf{x}_1^H} [\left(\mathbf{p}_E^{1,1}\right)' \mathbf{x}_1^1 + \left(\mathbf{p}_E^{1,2}\right)' \mathbf{x}_1^2 + \left(\mathbf{p}_E^{1,H}\right)' \mathbf{x}_1^H | \mathbf{x}_1 R_o^{\mathbf{g}^1} \mathbf{q}_E] \text{ and} \\ y_E^2 &< \min_{\mathbf{x}_2^1,\mathbf{x}_2^2,\mathbf{x}_2^H} [\left(\mathbf{p}_E^{2,1}\right)' \mathbf{x}_2^1 + \left(\mathbf{p}_E^{2,2}\right)' \mathbf{x}_2^2 + \left(\mathbf{p}_E^{2,H}\right)' \mathbf{x}_2^H | \mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{q}_E]. \end{aligned} \tag{5}$$

This necessary condition for a collective rationalization of $\mathbf{g}$ directly follows from the WARP conditions in Proposition 1. The right hand sides of the inequalities in (5) define upper bounds y_E^{u1} and y_E^{u2} for the income shares y_E^1 and y_E^2, i.e. $y_E^1 < y_E^{u1}$ and $y_E^2 < y_E^{u2}$. Further, we have that $y_E^{l1} = y_E - y_E^{u2}$ and $y_E^{l2} = y_E - y_E^{u1}$.

In empirical applications, we do not know the functions $\mathbf{g}^1$ and $\mathbf{g}^2$. Therefore, our method concentrates on all admissible individual demands $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$. The tightest bounds will be obtained if the (positive) differences $y_E^{u1} - y_E^{l1}$ and $y_E^{u2} - y_E^{l2}$ are as small as possible. Substituting $y_E - y_E^{u2}$ for y_E^{l1} in $y_E^{u1} - y_E^{l1}$ (or substituting $y_E - y_E^{u1}$ for y_E^{l2} in $y_E^{u2} - y_E^{l2}$) obtains that these sharpest bounds correspond to the smallest value of $y_E^{u1} + y_E^{u2} - y_E$. Since y_E is a constant, we will ultimately aim at minimizing the sum $y_E^{u1} + y_E^{u2}$ in what follows. (In this respect, also observe that minimizing $(y_E^{u1} + y_E^{u2})$ is equivalent to maximizing $(y_E^{l1} + y_E^{l2})$.)

Summarizing, the upper bounds y_E^{u1} and y_E^{u2} need to solve

$$\max_{(\mathbf{g}^1,\mathbf{g}^2)\in Q(\mathbf{g})} \min_{\substack{\mathbf{x}_1^1,\mathbf{x}_1^2,\mathbf{x}_1^H \\ \mathbf{x}_2^1,\mathbf{x}_2^2,\mathbf{x}_2^H}} (y_E^{u1} + y_E^{u2}) \tag{P.0}$$

s.t.

$$\begin{aligned} y_E^{u1} &= \left(\mathbf{p}_E^{1,1}\right)' \mathbf{x}_1^1 + \left(\mathbf{p}_E^{1,2}\right)' \mathbf{x}_1^2 + \left(\mathbf{p}_E^{1,H}\right)' \mathbf{x}_1^H, \\ y_E^{u2} &= \left(\mathbf{p}_E^{2,1}\right)' \mathbf{x}_2^1 + \left(\mathbf{p}_E^{2,2}\right)' \mathbf{x}_2^2 + \left(\mathbf{p}_E^{2,H}\right)' \mathbf{x}_2^H, \\ &\mathbf{x}_1 R_o^{\mathbf{g}^1} \mathbf{q}_E, \\ &\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{q}_E. \end{aligned}$$

The max operator in the objective makes that the upper bounds y_E^{u1} and y_E^{u2} apply to any possible specification of $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$. Corresponding lower bounds are defined as $y_E^{l1} = y_E - y_E^{u2}$ and $y_E^{l2} = y_E - y_E^{u1}$. It directly follows that a collective rationalization of the data is possible only for individual income shares y_E^1 and y_E^2 that meet (4).

One final remark is in order. It follows from our discussion preceding program P.0 that, if the solution value of program P.0 does not exceed y_E, it is impossible to specify income shares y_E^1 and y_E^2 that meet (4). The interpretation is that, in such a case, the demand function $\mathbf{g}$ cannot be collectively rationalized. Or, in other words, we conclude that the collective model is rejected for the function $\mathbf{g}$ at hand. Analogous results apply to the programs that we present below, which will be relevant for our application in Section 6.

4.2 Identification in practice

The above program P.0 is not directly useful in practice, because it requires considering infinitely many combinations of the individual demands $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$. In this section, we present a program that has practical usefulness, because it allows for sharing rule identification by solely using information on the household demand $\mathbf{g}$.

A starting result characterizes the bundles $\mathbf{x}_1$ and $\mathbf{x}_2$ that satisfy $\mathbf{x}_1 R_o^{\mathbf{g}^1} \mathbf{q}_E$ and $\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{q}_E$ in program P.0 in terms of the household demand function $\mathbf{g}$.

Proposition 2 *Let* $\mathbf{g}$ *be a household demand function. Then, we have* $\mathbf{x}_1 R_o^{\mathbf{g}^m} \mathbf{q}_E$ *and* $\mathbf{x}_2 R_o^{\mathbf{g}^l} \mathbf{q}_E$ *(*$m, l \in \{1,2\}$*,* $m \neq l$*) for all admissible individual demand functions* $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$ *that satisfy WARP if* $\mathbf{x}_1 = \mathbf{g}(\mathbf{p}_1, y_1)$ *and* $\mathbf{x}_2 = \mathbf{g}(\mathbf{p}_2, y_2)$ *such that*

$$y_1 \geq \mathbf{p}_1' (\mathbf{q}_E + \mathbf{x}_2) \text{ and } y_2 \geq \mathbf{p}_2' (\mathbf{q}_E + \mathbf{x}_1). \tag{C}$$

Thus, as soon as condition C holds, we conclude that $\mathbf{x}_1 R_o^{\mathbf{g}^m} \mathbf{q}_E$ and $\mathbf{x}_2 R_o^{\mathbf{g}^l} \mathbf{q}_E$. This result makes it possible to compute the upper bounds y_E^{u1} and y_E^{u2} (and corresponding lower bounds $y_E^{l1} = y_E - y_E^{u2}$ and $y_E^{l2} = y_E - y_E^{u1}$) through the following programming problem:

$$\min_{\mathbf{p}_1, \mathbf{p}_2, y_1, y_2} (y_E^{u1} + y_E^{u2}) \tag{P.1}$$

s.t.

$$y_E^{u1} = \mathbf{p}_E' \mathbf{x}_1,\ y_E^{u2} = \mathbf{p}_E' \mathbf{x}_2, \tag{P.1-1}$$
$$y_1 \geq \mathbf{p}_1' (\mathbf{q}_E + \mathbf{x}_2),\ y_2 \geq \mathbf{p}_2' (\mathbf{q}_E + \mathbf{x}_1), \tag{P.1-2}$$
$$\mathbf{x}_1 = \mathbf{g}(\mathbf{p}_1, y_1),\ \mathbf{x}_2 = \mathbf{g}(\mathbf{p}_2, y_2). \tag{P.1-3}$$

