CentER

Discussion Paper

No. 2005–51

OPENING THE BLACK BOX OF INTRA-HOUSEHOLD DECISION-MAKING: THEORY AND NON-PARAMETRIC EMPIRICAL TESTS OF GENERAL COLLECTIVE CONSUMPTION MODELS

By Laurens Cherchye, Bram de Rock, Frederic Vermeulen

March 2005

ISSN 0924-7815

Tilburg University

Opening the black box of intra-household decision-making: Theory and non-parametric empirical tests of general collective consumption models*

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

March 21, 2005

Abstract

We non-parametrically test a general collective consumption model with public consumption and externalities inside the household. We further propose a novel approach to model special cases of the general collective model. These special cases include alternative restrictions on the 'sharing rule' that applies to each household, and which defines the distribution of the household budget over the household members. A limiting case is the unitary model. Our application uses data from the Russia Longitudinal Monitoring Survey (RLMS); the panel structure of this data set allows non-parametric testing of the behavioral models without relying on preference homogeneity assumptions across similar individuals. This application includes test results but also a power analysis for different specifications of the collective consumption model. Our main findings are that the most general collective model, together with a large class of special but still fairly general cases, cannot be rejected by the data, while other restricted versions of the general model, including the unitary alternative, are rejected. Since these tests are entirely non-parametric, this provides strong evidence in favor of models focusing on intra-household decision-making.

Key words: collective household models, intra-household allocation, revealed preferences, non-parametric analysis.

JEL-classification: D11, D12, C14.

---

*We are grateful to André Decoster for giving an introduction to the RLMS.

[†]University of Leuven, Campus Kortrijk and Center for Economic Studies; Fund for Scientific Research - Flanders (FWO-Vlaanderen). E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: laurens.cherchye@kulak.ac.be.

[‡]University of Leuven, Campus Kortrijk and Department of Mathematics. E. Sabbelaan 53, B-8500 Kortrijk, Belgium. E-mail: bram.derock@kulak.ac.be.

[§]Tilburg University, CentER. P.O. Box 90153, NL-5000 LE Tilburg, The Netherlands. E-mail: frederic.vermeulen@uvt.nl. Frederic Vermeulen acknowledges the financial support provided through the European Community's Human Potential Programme under contract HPRN-CT-2002-00235 (AGE).

1. Introduction

Microeconomists are increasingly taking an interest in intra-household decision-making models. This interest is, *inter alia*, fed by the methodological argument that *individuals,* and not *households*, have preferences. As a consequence, the standard assumption that a household consisting of several individuals behaves as if it were a single decision-maker seems overly restrictive. This methodological argument is supported by the growing empirical evidence that the standard unitary model indeed does not provide an adequate description of observed multi-person household behavior (see, *e.g.*, Fortin and Lacroix, 1997, Browning and Chiappori, 1998, and Vermeulen, 2004).

The so-called *collective* model for household consumption behavior (after Chiappori, 1988, 1992) has become growingly popular for analyzing intra-household decision-making processes. Apart from the evident point of departure that a household is formed by individuals with own rational preferences, this collective approach uses the mere assumption that intra-household allocations are Pareto-efficient. Contrary to the unitary model, the collective model entails empirical restrictions that seem more difficult to reject when tested on multi-person household data (see, again, Fortin and Lacroix, 1997, Browning and Chiappori, 1998, and Vermeulen, 2004).

One inherent difficulty in the usual testing of behavioral models is that they implicitly rely on non-verifiable assumptions concerning the functional structure of preferences and/or the intra-household bargaining process. Standard parametric tests do not only check a model's theoretical implications, but also the ad-hoc functional specification assumed. A rejection of a behavioral model may thus well be due to ill-specification.

By definition, non-parametric tests do not assume any functional specification regarding the household consumption process. They directly test the adequacy of a theory on the raw quantity and price data by means of revealed preferences axioms. Building further on seminal work of, among others, Afriat (1967), Varian (1982) conceived a necessary and sufficient condition for observed household consumption to be consistent with the unitary model, which is captured in the well-known *Generalized Axiom of Revealed Preference* (*GARP*). Along with parametric restrictions, Chiappori (1988) derived the non-parametric implications of a collective labor supply model. Still, such a labor supply setting involves a number of convenient simplifications for the empirical analyst, such as observability of the individuals' labor supply and the exclusion of public consumption within the household.

It is clear that a realistic modeling of the household consumption process should account for (partial) public consumption of certain commodities (*e.g.*, rent and car use) and externalities (*e.g.*, related to clothing) within the household. Therefore, Cherchye, De Rock and Vermeulen (2004) provided a non-parametric characterization of a very general collective consumption model.[1] Its generality lies in the incorporation of public consumption and externalities within the household, while it does not require any information on the intra-household allocation of the observed consumption bundle. This non-parametric characterization can be viewed as the counterpart of the parametric characterization of Browning and Chiappori (1998), which equally applies to a general collective consumption model that includes public consumption and externalities inside the household.