The explanation is as follows. Like in program P.0, the objective minimizes the sum $(y_E^{u1} + y_E^{u2})$ by suitable selecting $\mathbf{x}_1$ and $\mathbf{x}_2$ (defined by $\mathbf{p}_1$, $\mathbf{p}_2$, y_1 and y_2; see (P.1-3)): a lower objective function value corresponds to tighter sharing rule bounds. Next, because of Proposition 2, the constraint (P.1-2) implies $\mathbf{x}_1 R_o^{\mathbf{g}^1} \mathbf{q}_E$ and $\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{q}_E$ or $\mathbf{x}_1 R_o^{\mathbf{g}^2} \mathbf{q}_E$ and $\mathbf{x}_2 R_o^{\mathbf{g}^1} \mathbf{q}_E$. Without loss of generality, we assume $\mathbf{x}_1 R_o^{\mathbf{g}^1} \mathbf{q}_E$ and $\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{q}_E$. Because $\mathbf{x}_m R_o^{\mathbf{g}^m} \mathbf{q}_E$, we need

$$y_E^m < \left(\mathbf{p}_E^{m,1}\right)' \mathbf{x}_m^1 + \left(\mathbf{p}_E^{m,2}\right)' \mathbf{x}_m^2 + \left(\mathbf{p}_E^{m,H}\right)' \mathbf{x}_m^H; \tag{6}$$

compare with (5) above. By construction, we have $\left(\mathbf{p}_E^{m,1}\right)'\mathbf{x}_m^1 + \left(\mathbf{p}_E^{m,2}\right)'\mathbf{x}_m^2 + \left(\mathbf{p}_E^{m,H}\right)'\mathbf{x}_m^H \leq \mathbf{p}_E'\mathbf{x}_m$ for any specification of $\mathbf{p}_E^{m,c}$ and $\mathbf{x}_m^c$ $(c = 1, 2, H)$. We avoid a specification of $\mathbf{p}_E^{m,c}$ and $\mathbf{x}_m^c$ in our practical approach developed here by using the upper bound $y_E^{um} = \mathbf{p}_E'\mathbf{x}_m$ in (P.1-1). This parallels the fact that we do not consider a particular specification of the individual functions $\mathbf{g}^m$ in the program P.1.

Importantly, the objective as well as most constraints in program P.1 are linear. The only nonlinear constraint is (P.1-3) (for a nonlinear function $\mathbf{g}$). This implies that the bounds y_E^{u1}, y_E^{u2}, y_E^{l1}, y_E^{l2} are to be computed through nonlinear programming techniques.

In Appendix 2, we present a stylized example that demonstrates the mechanics of program P.1. Interestingly, this example shows that the method discussed in this section can obtain arbitrarily tight bounds on the individual incomes for the general version of the collective model, even if no good is specified as public or private a priori. This is a noteworthy result, as it stands in sharp contrast with the finding of Chiappori and Ekeland (2009), who conclude that the differential approach does not allow for such sharing rule identification (see our discussion in the introductory section).

We conclude from the example in Appendix 2 that, in principle, the method described here may work well for sharing rule identification. In practice, however, additional information on the private nature of goods may help to strengthen the analysis. For example, this will be the case for the empirical application that we present in Section 6. In the next section, we show how to extend our method to include such information on private goods.

5 Extensions

In empirical applications, it is often reasonable to assume that a subset of goods is privately consumed without externalities, while the nature of the other goods is unknown. In fact, applications of the collective consumption model usually include assignability of particular goods to individual household members. Such goods are then called exclusive goods, because they exclusively benefit the utility of single household members; see Bourguignon, Browning and Chiappori (2009). This section considers including such information on privately consumed goods in our method for sharing rule identification. First, we will reformulate program P.1 to account for goods that are privately consumed without externalities. Subsequently, we will consider assignability of such goods to individual members.

5.1 Private goods without externalities

Formally, let N_A be the subset of private goods without externalities and N_B the subset of other goods, so that $N = N_A \cup N_B$. For a good $n \in N_A$, we get the following extension of the collective rationalization condition in Definition 4 (for $(\mathbf{a})_n$ the nth entry of vector **a**):

$$\left(\mathbf{p}^{1,2}\right)_n = 0 \text{ (or } \left(\mathbf{p}^{1,1}\right)_n = (\mathbf{p})_n \text{) and } \left(\mathbf{p}^{2,1}\right)_n = 0 \text{ (or } \left(\mathbf{p}^{2,2}\right)_n = (\mathbf{p})_n \text{)} \qquad (7)$$

Intuitively, because there are no consumption externalities for good n, the willingness to pay of household member m for member l's consumption is zero.[5]

Using (7), we can reformulate program P.1 as follows:

$$\min_{\mathbf{p}_1,\mathbf{p}_2,y_1,y_2} (y_E^{u1} + y_E^{u2}) \qquad \text{(P.2)}$$

s.t.

$$y_E^{u1} = \sum\nolimits_{n\in N_A} (\mathbf{p}_E)_n \left(\mathbf{x}_1^1\right)_n + \sum\nolimits_{n\in N_B} (\mathbf{p}_E)_n (\mathbf{x}_1)_n\,, \qquad \text{(P.2-1)}$$

$$y_E^{u2} = \sum\nolimits_{n\in N_A} (\mathbf{p}_E)_n \left(\mathbf{x}_2^2\right)_n + \sum\nolimits_{n\in N_B} (\mathbf{p}_E)_n (\mathbf{x}_2)_n\,, \qquad \text{(P.2-2)}$$

$$y_1 \geq \mathbf{p}_1' (\mathbf{q}_E + \mathbf{x}_2)\,,\ y_2 \geq \mathbf{p}_2' (\mathbf{q}_E + \mathbf{x}_1)\,, \qquad \text{(P.2-3)}$$

$$(\mathbf{x}_k)_n = \left(\mathbf{x}_k^1\right)_n + \left(\mathbf{x}_k^2\right)_n \quad (k = 1, 2;\ n \in N_A)\,, \qquad \text{(P.2-4)}$$

$$\mathbf{x}_1 = \mathbf{g}(\mathbf{p}_1, y_1)\,,\ \mathbf{x}_2 = \mathbf{g}(\mathbf{p}_2, y_2)\,.$$

Similar to before, the constraint (P.2-3) implies $\mathbf{x}_1 R_o^{\mathbf{g}^1} \mathbf{q}_E$ and $\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{q}_E$. Thus, we again get the condition (6). In this case, we have

$$\left(\mathbf{p}_E^{m,m}\right)_n (\mathbf{x}_m^m)_n = (\mathbf{p}_E)_n (\mathbf{x}_m^m)_n \text{ and } \left(\mathbf{p}_E^{m,l}\right)_n (\mathbf{x}_m^m)_n = 0\ (m \neq l) \text{ if } n \in N_A,$$

while

$$\left(\mathbf{p}_E^{m,1}\right)_n \left(\mathbf{x}_m^1\right)_n + \left(\mathbf{p}_E^{m,2}\right)_n \left(\mathbf{x}_m^2\right)_n + \left(\mathbf{p}_E^{m,H}\right)_n \left(\mathbf{x}_m^H\right)_n \leq (\mathbf{p}_E)_n (\mathbf{x}_m)_n \text{ if } n \in N_B.$$

Therefore, we can use

$$\left(\mathbf{p}_E^{m,1}\right)' \mathbf{x}_m^1 + \left(\mathbf{p}_E^{m,2}\right)' \mathbf{x}_m^2 + \left(\mathbf{p}_E^{m,H}\right)' \mathbf{x}_m^H \leq \sum\nolimits_{n\in N_A} (\mathbf{p}_E)_n (\mathbf{x}_m^m)_n + \sum\nolimits_{n\in N_B} (\mathbf{p}_E)_n (\mathbf{x}_m)_n\,,$$

which obtains (P.2-1) and (P.2-2) instead of (P.1-1) before. We note that for $n \in N_A$ the privately consumed quantities $(\mathbf{x}_k^1)_n$ and $(\mathbf{x}_k^2)_n$ are not given a priori and therefore defined within program P.2 (subject to the constraint (P.2-4)). Like before, program P.2 is solved by nonlinear programming techniques.

[5] For compactness, we do not include a formal argument showing (7). But the result follows quite directly from our definition of $\left(\mathbf{p}^{m,c}\right)_n$ in (3).

5.2 Exclusive goods

So far, we have abstracted from assignability of the goods $n \in N_A$ to individual household members. To account for such assignability, let $N_{Am} \subseteq N_A$ represent the set of goods that are assignable (or exclusive) to member m. Then, we get

$$(\mathbf{x}_k^m)_n = (\mathbf{x}_k)_n \text{ if } n \in N_{Am}.$$

Using this, we obtain the following extension of Proposition 2.