Non-parametric tests of collective models have been very scarce up to now. Snyder (2000) tests Chiappori's (1988) labor supply model with egoistic agents and observed labor supply. Using semi-algebraic theory for quantifier elimination, she develops a strict necessary and sufficient test for data consistency with collective rationality. This test, however, only applies to settings with two observations. Cherchye and Vermeulen (2004) test unitary and collective models by starting from the respective *GARP* conditions. Their collective test procedure first draws possible, but not observed, intra-household consumption allocations for each household; for each selected allocation, it subsequently tests *GARP* consistency at the level of each individual household member. Conveniently, this procedure can be applied in settings with T observations (where, possibly, $T > 2$). But, again, the study of Cherchye and Vermeulen focuses on the restrictive setting of labor supply behavior of egoistic individuals.

To the best of our knowledge, general collective models with public goods and/or externalities have not yet been tested non-parametrically. A first objective of this paper is to fill this gap. More specifically, we test the non-parametric necessary and sufficient conditions for general collective rationality derived in Cherchye *et alii* (2004) on data that is drawn from the Russia Longitudinal Monitoring Survey (RLMS). The RLMS is one of the few surveys that enables constructing a very detailed panel of household consumption. This is of course interesting, since it permits non-parametric tests without having to assume that preferences are homogenous across similar individuals (or, in the unitary case, households). Moreover, although our sample covers a time series of only eight observations, there is enough relative price variation over time to test behavioral models in a meaningful way.

As for the practical implementation of the general collective rationality restrictions, it is to be recognized that the basic tests derived in Cherchye *et alii* may be computationally burdensome if there are many observations. Still, as we will show, some basic theoretical insights enable considerable efficiency gains in practical applications. The derivation and application of efficient testing algorithms constitutes a second objective of this study.

If the general collective model cannot be rejected, a further step may consist of test-

[1] While our following analysis concentrates on two-person households, it is worth indicating that Cherchye *et alii* also considered the more general case of M-person households (with, possibly, $M > 2$). Still, much of our discussion is easily translated towards such a more general setting.

ing more restrictive versions of the collective model. Evidently, such a more restrictive model implies a higher probability of rejection. Usually, restrictions to the general collective model are defined with respect to individual preferences or to the observability of certain intra-household allocations. An example is Chiappori's (1988, 1992) collective labor supply model with egoistic preferences and observed individual labor. In the current study, we propose a novel (and more general) approach to model restricted versions of the general collective model. Specifically, we consider the possibility of including alternative positions regarding the 'sharing rule' that applies to each household; this sharing rule defines the within-household distribution of the household budget, so reflecting the intra-household bargaining power of the different household members.[2] In fact, as we will demonstrate, this approach obtains the unitary model as a special case of the general model in a rather unexpected manner. A third objective of this paper is to explore and non-parametrically test plausible but more restrictive alternatives of the general collective consumption model.

Finally, one potential drawback of a collective consumption model that takes into account externalities and public consumption inside the household, is that its generality makes it hardly rejectable. Although Cherchye *et alii* (2004) established that such a model can be rejected on the basis of couples' data when there are at least three commodities and three observations, the question remains how powerful the theoretical implications are in real-life empirical applications. Therefore, in addition to the non-parametric tests, we also include a power analysis of the various specifications of the collective consumption model. Such an analysis focuses on the probability of detecting an alternative hypothesis (*e.g.*, based on Becker's (1962) notion of irrational behavior) to the detriment of the behavioral model under study. See also Bronars (1987), who introduced power assessment tools for the non-parametric unitary ($GARP$) condition.

The remainder of this study is structured as follows. In Section 2, we shortly reiterate the general collective consumption model, and introduce the non-parametric necessity condition of Cherchye *et alii* (2004). We also propose some efficiency-enhancing algorithms for practical application of the necessity tests. Section 3 introduces the data of our empirical exercises, and includes the associated test results for the necessary collective rationality condition. Section 4 institutes the 'sharing rule'-based approach for modeling special cases of the general collective consumption model. Section 5 then contains the empirical results for alternative specifications of the collective consumption model, with a special focus on the power of the different specifications. Section 6 concludes.

[2] See Browning, Chiappori and Lewbel (2004) for a recent discussion of this sharing rule concept in a parametric treatment of the collective consumption model.

2. Non-parametric tests of the general collective model

This section presents the general collective consumption model, which includes public consumption and externalities within the household. It also introduces a necessary condition for data consistency with the collective rationality condition, and discusses the (efficient) empirical implementation of the associated necessity tests in practical applications.