Proposition 3 *Let* $\mathbf{g}$ *be a household demand function. Then, we have* $\mathbf{x}_1 R_o^{\mathbf{g}^1} \mathbf{q}_E$ *and* $\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{q}_E$ *for all admissible individual demand functions* $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$ *that satisfy WARP if* $\mathbf{x}_1 = \mathbf{g}(\mathbf{p}_1, y_1)$ *and* $\mathbf{x}_2 = \mathbf{g}(\mathbf{p}_2, y_2)$ *such that one of the following conditions holds:*

$$\begin{aligned} y_1 &\geq \mathbf{p}_1' (\mathbf{q}_E + \mathbf{x}_2) \text{ and} \\ \sum\nolimits_{n \in N_{A_2}} (\mathbf{p}_2)_n (\mathbf{x}_2)_n &\geq \mathbf{p}_2' \mathbf{x} - \sum\nolimits_{n \in N_{A_1}} (\mathbf{p}_2)_n (\mathbf{x})_n \text{ for } \mathbf{x} = \mathbf{x}_1, \mathbf{q}_E, \end{aligned} \tag{C.1}$$

$$\begin{aligned} y_2 &\geq \mathbf{p}_2' (\mathbf{q}_E + \mathbf{x}_1) \text{ and} \\ \sum\nolimits_{n \in N_{A_1}} (\mathbf{p}_1)_n (\mathbf{x}_1)_n &\geq \mathbf{p}_1' \mathbf{q}_E - \sum\nolimits_{n \in N_{A_2}} (\mathbf{p}_1)_n (\mathbf{x})_n \text{ for } \mathbf{x} = \mathbf{x}_2, \mathbf{q}_E, \end{aligned} \tag{C.2}$$

or

$$\begin{aligned} \sum\nolimits_{n \in N_{A_1}} (\mathbf{p}_1)_n (\mathbf{x}_1)_n &\geq \mathbf{p}_1' \mathbf{q}_E - \sum\nolimits_{n \in N_{A_2}} (\mathbf{p}_1)_n (\mathbf{q}_E)_n \text{ and} \\ \sum\nolimits_{n \in N_{A_2}} (\mathbf{p}_2)_n (\mathbf{x}_2)_n &\geq \mathbf{p}_2' \mathbf{q}_E - \sum\nolimits_{n \in N_{A_1}} (\mathbf{p}_2)_n (\mathbf{q}_E)_n . \end{aligned} \tag{C.3}$$

It is interesting to compare this result to the one in Proposition 2. The essential difference is that, in contrast to condition C, the new conditions C.1, C.2 and C.3 'assign' preference relations to individual household members (i.e. we get $\mathbf{x}_1 R_o^{\mathbf{g}^1} \mathbf{q}_E$ and $\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{q}_E$). This assigning of preference relations is possible because we can use information on assignable goods.

Because of Proposition 3, we must consider three nonlinear programs (in addition to program P.2) to define y_E^{u1}, y_E^{u2}, y_E^{l1}, y_E^{l2}. Each program has the same structure as P.2, except that the condition C (in the constraint (P.2-3)) is replaced by one of the conditions C.1-C.3. 'Best' upper (resp. lower) bounds then correspond to minimum (resp. maximum) values defined over the different programs.

6 Application

We apply our methods to a labor supply setting. Specifically, at observed individual wage rates we consider the allocation of a household's full income (the sum of both spouses' maximum possible labor income and the household's non-labor income) to both spouses' leisure and consumption. Household consumption is here treated as an aggregate Hicksian commodity with price normalized at unity. This setting contains substantial price (i.e. wage) variation, which is useful for obtaining informative sharing rule bounds. Also, this application allows us to consider various assumptions regarding the nature of the different goods. For example, we can see how much the bounds tighten if we treat each individual's leisure as an exclusive private good, or if we treat the aggregate Hicksian commodity as a private good without externalities. Lise and Seitz (2011) similarly use labor supply to identify resource shares, but for identification their results depend on strong functional form assumptions, as well as restrictions across households like those in Browning, Chiappori, and Lewbel (2006).

Some implementations of collective household models treat wages as distribution factors, thereby assuming they only affect sharing rules and not preferences. In our application, wages are prices (of leisure). Our methodology provides sharing rule bounds without requiring the presence of any distribution factors.

6.1 Set-up

We apply our sharing rule identification method to a sample of Dutch households drawn from the 2009 wave of the LISS (Longitudinal Internet Studies for the Social sciences) panel that is gathered by CentERdata. This survey, which is representative for the Dutch population, contains a rich variety of economic and socio-demographic variables.[6] The set of households used for this study was subject to the following sample selection rules. First we selected couples where both adult members participate in the labor market. We include both couples with and without children.[7] Next, we excluded the self-employed to avoid issues regarding imputation of wages and the separation of consumption from work related expenditures. After deleting the households with important missing information (mostly, incomplete information on one of the spouses), we obtained a sample of 211 observations. This sample is rather small for our purposes, but as we show below, it still contains enough information to yield meaningful results.

Table 1 provides summary statistics on the relevant price and quantity data for the sample at hand. Wages are net hourly wages. Leisure is measured in hours per week. To compute leisure hours we assume an individual needs 10 hours per day for sleeping and eating (i.e. leisure = 168 - 70 - hours worked). Full income and consumption are measured in euros per week. For completeness, Table 1 (last column) also reports on the number of children in the households under consideration. Table 1 reveals considerable variation in relative prices (wages) and full income, which is what allows

[6] Households without any Internet access are provided with a basic computer (a 'SimPC') that enables them to connect to the Internet and thereby participate in the survey.

[7] Unlike Dunbar, Lewbel and Pendakur (2012), Cherchye, De Rock and Vermeulen (2012a), or Bargain and Donni (2012), we do not attempt to separately identify resource shares devoted to specifically children. Instead, we implicitly assume that expenditures on children are internalized in the parents' preferences through individual or public consumption.

		Wage		Leisure			
	Full income	Female	Male	Female	Male	Consumption	Children
Mean	2852.00	13.20	15.43	70.48	57.62	1028.74	1.05
St.dev.	1418.83	9.63	9.44	9.23	6.78	517.06	1.04
Max.	15677.82	138.89	128.97	88.00	78.00	5872.15	4.00
Min.	1209.03	4.06	5.08	38.00	38.00	325.04	0.00

Table 1: Descriptive statistics; prices and quantities for the LISS sample

us to obtain reasonably tight bounds despite our relatively small sample size.

A crucial ingredient to our identification method is the household's vector-valued demand function with respect to both spouses' leisure and total household consumption (i.e. the function $\mathbf{g}$ in the earlier sections), which we need to estimate. To avoid specification errors, we use our sample to estimate this function nonparametrically. Specifically, we define the first two elements of $\mathbf{g}$ to be the fitted values of nonparametrically regressing both spouses' leisure on their wage rates and on the household's full income, using a Nadaraya-Watson kernel estimator with a multivariate Gaussian kernel. The remaining element of $\mathbf{g}$, the household's consumption demand function, is then obtained by the adding-up restriction. Our entire model, including the nonparametric regression step, was coded in MATLAB, using the TOMLAB/SNOPT code to solve our nonlinear programs.

We first experimented with three different general model specifications. The first was the most general specification of the collective model, which makes no prior assumption regarding the public or private nature of any good, thereby allowing any of our three goods (i.e. the aggregate Hicksian commodity and both spouses' leisure) to be private (with or without externalities), public, or both. We found that this specification did not provide useful bounds on income shares. Since we know from our theory that this most general model can yield informative bounds, our empirical results with this specification are likely due to the limited size of our available data set.

At the other extreme, we also considered the model specification in which all three goods are assumed to be privately consumed without externalities, and that male and female leisure are exclusive (assignable), and so do not generate externalities within the household. Essentially, this corresponds to the original 'egoistic' labor supply model of Chiappori (1988). Here, we found that this collective model is systematically rejected, implying that a purely egoistic model cannot rationalize our data.[8]

Given the above results, we did all of our remaining analyses assuming an in-between specification where male and female leisure are assumed to be private and assignable, generating no externalities, while no restriction is placed on the nature of the Hicksian aggregate commodity. This allows consumption to be private (with or

[8] Specifically, we conclude that the collective model is rejected if our estimated function $\mathbf{g}$ cannot be collectively rationalized, meaning that the inequalities associated with the model cannot be satisfied for any values of utility or sharing rule. See our remark at the end of Section 4.1 for details.

without externalities), public, or both.