2.1. A necessity condition for collective rationality

A fundamental characteristic of the collective approach to household behavior is that each individual household member has her or his own rational preferences, which can be represented by well-behaved individual utility functions. Since our most general collective consumption model allows for both externalities and public consumption inside the household, it is not the *observed* household consumption bundle that generates utility for the individuals, but rather the usually *unobserved* intra-household allocation of this bundle. Let us denote the observed household consumption bundle by $\mathbf{q} \in \mathbb{R}^n_+$, with the associated price vector $\mathbf{p} \in \mathbb{R}^n_{++}$. The intra-household allocation of the consumption bundle of a couple can be expressed as:

$$\mathbf{q} = \mathbf{q}^1 + \mathbf{q}^2 + \mathbf{Q},$$

where $\mathbf{q}^m$ is the private consumption bundle of household member m $(m = 1, 2)$ and where $\mathbf{Q}$ denotes the public consumption bundle within the couple. It is clear that observed consumption can be used for both private and public consumption (or combinations of both). We will denote the set of observed household quantity and price bundles by $S = \{(\mathbf{q}_j; \mathbf{p}_j), j = 1, ..., T\}$. Individual preferences defined over the intra-household allocation are represented by the following monotonously increasing utility functions:[3]

$$U^m(\mathbf{q}^1_j, \mathbf{q}^2_j, \mathbf{Q}_j),\ m = 1, 2.$$

A second fundamental characteristic of collective household models is the assumption that intra-household allocations are Pareto-efficient. Given our non-parametric setting

[3] The fact that we regard monotonously increasing utility functions in principle excludes negative externalities within the household. Admittedly, this assumption may be restrictive in some instances (*e.g.*, tobacco consumption). However, its restrictive nature should not be overestimated, especially within the specific context of a household micro-economy (which differs substantially from that of a macro-level economy). Even though a negative externality may be associated with, *e.g.*, tobacco consumption, the non-smoker's positive valuation of the smoker's utility generated by smoking might well outweigh that negative externality. In addition, within-household mechanisms may be instituted that decrease or even eliminate the negative externalities; see, *e.g.*, the widespread practice of smoking outside in households consisting of smokers as well as non-smokers.

(with a finite set of observations), this assumption can be more exactly stated as (using $\mathbf{0}^n$ for the n-vector of zeroes):

Definition 1. *A pair of utility functions U^1 and U^2 provides a* collective rationalization (CR-2) *of the observed set S, if there exist T combinations of two vectors $\mathbf{q}_j^1$ and $\mathbf{q}_j^2$, both $\in \mathbb{R}_+^n$, and a scalar $\mu_j \in \mathbb{R}_{++}$ such that:*

$$\text{(i) } \mathbf{0}^n \leq \mathbf{q}_j^m \leq \mathbf{q}_j, m = 1, 2;$$

$$\text{(ii) } \mathbf{0}^n \leq \mathbf{q}_j - \mathbf{q}_j^1 - \mathbf{q}_j^2;$$

$$\text{(iii) } U^1\left(\mathbf{q}_j^1, \mathbf{q}_j^2, \mathbf{q}_j - \mathbf{q}_j^1 - \mathbf{q}_j^2\right) + \mu_j U^2\left(\mathbf{q}_j^1, \mathbf{q}_j^2, \mathbf{q}_j - \mathbf{q}_j^1 - \mathbf{q}_j^2\right) \geq$$
$$U^1\left(\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^H\right) + \mu_j U^2\left(\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^H\right)$$
$$\text{for all } \left(\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^H\right) \in \left(\mathbb{R}_+^n\right)^3 \text{ with } \mathbf{p}_j'\left(\mathbf{z}^1 + \mathbf{z}^2 + \mathbf{z}^H\right) \leq \mathbf{p}_j'\mathbf{q}_j.$$

In this definition, the scalars μ_j $(j = 1, ..., T)$ represent the relative bargaining power (*vis-à-vis* individual 1) of household member 2. A greater bargaining power implies, *ceteris paribus*, a higher utility level for the corresponding individual. This higher utility is not necessarily 'produced' by a more favorable own private consumption bundle: it may also follow from individual 1's private consumption (through externalities) or from publicly consumed quantities.

Cherchye *et alii* (2004) (see their Proposition 1; also Chiappori, 1988, for the special case of household labor supply) established that the above representation of collective rationality can be *decentralized* in the following sense. If the *intra-household allocation* of each observed household consumption bundle would be known, together with a corresponding set of *personalized* or *Lindahl* prices, then the general collective model can be tested by checking whether the respective data sets

$$\left\{(\mathbf{q}_j^1, \mathbf{q}_j^2, \mathbf{q} - \mathbf{q}_j^1 - \mathbf{q}_j^2; \boldsymbol{\pi}_j^1, \boldsymbol{\pi}_j^2, \boldsymbol{\pi}_j^H), j = 1, ..., T\right\}$$

and

$$\left\{(\mathbf{q}_j^1, \mathbf{q}_j^2, \mathbf{q} - \mathbf{q}_j^1 - \mathbf{q}_j^2; \mathbf{p}_j - \boldsymbol{\pi}_j^1, \mathbf{p}_j - \boldsymbol{\pi}_j^2, \mathbf{p}_j - \boldsymbol{\pi}_j^H), j = 1, ..., T\right\}$$

both satisfy the *GARP*, where $\boldsymbol{\pi}_j^i$ and $(\mathbf{p}_j - \boldsymbol{\pi}_j^i)$ $(i = 1, 2, H)$ are the vectors of Lindahl prices associated with the different commodity bundles.[4]

Unfortunately, the above observability condition is usually not satisfied. As shown in Cherchye *et alii* (2004) indirect operational necessity and sufficiency tests can be

[4] Note the similarity between the above decentralization result and the well-known decentralization result in the case of egoistic agents without public consumption; this result forms the theoretical basis for a great deal of collective rationality tests (see, *e.g.*, Chiappori, 1988, Fortin and Lacroix, 1997, and Cherchye and Vermeulen, 2004).

derived that exploit the limited (*observable*) price-quantity information available. (In fact, non-observability of the intra-household allocation characteristics implies that these operational necessary and sufficiency condition do not coincide in general.) The starting point of the necessity test for collective rationality is a set of T *directly revealed preferred sets* DRP_j $(j = 1, ..., T)$, which are defined as follows:

Definition 2. *If* $\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j$ *then* $\mathbf{q}_i \in DRP_j$, *where* DRP_j *represents the* directly revealed preferred set *associated with the bundle* $\mathbf{q}_j$.

Remark that these DRP_j sets are also fundamental to the non-parametric testing of the unitary model, where the $GARP$ condition should be satisfied for the *revealed preferred sets* RP_j, which are the transitive closures of the DRP_j sets (see Varian, 1982). In the collective case, the DRP_j sets are used to define *member-specific observable directly revealed preferred sets* $(\widehat{DRP}_j^m)$ and *revealed preferred sets* $(\widehat{RP}_j^m)$:

Definition 3. *The sets* $\widehat{RP}_j^m \subseteq \{\mathbf{q}_1, ..., \mathbf{q}_T\}$, $j \in \{1, ..., T\}$ *and* $m \in \{1, 2\}$, *represent a collection of* observable member-specific revealed preferred sets *if*

$$\begin{aligned}
&\textit{(i)}\ \mathbf{q}_i \in DRP_j \Rightarrow \mathbf{q}_i \in \widehat{DRP}_j^m\ (m = 1 \textit{ or } 2),\\
&\textit{(ii)}\ \mathbf{q}_i \in \widehat{DRP}_j^m \Rightarrow (\mathbf{q}_i \in \widehat{RP}_j^m \wedge \widehat{RP}_i^m \subseteq \widehat{RP}_j^m),\\
&\textit{(iii)}\ (\mathbf{q}_i \in \widehat{DRP}_j^m \wedge \mathbf{q}_j \in \widehat{RP}_i^m) \Rightarrow \mathbf{q}_i \in \widehat{DRP}_j^l\ (m, l \in \{1, 2\}; m \neq l), \textit{ and}\\
&\textit{(iv)}\ (\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'(\mathbf{q}_{j_1} + \mathbf{q}_{j_2}) \wedge \mathbf{q}_{j_1} \in \widehat{RP}_i^m) \Rightarrow \mathbf{q}_i \in \widehat{DRP}_{j_2}^l\ (m, l \in \{1, 2\}; m \neq l).
\end{aligned}$$

The intuition of Property (i) is that if a bundle $\mathbf{q}_i$ was observed when bundle $\mathbf{q}_j$ was also available, then the intra-household allocation associated with the former bundle must be revealed preferred over the latter bundle for at least one of the household members. Given this, we construct revealed preferred sets that (only) include observed directly revealed preferred bundles, *i.e.* $DRP_j = \bigcup_{m=1,2} \widehat{DRP}_j^m$. Further, Property (ii) is simply the transitive closure of the member-specific directly revealed preference relationship. Property (iii) states that if an individual is indifferent between the bundles $\mathbf{q}_i$ and $\mathbf{q}_j$ (each giving rise to a specific intra-household allocation), then the fact that $\mathbf{q}_i$ is observed while $\mathbf{q}_j$ could also be obtained reveals (in a direct manner) that the other household member prefers the former bundle to the latter. Finally, Property (iv) implies that if an individual prefers a bundle $\mathbf{q}_{j_1}$ to the bundle $\mathbf{q}_i$, while the expenses associated with bundle $\mathbf{q}_i$ exceed the expenditures needed to purchase both bundles $\mathbf{q}_{j_1}$ and $\mathbf{q}_{j_2}$, then the other individual should prefer bundle $\mathbf{q}_i$ above bundle $\mathbf{q}_{j_2}$. Intuitively, these properties extend the unitary revealed preference concepts by directly exploiting the Pareto efficiency assumption that underlies the collective approach.