6.2 Empirical results

6.2.1 RP based sharing rule bounds versus atheoretic bounds

As a first exercise, we compare the bounds on female income shares (male shares are one minus the female shares) that are obtained by our revealed preference (RP) methodology with 'atheoretic' bounds. These atheoretic bounds do not make use of the (theoretical) RP restrictions associated with the collective consumption model and are defined as follows: the lower bound for a female in a particular household equals the share of the value of her leisure in this household's full income; the corresponding upper bound adds the share of household consumption in the household's full income to this lower bound. In other words, the atheoretic lower (upper) bound corresponds to an (extreme) scenario where all the household's consumption is allocated to the male (female).

The results of this exercise are summarized in Table 2. The table shows that our RP based bounds provide a substantial gain in precision compared to the atheoretic bounds. The average difference between the upper and lower atheoretic bounds is about 36 percentage points, while this difference equals only about 11.5 percentage points for the RP based bounds. The median difference between the upper and lower RP based bounds is about 5 percentage points, whereas this difference is substantially larger (i.e. about 36 percentage points) for the atheoretic bounds. Qualitatively similar results are obtained for the other quantiles reported in the table.[9]

	RP based bounds	Atheoretic bounds
Mean	11.53	36.32
Minimum	0.00	22.08
First quartile	2.15	32.09
Median	5.15	36.18
Third quartile	17.02	39.73
Maximum	53.26	53.90

Table 2: Percentage point differences between upper and lower bounds on the female income share

6.2.2 The relation between sharing rule bounds, total income, and relative wages

We next focus on the relation between our RP based bounds and some observed individual and household characteristics. Figure 1 shows the relation between the absolute

[9] Interestingly, the smallest difference between the RP based bounds is about zero, which means that our set identification is actually very close to point identification.

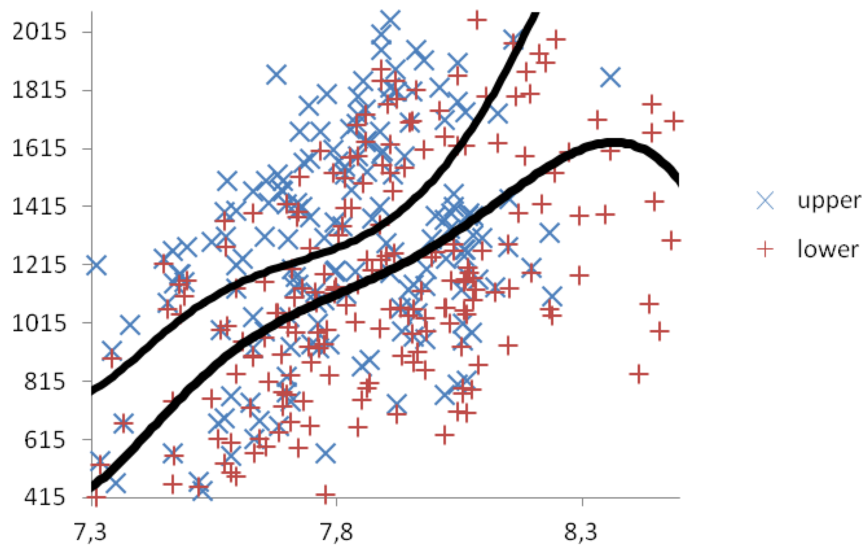

Figure 1: Absolute sharing rule bounds (Y-axis) and the logarithm of full income (X-axis)

RP based bounds on the female income share and the logarithm of the household's full income. Each X and + sign on the figure represents the upper and lower bound for a given household in our sample. To better visualize our results, we included trendlines showing local sample averages of the estimated upper and lower bounds.

Figure 1 shows that the average bounds are fairly tight, reflecting the results in Table 2. The trendlines are upward sloping, showing that female income share is a normal good (as opposed to an inferior good) for the household. They are also roughly 45 degrees, showing that the female's share of household income is not far from proportional to total household income.

Figure 2 confirms the rough proportionality of female income to total household income, by plotting the relative share of the female (her income share as a fraction of full income) against the household's full income. The trendlines give an average upper bound hovering around 60% and average lower bound around 40%. This finding lends empirical support to the assumption that relative income shares do not vary with the logarithm of total income, which Lewbel and Pendakur (2008) and Bargain and Donni (2012) used to help point identify resource shares.[10] Figure 2 further suggests that the income shares of females and males are not far from equal on average. This is confirmed by the average lower and upper bound for all the households in the sample, which equal

[10]Dunbar, Lewbel and Pendakur (2012) used a weaker version of the same assumption, by requiring it to hold (only) at a range of low income levels. Here, we must also add that these authors define publicness of goods differently than we do. Next, their empirical applications focus on a different definition of full income, since they do not include leisure in their consumption models. See also Menon, Pendakur, and Perali (2012) for some direct estimates of intrahousehold resource shares.

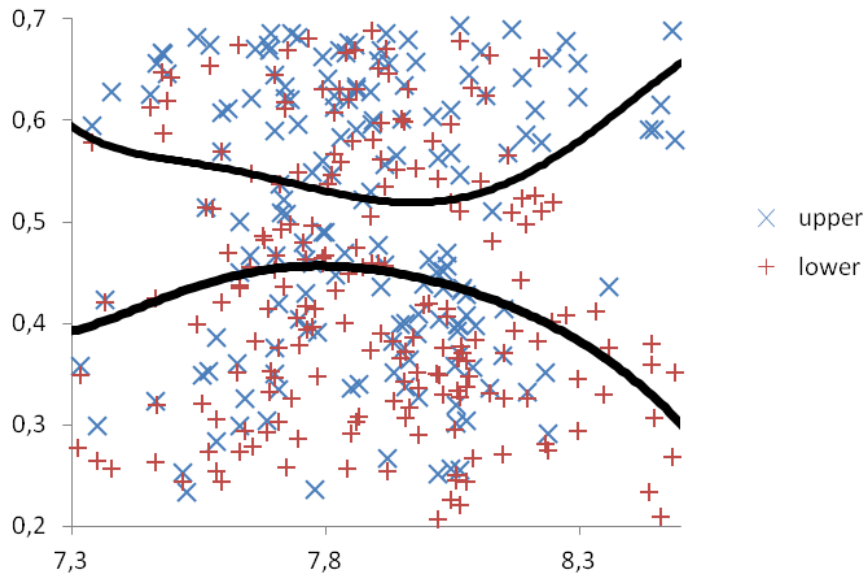

Figure 2: Relative sharing rule bounds (Y-axis) and the logarithm of full income (X-axis)

43.2% and 54.7%, respectively. However, the figure also shows that some households divide income very unequally, e.g., one household has upper and lower bounds of the relative female income share equal to 22.1% and 25.6%, while in another household these bounds are 76.3% and 83.1%.[11]

We next look at the relationship between the bounds on the relative female income share and the relative wage (defined as female wage divided by male wage). Figure 3 shows the household specific upper and lower bounds and the corresponding trendlines. In line with our prior expectations, both bounds clearly increase when the relative wage of females goes up. This result, which we obtained through our robust RP based approach, confirms earlier evidence found in the literature, which shows that a household member's bargaining power generally increases with her/his wage (see Chiappori, Fortin and Lacroix, 2002, Blundell, Chiappori, Magnac and Meghir, 2007, and Oreffice, 2011, among many others).

6.2.3 Poverty analysis

A unique advantage of models that focus on the intrahousehold allocation of resources is that they allow one to conduct welfare analyses at the level of individuals rather

[11] We tried investigating whether variation in bounds correlated with other observable characteristics such as age, education, number of children, and nonlabor income, but we did not find any systematic relationship between these variables and income shares. This might reflect the limits of our relatively small number of households.