The necessity test for general collective rationality can now be summarized as follows:[5]

Proposition 1. *A* necessary condition *for the existence of utility functions U^1 and U^2 that provide a* CR-2 *of the observed set S is that there exists a collection of observable revealed preferred sets $\widehat{RP}_j^m$ $(m = 1, 2)$, $j \in \{1, ..., T\}$ such that $\mathbf{p}_j'\mathbf{q}_j \leq \mathbf{p}_j'(\mathbf{q}_{r_1} + \mathbf{q}_{r2})$ for all combinations $\mathbf{q}_{r_m} \in \widehat{RP}_j^m$ $(m = 1, 2)$ and $\mathbf{p}_j'\mathbf{q}_j \leq \mathbf{p}_j'\mathbf{q}_r$ for $\mathbf{q}_r \in \bigcap_{m=1,2} \widehat{RP}_j^m$.*

The interpretation of the necessary condition complements the above Property (iv): if household members 1 and 2 prefer respectively $\mathbf{q}_{r_1}$ and $\mathbf{q}_{r_2}$ over $\mathbf{q}_j$, then the choice of $\mathbf{q}_j$ can be rationalized only if it cannot be exchanged for the sum of $\mathbf{q}_{r_1}$ and $\mathbf{q}_{r_2}$. Generally, three different consumption bundles will be involved in the inequality conditions in Proposition 1. Only two bundles are possible, though, if both members prefer bundle $\mathbf{q}_r$ over bundle $\mathbf{q}_j$; implying that $\mathbf{q}_j$ should not be exchangeable for that (single) bundle $\mathbf{q}_r$.

As a final note, we recall that Cherchye *et alii* (2004) also introduced a sufficiency condition for data consistency with the collective rationality model. This condition will be recaptured in Section 4, where we also discuss convergence of the necessity condition towards the sufficiency condition.[6]

2.2. Efficient tests of the necessity condition

The necessity test of the general collective model checks the condition in Proposition 1 for each possible configuration of the member-specific revealed preferred sets (which are by definition consistent with Properties (i) to (iv) in Definition 3). More formally, the necessity test starts from the DRP_j sets ($j = 1, ..., T$), of which each element should be allocated to at least one member-specific directly revealed preferred set ($\widehat{DRP}_j^m$; $m = 1$ or 2). The resulting $\widehat{DRP}_j^m$ sets are then used to construct $\widehat{RP}_j^m$ sets (according to the properties stated in Definition 3). Finally, the necessity condition stated in Proposition 1 should be satisfied for at least one configuration of possible $\widehat{RP}_j^m$ sets ($m = 1, 2$ and $j = 1, ..., T$) in order not to reject the general collective consumption model.

Given all this, the test potentially implies checking the necessity condition on 3^{T^2} configurations of member-specific revealed preferred sets. This may entail a huge computational burden if there are many observations, which may be problematic even for present-day computers. We next present a procedure that may considerably enhance the efficiency of the necessity tests. Essentially, starting from the unitary $GARP$ test,

[5] We refer to Cherchye *et alii* (2004) for the proof of this result.

[6] Cherchye *et alii* obtain that the necessity condition is sufficient for $T = 3$. Hence, their sufficiency tests become relevant as soon as $T \geq 4$, which is the case in our empirical application.

this procedure constructs mutually independent subsets of observations for which the necessity condition can be tested separately.[7]

Unitary GARP testing

As a preliminary step, we recall that the standard (unitary) *GARP* test starts from the DRP_j sets in Definition 2, which subsequently form the basis for constructing the (unitary) revealed preferred sets RP_j (via Warshall's algorithm; see Varian, 1982). The *GARP* condition then states that each observation j should be cost minimizing over its revealed preferred set RP_j.

Our efficiency enhancing procedure concentrates on the *GARP* violating condition for a couple of observations $(i,\, j) \in \{1, \ldots , T\} \times \{1, \ldots , T\}$, *i.e.*

$$i \in RP_j \wedge \mathbf{p}_j'\mathbf{q}_j > \mathbf{p}_j'\mathbf{q}_i. \tag{2.1}$$

If (2.1) is not met, the couple (i, j) cannot be involved (at the level of the individual household members) in a rejection of the necessity condition for a *CR-2* of the data.[8] Specifically, in such a case each constellation of the member-specific revealed preferred sets consistent with Definition 3 (that starts from $DRP_j = \bigcup_{m=1,2} \widehat{DRP}_j^m$) will never imply a violation of the closing condition in Proposition 1 that involves i and j. This is obtained by noting that $i \in \widehat{RP}_j^m$ ($m = 1$ and/or 2) only if $i \in RP_j$, which in turn entails $\mathbf{p}_j'\mathbf{q}_j \leq \mathbf{p}_j'\mathbf{q}_i$ given that (2.1) is not met.

Trimming

The first step of the procedure 'trims out' the observations from the original data set that are 'irrelevant' for the necessity test. Specifically, it uses the above insight to concentrate exclusively on couples of observations (i, j) that satisfy (2.1). Of course, given the construction of the member-specific revealed preferred sets, we should also include the sequence(s) of observations k that lie between i and j (in the sense that $i \in RP_k$ and $k \in RP_j$). More generally, for each couple $(i, j) \in \{1, ..., T\} \times \{1, ..., T\}$ we define the set

$$\begin{aligned} Seq(i,j) &= \{k \mid i \in RP_k \wedge k \in RP_j\} \text{ if (2.1)}, \\ Seq(i,j) &= \emptyset \text{ if not (2.1)}. \end{aligned}$$

It follows from our above argument that we may concentrate on the union of the sets $Seq(i, j)$; we further refer to that union as $Useq$. Intuitively, this means focusing

[7] We programmed all our algorithms (for the necessity as well as the sufficiency conditions) in Fortran. They can be obtained from the authors upon simple request.

[8] We note that the order of the *GARP* violating couple (i, j) is relevant. Specifically, the empirical content of the condition ($i \in RP_j \wedge\, \mathbf{p}_j'\mathbf{q}_j > \mathbf{p}_j'\mathbf{q}_i$) is clearly different from that of ($j \in RP_i \wedge\, \mathbf{p}_i'\mathbf{q}_i > \mathbf{p}_i'\mathbf{q}_j$).

on the couples of observations i and j (and the associated 'in between' observations k) that cannot be rationalized by the unitary model.

In fact, the observations that do not belong to some $Seq(i, j)$ as defined above become 'irrelevant' for the necessity test. Given the exponential increase of the number of computations needed to test collective rationality for larger data sets, excluding these observations may considerably shorten the time needed to come to a verdict. In fact, *GARP* consistency of a particular sample means that all observations are 'trimmed out' by construction. In that case, all observations become 'irrelevant' for the collective necessity test, meaning that the test itself becomes redundant.

Subset testing

The second step, which we refer to as 'subset testing', partitions the (trimmed) data set $Useq$ into subsets that are mutually independent when testing the necessity condition. In this context, 'mutual independence' indicates that for any two subsets, say $Useq^1$ and $Useq^2$ (with $\bigcup_{l=1,2} Useq^l \subseteq Useq$ and $\bigcap_{l=1,2} Useq^l = \emptyset$), we have that $Useq^1$ does not include observations that are implicated in some *GARP* violation contained in the subset $Useq^2$, and *vice versa*. Formally, this means that for each combination of couples $(i_1, j_1) \in Useq^1 \times Useq^1$ and $(i_2, j_2) \in Useq^2 \times Useq^2$ that $i_l, j_l \notin Seq\,(i_m, j_m)$ $(l, m \in \{1, 2\},\, l \neq m)$.

Indeed, a similar argument as before makes that we may restrict to testing the necessity condition for a *CR-2* of the data at the level of the separate subsets rather than the (unpartitioned) union $Useq$. Again, this subsetting may considerably reduce the computational burden of the necessity test, especially when the number of mutually independent subsets gets large.

To conclude, we remark that the partitioning of $Useq$ can proceed efficiently by starting from the sets $Seq\,(i, j)$. Specifically, using the earlier definitions, it can be imposed that for any $(i_1, j_1) \in Useq^1 \times Useq^1$ and $(i_2, j_2) \in Useq^2 \times Useq^2 : Seq\,(i_1, j_1) \cap Seq\,(i_2, j_2) = \emptyset$. Evidently, one should then only focus on *GARP* violating couples of observations (which satisfy condition (2.1)). As a result, a simple enumeration algorithm, which consecutively considers the different violations of the unitary *GARP* test, can identify the maximum number of independent subsets of $Useq$.

3. Empirical application of the necessity tests for collective rationality

This section presents the empirical results for our application of the necessity tests to data taken from the Russia Longitudinal Monitoring Survey (RLMS). Before presenting these results, we discuss some particularities of the RLMS data set.

3.1. Data

The data are drawn from the RLMS. More specifically, they come from Phase II of the RLMS, which covers the time period between 1994 and 2003 (Rounds V-XII). The RLMS contains a lot of socio-economic information like detailed expenditures, incomes, assets and health from a nationally representative sample of Russian households. It was designed to measure the impact of Russian reforms on the economic well-being of households and individuals. Although the RLMS survey design focuses on a longitudinal study of populations of dwelling units, it allows a panel analysis of those households remaining in the original dwelling unit over time.

The sample selection for the present study is for couples with no one else in the household. We select households where both members are employed to mitigate the issue of the non-separability between consumption and leisure (see Browning and Meghir, 1991). Finally, we only consider households that were observed in all the available rounds of Phase II of the RLMS. This results in a sample of 148 couples that are 8 times observed.