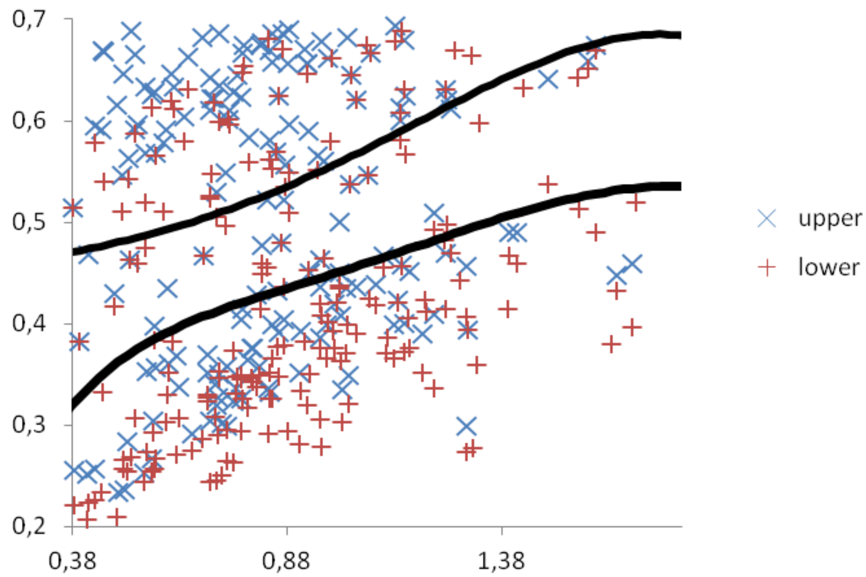

Figure 3: Relative sharing rule bounds (Y-axis) and the relative wage (wage female/wage male) (X-axis)

than at the level of households (see Chiappori, 1992, Blundell, Chiappori and Meghir, 2005, and Browning, Chiappori and Lewbel, 2006, for further discussion, and Lise and Seitz, 2011, and Cherchye, De Rock and Vermeulen, 2012b, for a few examples).

In what follows, we conduct a poverty analysis at the level of individuals by means of our RP based bounds. Unlike previous studies that are based on point identified sharing rules, we do not require any assumptions regarding similarity of preferences across individuals or on functional forms for our analysis.[12] Our analysis is based entirely on the separate observed choice behavior of each couple, without any assumption regarding the functional form of preferences or the sharing rule.

Table 3 summarizes the results of our poverty analysis. The first column contains the poverty rate among unitary households. This poverty rate is calculated in the usual way: it is the percentage of households of which the full income falls below the poverty line, which we define as 60% of the median full income in our sample of households. Note that while 60% of median income is a standard measure of relative poverty, in our case the poverty rate is calculated on the basis of full income instead of (the more commonly considered) earnings or total expenditures. Also, we restrict attention to couples where both spouses participate to the labor market. Such households should be less subject to poverty than households containing an unemployed, retired or disabled spouse.

[12]For example, Browning, Chiappori and Lewbel (2006) and Lise and Seitz (2011) assume similarity of preferences of (fe)male singles and preferences of (fe)male individuals in couples, while Dunbar, Lewbel and Pendakur (2012) assume restrictions upon individual preferences.

The second column of Table 3 shows the incidence of poverty at the level of individuals in our sample. Here our RP based income shares come into play. Similar to before, an individual is labelled as poor if her/his income share falls below the corresponding individual poverty line. The results in Table 3 use half the households' poverty line as the individual poverty line. Based on our income share bounds calculations, we can compute upper and lower bounds for the individual poverty rates. If all couples split income perfectly equally, then these poverty rates would equal those of column 1. However, despite our earlier finding that many couples appear to have close to equal splits, we obtain lower and upper poverty rate bounds of 8.53% and 21.33%, respectively, compared to the unitary household rate of 4.27%.

These bounds indicate that, due to unequal sharing of resources within households, the fraction of individuals living below the poverty line is two to five times greater than those obtained by standard measures that ignore intrahousehold allocations. When we focus on females and males separately, we see that the lower bound on the poverty rate is a bit lower for males than for females, showing that households tend to devote somewhat more resources towards males. Based on our previous figures, this difference is not surprisingly related to males tending to have higher wages.

	Households	All individuals	Females	Males
Household poverty rate	4.27%	-	-	-
Lower bound	-	8.53%	9.48%	7.58%
Upper bound	-	21.33%	21.33%	21.33%

Table 3: Poverty rates

7 Conclusion

It has long been known that, under the standard Pareto efficient collective household model, the income sharing rule is not identified. Past responses to this result have been to focus on features of the model that are identified (like the impacts of distribution factors), or to add additional strong identifying assumptions on preferences and behavior. In contrast, we first show at a theoretical level that, given just household level demand functions, bounds on the sharing rule can be obtained. Moreover, informative bounds are possible even when nothing is known about the privateness or assignability of the goods being consumed by household members, and when no distribution factors are observed. We also show how these bounds can be implemented using standard programming methods, with household level demand functions that are estimated by standard nonparametric regression methods.

At the practical level, we demonstrate that our identification methods are empirically tractable, yielding meaningfully narrow bounds when applied to a small data set of Dutch households. These bounds enable analyses of the effects of household characteristics like income and relative wages on income shares, and a distributional analysis of the incidence of poverty at the level of individuals rather than at the level of observed households.

In our analyses, the household demand function needed to be estimated, and we have not explicitly taken estimation errors in these functions into account. In our empirical application the sample size (both in number of households and number of observations per household) is relatively small, which would in any case limit the applicability of asymptotic distribution theory. Still, in principle it might be possible to base inference on set identification methods like Manski (2003), Chernozhukov, Hong and Tamer (2007), Beresteanu and Molinari (2008), or Galichon and Henry (2009). Note however that unlike most of the results in this literature to date, in our case we are estimating bounds on nonparametrically estimated functions rather than on parameters of finitely parameterized models. This suggests a useful direction for future research. Similar issues arise in other applications that combine nonparametric or semiparametric estimation with revealed preference restrictions, such as Blundell, Browning and Crawford (2008) and Blundell, Kristensen and Matzkin (2011).

For simplicity, our analysis was based on households with two members, implicitly treating expenditures on children as consumption yielding utility for the parents. However, our methods immediately extend to handle more than two consumers per household, and therefore could (given a larger sample) be used to estimate bounds on children's resource shares as well, treating children as additional consumers with their own utility functions and Pareto weights in the collective model.[13]

Appendix 1: proofs

Proof of Proposition 1

The result readily follows from adapting the original reasoning of Samuelson (1938) to our particular collective setting. Specifically, Definition 4 states that a pair of utility functions U^1 and U^2 provides a collective rationalization of $\mathbf{g}$ if there exist admissible individual demand functions $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$ such that, for each m,

$$\mathbf{g}^m\left(\mathbf{p}^{m,1}, \mathbf{p}^{m,2}, \mathbf{p}^{m,H}, y^m\right) = \mathbf{q}^1 + \mathbf{q}^2 + \mathbf{q}^H$$

for

$$\begin{aligned}(\mathbf{q}^1, \mathbf{q}^2, \mathbf{q}^H) = \arg\max_{\mathbf{x}^1,\mathbf{x}^2,\mathbf{x}^H} [U^m\left(\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^H\right) \text{ s.t.}\\ \left(\mathbf{p}^{m,1}\right)' \mathbf{x}^1 + \left(\mathbf{p}^{m,2}\right)' \mathbf{x}^2 + \left(\mathbf{p}^{m,H}\right)' \mathbf{x}^H \leq y^m].\end{aligned}$$

Thus, for each individual m there must exist a utility function U^m such that the function $\mathbf{g}^m$ solves the corresponding maximization problem for any prices $\mathbf{p}^{m,1}$, $\mathbf{p}^{m,2}$,

[13] Note that since the prices of leisure (wages) are not observed for children, utility for children would need to be based only on consumption of goods.

$\mathbf{p}^{m,H}$ and income y^m. Samuelson's (1938) argument obtains that this is possible only if the function $\mathbf{g}^m$ satisfies the WARP condition in Definition 2.