In our empirical exercises, we focus on a fairly detailed commodity bundle that consists of 21 nondurable goods: (1) bread, (2) potatoes, (3) vegetables, (4) fruit, (5) meat, (6) dairy products, (7) fat, (8) sugar, (9) eggs, (10) fish, (11) other food items, (12) alcohol, (13) tobacco, (14) food outside the home, (15) clothing, (16) car fuel, (17) wood fuel, (18) gas fuel, (19) luxury goods, (20) services and (21) rent. Although the disaggregation of food items may appear far too detailed, it should be noted that the average budget share of food equals 40% for the selected sample (see also the Appendix). Prices are obtained by averaging recorded prices across the households in a given census region. Some of the commodities that we use are aggregate commodities. The price index for a composite commodity is the weighted geometric mean of the prices of the different items in the aggregate good, with weights equal to the average budget shares in a given census region (*i.e.*, we use the Stone price index). Some summary statistics for our sample are given in the Appendix.

Anticipating the empirical results, it should be stressed that we apply the non-parametric collective rationality test to each separate household, which implies that each household's quantity and price observations form an apart set S. As already mentioned in the introduction, falsification of the general collective model requires (at least) three commodities and three observations. Hence, given that each household is eight times observed and the commodity bundle consists of more than three goods, the general collective model is potentially rejectable. Another advantage of testing at the household level is that we do not need to rely on possibly controversial preference homogeneity assumptions across individuals in different households.

3.2. Empirical results

Table 3.1 summarizes the empirical results for the necessity test of the general collective consumption model. It is clear from the upper panel of the table, that all couples in our sample pass the necessity test. From the middle panel, it can deduced that the consumption behavior of 117 couples, or 79%, can be described by means of a unitary model. In other words, their set of observed quantity-price bundles satisfies the *GARP* condition. Following our trimming procedure, all eight observations of these households are irrelevant for the necessity test in the sense described above.

Next, all households that do not pass the *GARP* test have at least three irrelevant observations. In fact, most of these households have five or six irrelevant observations, which considerably favors the efficiency of the necessity test algorithm. This indicates that the suggested trimming procedure may considerably enhance the efficiency of the testing algorithm in practical applications.

The results of the subset testing procedure are given in the bottom panel of Table 3.1. For most households that do not satisfy the *GARP*, only a single subset can be created from the original data sets. In such cases, all 'relevant' observations are linked to each other via revealed preference relationships; which makes a separate analysis of subsets impossible. For one household, we can distinguish two subsets for which the necessity conditions can be tested separately. It has five relevant observations, which are allocated to subsets with respectively two and three observations. More generally, one may expect this subset procedure to be particularly useful for larger data sets.

What can we infer from these results? A first conclusion is that the general collective consumption model seems to provide an adequate description of the observed consumption behavior of the couples in our sample, at least if the evaluation criterion is non-rejection of its theoretical implications when tested on real data. The unitary model's *GARP* condition, on the other hand, is rejected for about 20% of the households in our sample. In view of the fact that we only have eight observations per household, this gives fairly strong evidence in favor of the collective approach. Moreover, we performed the non-parametric tests at the individual household level, which excludes the interpretation of *GARP* violations as revealing unobserved heterogeneity.

Another conclusion may be that the theoretical implications of the general collective consumption model are simply too 'generous' to obtain violations from realistic data. This alternative interpretation motivates our next section, which discusses how far we can go in restricting the general model. Our following empirical assessment in Section 5 also includes a power analysis of the restricted collective consumption model, which should give a deeper insight into the effective 'generosity' of the alternative model specifications.

Finally, and perhaps most importantly, the following empirical analysis focuses on sufficiency conditions for collective rationality, which naturally complements this first-

Table 3.1: Necessity test results

	Frequency	Percentage
Necessity test		
CR-2 rejected	0	0.00
CR-2 not rejected	148	100.00
Number of irrelevant observations		
0	0	0.00
1	0	0.00
2	0	0.00
3	1	0.68
4	1	0.68
5	8	5.41
6	21	14.19
7	0	0.00
8	117	79.05
Number of subsets (of relevant observations)		
0	117	79.05
1	30	20.27
2	1	0.68

step assessment of the necessity conditions. In particular, while the results in Table 3.1 yield the conclusion that we cannot exclude that a collective rationalization of the data is possible, these further sufficiency results will reveal whether or not it is *certainly* feasible to define collective consumption models that rationalize the observed couples' behavior.

4. Restricting the general collective model: a new approach

If the general collective model cannot be rejected, one may investigate the extent to which more restrictive models are equally plausible. This may be an interesting research question from an empirical point of view. For example, it may examine the validity of (frequently employed) restrictions with respect to individual preferences or to the observability of certain intra-household allocations. But such an investigation may also be interesting from a theoretical perspective. For example, a great deal of the (parametric) testing and identification results in the collective literature are derived for the case of egoistic agents.