Proof of Proposition 2

As a first step, we prove (for all $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$)

$$(y_1 \geq \mathbf{p}_1' \mathbf{x}_2 \text{ and } y_2 \geq \mathbf{p}_2' \mathbf{x}_1) \Rightarrow \left(\mathbf{x}_1 R_o^{\mathbf{g}^m} \mathbf{x}_2 \text{ and } \mathbf{x}_2 R_o^{\mathbf{g}^l} \mathbf{x}_1 \ (l \neq m)\right). \tag{8}$$

To obtain this result, we note that $y_1 \geq \mathbf{p}_1' \mathbf{x}_2$ implies by construction

$$\begin{aligned} & \sum\nolimits_{m=1}^{2} \left(\mathbf{p}_1^{m,1}\right)' \mathbf{x}_1^1 + \left(\mathbf{p}_1^{m,2}\right)' \mathbf{x}_1^2 + \left(\mathbf{p}_1^{m,H}\right)' \mathbf{x}_1^H \\ \geq \ & \sum\nolimits_{m=1}^{2} \left(\mathbf{p}_1^{m,1}\right)' \mathbf{x}_2^1 + \left(\mathbf{p}_1^{m,2}\right)' \mathbf{x}_2^2 + \left(\mathbf{p}_1^{m,H}\right)' \mathbf{x}_2^H \end{aligned} \tag{9}$$

for all possible specifications of $\mathbf{p}_1^{m,c}$, $\mathbf{x}_1^c$ and $\mathbf{x}_2^c$ $(m = 1, 2;\ c = 1, 2, 3)$. The inequality (9) necessarily obtains

$$\left(\mathbf{p}_1^{m,1}\right)' \mathbf{x}_1^1 + \left(\mathbf{p}_1^{m,2}\right)' \mathbf{x}_1^2 + \left(\mathbf{p}_1^{m,H}\right)' \mathbf{x}_1^H \geq \left(\mathbf{p}_1^{m,1}\right)' \mathbf{x}_2^1 + \left(\mathbf{p}_1^{m,2}\right)' \mathbf{x}_2^2 + \left(\mathbf{p}_1^{m,H}\right)' \mathbf{x}_2^H$$

for $m = 1$ *or* 2, which can also be expressed as $\mathbf{x}_1 R_o^{\mathbf{g}^1} \mathbf{x}_2$ *or* $\mathbf{x}_1 R_o^{\mathbf{g}^2} \mathbf{x}_2$ for all $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$.

Now, without loss of generality, let us assume $\mathbf{x}_1 R_o^{\mathbf{g}^1} \mathbf{x}_2$ (i.e. $m = 1$ in (8)). Then, because the functions $(\mathbf{g}^1, \mathbf{g}^2)$ satisfy WARP, we must have

$$\left(\mathbf{p}_2^{1,1}\right)' \mathbf{x}_1^1 + \left(\mathbf{p}_2^{1,2}\right)' \mathbf{x}_1^2 + \left(\mathbf{p}_2^{1,H}\right)' \mathbf{x}_1^H > \left(\mathbf{p}_2^{1,1}\right)' \mathbf{x}_2^1 + \left(\mathbf{p}_2^{1,2}\right)' \mathbf{x}_2^2 + \left(\mathbf{p}_2^{1,H}\right)' \mathbf{x}_2^H. \tag{10}$$

In turn, because $y_2 \geq \mathbf{p}_2' \mathbf{x}_1$, this implies

$$\left(\mathbf{p}_2^{2,1}\right)' \mathbf{x}_1^1 + \left(\mathbf{p}_2^{2,2}\right)' \mathbf{x}_1^2 + \left(\mathbf{p}_2^{2,H}\right)' \mathbf{x}_1^H \leq \left(\mathbf{p}_2^{2,1}\right)' \mathbf{x}_2^1 + \left(\mathbf{p}_2^{2,2}\right)' \mathbf{x}_2^2 + \left(\mathbf{p}_2^{2,H}\right)' \mathbf{x}_2^H,$$

or $\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{x}_1$. This proves (8).

As a second step, we show (for all $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$)

$$\left(\mathbf{x}_1 R_o^{\mathbf{g}^m} \mathbf{x}_2 \text{ and } y_2 \geq \mathbf{p}_2' (\mathbf{q}_E + \mathbf{x}_1)\right) \Rightarrow \left(\mathbf{x}_2 R_o^{\mathbf{g}^l} \mathbf{q}_E \ (l \neq m)\right). \tag{11}$$

To prove this result, we first observe that $y_2 \geq \mathbf{p}_2' (\mathbf{q}_E + \mathbf{x}_1)$ implies

$$\begin{aligned} & \sum\nolimits_{m=1}^{2} \left(\mathbf{p}_2^{m,1}\right)' \mathbf{x}_2^1 + \left(\mathbf{p}_2^{m,2}\right)' \mathbf{x}_2^2 + \left(\mathbf{p}_2^{m,H}\right)' \mathbf{x}_2^H \\ \geq \ & \sum\nolimits_{m=1}^{2} \left(\mathbf{p}_2^{m,1}\right)' \left(\mathbf{q}_E^1 + \mathbf{x}_1^1\right) + \left(\mathbf{p}_2^{m,2}\right)' \left(\mathbf{q}_E^2 + \mathbf{x}_1^2\right) + \left(\mathbf{p}_2^{m,H}\right)' \left(\mathbf{q}_E^H + \mathbf{x}_1^H\right), \end{aligned} \tag{12}$$

for all possible specifications of $\mathbf{p}_1^{m,c}$, $\mathbf{x}_1^c$, $\mathbf{x}_2^c$ and $\mathbf{q}_E^c$ $(m = 1, 2;\ c = 1, 2, H)$.

Without loss of generality, let us now assume $\mathbf{x}_1 R_o^{\mathbf{g}^1} \mathbf{x}_2$ (i.e. $m = 1$ in (11)). Like before, WARP consistency then requires (10), and combining this last inequality with (12) yields

$$\left(\mathbf{p}_2^{2,1}\right)' \mathbf{x}_2^1 + \left(\mathbf{p}_2^{2,2}\right)' \mathbf{x}_2^2 + \left(\mathbf{p}_2^{2,H}\right)' \mathbf{x}_2^H > \left(\mathbf{p}_2^{2,1}\right)' \mathbf{q}_E^1 + \left(\mathbf{p}_2^{2,2}\right)' \mathbf{q}_E^2 + \left(\mathbf{p}_2^{2,H}\right)' \mathbf{q}_E^H,$$

or $\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{q}_E$. This proves (11).

Now, a directly similar reasoning as the one leading up to (11) yields

$$\left(\mathbf{x}_2 R_o^{\mathbf{g}^l} \mathbf{x}_1 \text{ and } y_1 \geq \mathbf{p}_1' \left(\mathbf{q}_E + \mathbf{x}_2\right)\right) \Rightarrow \left(\mathbf{x}_1 R_o^{\mathbf{g}^m} \mathbf{q}_E \ (m \neq l)\right). \tag{13}$$

Combining (8), (11) and (13) gives the wanted result: we have $\mathbf{x}_1 R_o^{\mathbf{g}^m} \mathbf{q}_E$ and $\mathbf{x}_2 R_o^{\mathbf{g}^l} \mathbf{q}_E$ for all admissible individual demand functions $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$ that satisfy WARP if $y_1 \geq \mathbf{p}_1' (\mathbf{q}_E + \mathbf{x}_2)$ and $y_2 \geq \mathbf{p}_2' (\mathbf{q}_E + \mathbf{x}_1)$. ■

Proof of Proposition 3

In what follows, we only give the proof for condition C.1. The arguments for the remaining conditions C.2 and C.3 are readily analogous.

As a first step, we prove that (for $\mathbf{x} = \mathbf{x}_1, \mathbf{q}_E$ and for all $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$)

$$\begin{aligned} \sum\nolimits_{n \in N_{A_2}} (\mathbf{p}_2)_n (\mathbf{x}_2)_n &\geq \mathbf{p}_2' \mathbf{x} - \sum\nolimits_{n \in N_{A_1}} (\mathbf{p}_2)_n (\mathbf{x})_n \\ &\Rightarrow \left(\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{x}_1 \text{ and } \mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{q}_E\right) \end{aligned} \tag{14}$$

To obtain this result, we note that

$$\left(\mathbf{p}_2^{2,1}\right)' \mathbf{x}_2^1 + \left(\mathbf{p}_2^{2,2}\right)' \mathbf{x}_2^2 + \left(\mathbf{p}_2^{2,H}\right)' \mathbf{x}_2^H \geq \sum\nolimits_{n \in N_{A_2}} (\mathbf{p}_2)_n (\mathbf{x}_2)_n \tag{15}$$

for any possible specification of $\mathbf{p}_2^{2,c}$ and $\mathbf{x}_2^c$ $(c = 1, 2, H)$. Indeed, $(\mathbf{x}_k^2)_n = (\mathbf{x}_k)_n$ and $\left(\mathbf{p}_2^{2,2}\right)_n = (\mathbf{p}_2)_n$ for the assignable goods $n \in N_{A_2}$, which yields

$$\left(\mathbf{p}_2^{2,2}\right)' \mathbf{x}_2^2 \geq \sum\nolimits_{n \in N_{A_2}} (\mathbf{p}_2)_n (\mathbf{x}_2)_n ,$$

while $\left(\mathbf{p}_2^{2,1}\right)' \mathbf{x}_2^1 + \left(\mathbf{p}_2^{2,H}\right)' \mathbf{x}_2^H \geq 0$ by construction.

Similarly, we have (for $\mathbf{x} = \mathbf{x}_1, \mathbf{q}_E$)

$$\mathbf{p}_2'\mathbf{x} - \sum\nolimits_{n \in N_{A_1}} (\mathbf{p}_2)_n (\mathbf{x})_n \geq \left(\mathbf{p}_2^{2,1}\right)' \mathbf{x}^1 + \left(\mathbf{p}_2^{2,2}\right)' \mathbf{x}^2 + \left(\mathbf{p}_2^{2,H}\right)' \mathbf{x}^H. \tag{16}$$

To see this, we can use an analogous argument as before to get

$$\left(\mathbf{p}_2^{1,1}\right)' \mathbf{x}^1 + \left(\mathbf{p}_2^{1,2}\right)' \mathbf{x}^2 + \left(\mathbf{p}_2^{1,H}\right)' \mathbf{x}^H \geq \sum\nolimits_{n \in N_{A1}} (\mathbf{p}_2)_n (\mathbf{x}_1)_n ,$$

and thus

$$\begin{aligned} & \mathbf{p}_2'\mathbf{x} - \sum\nolimits_{n \in N_{A_1}} (\mathbf{p}_2)_n (\mathbf{x})_n \\ \geq\; & \mathbf{p}_2\mathbf{x} - \left[\left(\mathbf{p}_2^{1,1}\right)' \mathbf{x}^1 + \left(\mathbf{p}_2^{1,2}\right)' \mathbf{x}^2 + \left(\mathbf{p}_2^{1,H}\right)' \mathbf{x}^H\right] \\ =\; & \left(\mathbf{p}_2^{2,1}\right)' \mathbf{x}^1 + \left(\mathbf{p}_2^{2,2}\right)' \mathbf{x}^2 + \left(\mathbf{p}_2^{2,H}\right)' \mathbf{x}^H. \end{aligned}$$

Combining (15) and (16) gives (for $\mathbf{x} = \mathbf{x}_1, \mathbf{q}_E$)

$$\left(\mathbf{p}_2^{2,1}\right)' \mathbf{x}_2^1 + \left(\mathbf{p}_2^{2,2}\right)' \mathbf{x}_2^2 + \left(\mathbf{p}_2^{2,H}\right)' \mathbf{x}_2^H \geq \left(\mathbf{p}_2^{2,1}\right)' \mathbf{x}^1 + \left(\mathbf{p}_2^{2,2}\right)' \mathbf{x}^2 + \left(\mathbf{p}_2^{2,H}\right)' \mathbf{x}^H,$$

which yields $\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{x}_1$ and $\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{q}_E$ for all $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$. This proves (14).

From this first step we conclude $\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{x}_1$ under condition C.1. Next,we can use a similar argument as in Proposition 2 (for (11)) to obtain (for all $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$)

$$\left(\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{x}_1 \text{ and } y_1 \geq \mathbf{p}_1' (\mathbf{q}_E + \mathbf{x}_2)\right) \Rightarrow \left(\mathbf{x}_1 R_o^{\mathbf{g}^1} \mathbf{q}_E\right). \tag{17}$$

This gives the wanted result: we have $\mathbf{x}_1 R_o^{\mathbf{g}^1} \mathbf{q}_E$ and $\mathbf{x}_2 R_o^{\mathbf{g}^2} \mathbf{q}_E$ for all admissible individual demand functions $(\mathbf{g}^1, \mathbf{g}^2) \in Q(\mathbf{g})$ that satisfy WARP if condition C.1 holds. ■

Appendix 2: A stylized example

The next example shows that our method based on program Program P.1 obtains sharing rule identification for the general version of the collective model, even if no good is specified as public or private a priori. Interestingly, we find that the method can, in principle, produce arbitrarily tight bounds (i.e. precise identification).

We focus on a setting with three goods, i.e. $|N| = 3$. For ease of the argument, we assume a stylized specification of the collective consumption model.[14] Importantly, we

[14] This specification implies household demand quantities that are zero under specific price configu-

recall that in practice the empirical analyst does not observe this consumption model, and so our identification method can only use the household demand that results from it (see further). Specifically, consider a household with individual utility functions (for $(\mathbf{q})_n$ the consumed quantity of the nth entry good)

$$U^1 = (\delta A) \cdot (\mathbf{q})_1 + (B - (\delta - 1)A) \cdot (\mathbf{q})_2 + C \cdot (\mathbf{q})_3 \text{ and} \quad (18)$$
$$U^2 = (B - (\delta - 1)A) \cdot (\mathbf{q})_1 + \delta A \cdot (\mathbf{q})_2 + C \cdot (\mathbf{q})_3, \quad (19)$$

where δ, A, B and C are positive real numbers specified below. Non-negative consumption externalities require $(B - (\delta - 1)A)$ to be positive. Intuitively, referring to our labor supply application in Section 6 of the main text, one may think of goods 1 and 2 as leisure of, respectively, the first and the second household member, while good 3 represents (other) household consumption.

Next, we assume the following Pareto weight specification:

$$\mu = 0 \text{ if } (\mathbf{p})_1 > (\mathbf{p})_2,\ \mu = \infty \text{ if } (\mathbf{p})_2 > (\mathbf{p})_1 \text{ and } \mu = 1 \text{ if } (\mathbf{p})_1 = (\mathbf{p})_2. \quad (20)$$

Essentially, this complies with a bargaining weight that is extremely sensitive to the ratio of the price of good 2 to the price of good 1. In a labor supply setting, the ratio $(\mathbf{p})_2/(\mathbf{p})_1$ represents the relative wages within the household. Then, if wages are different, the member with the higher wage gets full bargaining power, while both members have exactly the same bargaining weight if spouses' wages are equal.

In what follows, we consider $0 < \epsilon < 1$. For given ϵ, we specify δ, A, B and C such that

$$\frac{\delta A}{C} > 2 + \frac{2}{\epsilon} \text{ and } \frac{C}{A+B} > \frac{1}{1+\epsilon}$$

Then, for $(B - (\delta - 1)A)$ positive but sufficiently small, it is easily verified that the model specification in (18) and (20) generates exactly the demand function $\mathbf{g}$ used in our following argument.

Now, consider

$$\mathbf{p}_E = (0.5 + \epsilon/2, 0.5 + \epsilon/2, 1) \text{ and } y_E = 1.$$

Our collective model implies

$$\mathbf{q}_E = \mathbf{g}(\mathbf{p}_E, y_E) = (0, 0, 1).$$

Next, program P.1 obtains

$$\mathbf{p}_1 = (1 + \epsilon, 1, \epsilon/2) \text{ and } y_1 = 1 + \epsilon,$$
$$\mathbf{p}_2 = (1, 1 + \epsilon, \epsilon/2) \text{ and } y_1 = 1 + \epsilon,$$

rations, which considerably facilitates our following argument. At this point, it is worth to emphasize that we use these zero quantities only for mathematical elegance and this does not affect the core of our argument.

with

$$\mathbf{x}_1 = \mathbf{g}(\mathbf{p}_1, y_1) = (1,0,0) \text{ and } \mathbf{x}_2 = \mathbf{g}(\mathbf{p}_2, y_2) = (0,1,0).$$

We note that

$$\begin{aligned} y_1 &= 1+\epsilon,\ \mathbf{p}_1'\mathbf{q}_2 = 1,\ \mathbf{p}_1'\mathbf{q}_E = \epsilon/2, \\ y_2 &= 1+\epsilon,\ \mathbf{p}_2'\mathbf{q}_1 = 1,\ \mathbf{p}_2'\mathbf{q}_E = \epsilon/2, \end{aligned}$$

so that constraint (P.1-2) is indeed satisfied.

Thus, we get

$$\begin{gathered} y_E^{u1} = \mathbf{p}_E'\mathbf{x}_1 = 0.5 + \epsilon/2,\ y_E^{u2} = \mathbf{p}_E'\mathbf{x}_2 = 0.5 + \epsilon/2, \\ y_E^{l1} = y_E - y_E^{u2} = 0.5 - \epsilon/2,\ y_E^{l2} = y_E - y_E^{u1} = 0.5 - \epsilon/2. \end{gathered}$$

These upper and lower bounds will become arbitrarily tight (i.e. precise recovery) if ϵ gets arbitrarily small. In words, in observation E the household members 1 and 2 will apply (approximately) equal income sharing under collectively rational behavior.

References

[1] Afriat, S. (1967), "The construction of utility functions from expenditure data", *International Economic Review*, 8, 67-77.

[2] Bargain, O., and O. Donni (2012), "The measurement of child costs: A Rothbarth-type method consistent with scale economies and parents' bargaining", *European Economic Review*, forthcoming.

[3] Becker, G. (1973), "A theory of marriage: Part I.", *Journal of Political Economy*, 81, 813–846.

[4] Becker, G. (1981), *A Treatise on the Family*, Cambridge: Harvard University Press, 1981

[5] Beresteanu, A. and F. Molinari, (2008), "Asymptotic properties for a class of partially identified models", *Econometrica*, 76, 763–814.

[6] Blundell, R., M. Browning and I. Crawford (2008), "Best nonparametric bounds on demand responses", *Econometrica*, 76, 1227-1262.

[7] Blundell, R., P.-A. Chiappori and C. Meghir (2005), "Collective labor supply with children", *Journal of Political Economy*, 113, 1277-1306.

[8] Blundell, R., P.-A. Chiappori, T. Magnac and C. Meghir (2007), "Collective labor supply: heterogeneity and nonparticipation", *Review of Economic Studies*, 74, 417-445.

[9] Blundell, R. Kristensen, D. and R. Matzkin (2011), “Bounding quantile demand functions using revealed preference inequalities”, *University College London Working Paper.*

[10] Bossert, W. (1993), “Continuous Choice Functions and the Strong Axiom of Revealed Preference”, *Economic Theory*, 3, 379-85.

[11] Bourguignon, F., Browning, M and P.-A. Chiappori (2009), “Efficient intrahousehold allocations and distribution factors: implications and identification”, *Review of Economic Studies*, 76, 503–528.

[12] Browning, M, F. Bourguignon, P.-A. Chiappori and V. Lechene (1994), “Income and Outcomes: A Structural Model of Intrahousehold Allocations”, *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, 1241-1278.

[14] Browning, M., P.-A. Chiappori and A. Lewbel (2006), “Estimating consumption economies of scale, adult equivalence scales, and household bargaining power”, *Boston College Working Paper*, WP588, Department of Economics, Boston College.

[15] Cherchye, L., B. De Rock and F. Vermeulen (2007), “The collective model of household consumption: a nonparametric characterization”, *Econometrica*, 75, 553-574.

[16] Cherchye, L., B. De Rock and F. Vermeulen (2010), “An Afriat Theorem for the collective model of household consumption”, *Journal of Economic Theory*, 145, 1142–1163.

[17] Cherchye, L., B. De Rock and F. Vermeulen (2011), “The revealed preference approach to collective consumption behavior: testing and sharing rule recovery”, *Review of Economic Studies*, 78, 176–198.

[18] Cherchye, L., B. De Rock and F. Vermeulen (2012a), “Married with children: a collective labor supply model with detailed time use and intrahousehold expenditure information”, *American Economic Review*, forthcoming.

[19] Cherchye, L., B. De Rock and F. Vermeulen (2012b), “Economic well-being and poverty among the elderly: an analysis based on a collective consumption model”, *European Economic Review*, forthcoming.

[20] Chernozhukov, V., H. Hong and E. Tamer (2007), “Estimation and confidence regions for parameter sets in econometric models”, *Econometrica*, 75, 1243 -1284.

[21] Chiappori, P.-A. (1988), "Rational household labor supply", *Econometrica*, 56, 63-89.

[22] Chiappori, P.-A. (1992), "Collective labor supply and welfare", *Journal of Political Economy*, 100, 437-467.

[23] Chiappori, P.-A. and I. Ekeland (2006), "The micro economics of group behavior: general characterization", *Journal of Economic Theory*, 130, 1-26.

[24] Chiappori, P.-A. and I. Ekeland (2009), "The micro economics of efficient group behavior: identification", *Econometrica*, 77, 763-799.

[25] Chiappori, P.-A., B. Fortin and G. Lacroix (2002), "Marriage market, divorce legislation and household labor supply", *Journal of Political Economy*, 110, 37-72.

[26] Diewert, W. E. (1973), "Afriat and revealed preference theory", *Review of Economic Studies*, 40, 419–426.

[27] 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.

[28] Gale, D. (1960), "A note on revealed preference", *Economica* 27, 348-354.

[29] Houthakker, H. S. (1950), "Revealed preference and the utility function", *Economica*, 17, 159–174.

[30] Galichon, A. and M. Henry (2009), "A test of non-identifying restrictions and confidence regions for partially identified parameters," *Journal of Econometrics*, 152, 186-196.

[31] Kihlstrom, R., Mas-Colell, A., Sonnenschein, H. F.(1976), "The demand theory of the weak axiom of revealed preference", *Econometrica,* 44, 971-978.

[32] Lewbel, A. and K. Pendakur (2008), "Estimation of collective household models with Engel curves", *Journal of Econometrics*, 147, 350-358.

[33] Lise, J. and S. Seitz (2011), "Consumption inequality and intra-household allocations", *Review of Economic Studies*, 78, 328-355.

[34] Manski, C. (2003), *Partial identification of probability distributions*, New York: Springer Verlag.

[35] McElroy, M. B. (1990), The empirical content of Nash bargained household behavior, *Journal of Human Resources*, 25, 559–583.

[36] Menon, M., K. Pendakur and F. Perali (2012), "On the expenditure-dependence of children's resource shares", unpublished manuscript.

[37] Oreffice, S. (2011), "Sexual orientation and household decision making. Same-sex couples balance of power and labor supply choices", *Labour Economics*, 18, 145-158.

[38] Samuelson, P. A. (1938), "A Note on the pure theory of consumer behavior", *Economica*, 5, 61–71.

[39] Tamer, E. (2010), "Partial identification in econometrics", *Annual Review of Economics*, 2, 167 -195.

[40] Uzawa, H. (1960), "Preference and rational choice in the theory of consumption", in: Arrow, K. J., Karlin and S., Suppes, P. (eds.), *Mathematical methods in the social sciences*, Stanford: Stanford, University Press, 129-148.

[41] Uzawa, H. (1971), "Preference and rational choice in the theory of consumption", in: Chipman, J. S., Hurwicz, L., Richter, M. K., Sonnenschein, H. F. (eds.) Preferences, utility, and demand, New York, Harcourt Brace Jovanovich, 7-28.

[42] Varian, H.R. (1982), "The nonparametric approach to demand analysis", *Econometrica*, 50, 945-972.