This section proposes a novel way to define restrictions of the general collective model. Specifically, the restrictions directly constrain the *sharing rule* that applies within each household. After introducing some general concepts, we present operational sufficient conditions that enable testing data consistency with collective rationality under alternative specifications of the sharing rule restrictions. As we will indicate, these sufficiency tests actually boil down to verifying the unitary $GARP$ condition on a transformed data set. This suggest the consumption models that underly the sufficiency tests as direct collective extensions of the unitary model.

As a preliminary remark, we recall that Cherchye *et alii* (2004) argue, for their general collective consumption model, the convergence of their empirical necessity restrictions towards the sufficiency restrictions when T gets large. Of course, that argument is of limited value here given that our application contains only eight observations for each evaluated household. Still, it does justify at least to some extent an important focus on the empirical testing of the sufficient collective rationality conditions in our empirical exercises. In addition, one may generally expect the necessary condition to converge more rapidly towards the sufficiency condition when the range of possible sharing rules is increasingly restricted. In fact, the necessity condition coincides with the sufficiency condition in the limiting case where the sharing rule restriction implies the latter condition reducing to the unitary $GARP$ condition (see our discussion below).

4.1. Collective rationality under sharing rule restrictions

The philosophy of our approach is that we want to maintain as much as possible the generality with respect to the structure of individual preferences and the non-observability of intra-household allocations. The approach is directly based on the decentralization result discussed in Section 2, which essentially implies that observed household consumption results from a two-step allocation procedure. In the first step, individuals divide the household's total expenditures among each other. In the second step, each individual allocates her or his expenditure share to the household's decomposed (private and public) consumption bundles. An important difference with respect to the more usual decentralization result (that applies under egoistic preferences), is that this (individual) consumption may encompass not only own private consumption, but also the partner's private consumption and public consumption. Given the assumption of Pareto-efficiency, this implies that the intra-household allocation process involves Lindahl or personalized prices, which add-up to observed prices. See also our discussion following Definition 1 in Section 2.

Let us denote individual 1's share in the total household expenditures for observation j as η_j, which entails the following definition:

Definition 4. *The share of individual* 1 *for observation* j, η_j, *equals* $\frac{\boldsymbol{\pi}_j'\widehat{\mathbf{q}}_j}{\mathbf{p}_j'\mathbf{q}_j}$, *where* $\boldsymbol{\pi}_j = (\boldsymbol{\pi}_j^{1\prime}, \boldsymbol{\pi}_j^{2\prime}, \boldsymbol{\pi}_j^{H\prime})'$ *and* $\widehat{\mathbf{q}}_j = (\mathbf{q}_j^{1\prime}, \mathbf{q}_j^{2\prime}, (\mathbf{q}_j - \mathbf{q}_j^1 - \mathbf{q}_j^2)')'$.

Individual 1's share thus equals the ratio of that individual's expenditures on the own consumption bundle, valued at Lindahl prices, to the household's total expenditures. Since $\mathbf{p}_j'\mathbf{q}_j = \widehat{\mathbf{p}}_j'\widehat{\mathbf{q}}_j$, where $\widehat{\mathbf{p}}_j = (\mathbf{p}_j', \mathbf{p}_j', \mathbf{p}_j')'$, individual 2's share is defined as:

Definition 5. *The share of individual* 2 *for observation* j *equals* $\frac{(\widehat{\mathbf{p}}_j - \boldsymbol{\pi}_j)'\widehat{\mathbf{q}}_j}{\mathbf{p}_j'\mathbf{q}_j} = 1 - \eta_j$.

The rationality condition for the restricted collective consumption models includes the idea of a 'sharing rule', which defines the distribution of household expenditures over the different household members. (This sharing rule may be interpreted as reflecting the bargaining power of the different household members in the household allocation process; see in particular the duality result in Proposition 3 of Browning, Chiappori and Lewbel (2004).) A broad class of special cases of the general collective model may then be defined through restrictions on this sharing rule. The different models include various restrictions on the sharing rule in the form of $\alpha(\mathbf{p}_j'\mathbf{q}_j) \leq \boldsymbol{\pi}_j'\widehat{\mathbf{q}}_j \leq (1-\alpha)(\mathbf{p}_j'\mathbf{q}_j)$ (or, equivalently, $\alpha(\mathbf{p}_j'\mathbf{q}_j) \leq (\widehat{\mathbf{p}}_j - \boldsymbol{\pi}_j)'\widehat{\mathbf{q}}_j \leq (1-\alpha)(\mathbf{p}_j'\mathbf{q}_j)$). For example, imagine that individuals in the household can only survive if they receive an expenditure share that is at least equal to $\alpha \in [0, 0.5]$; α can then be literally interpreted as a subsistence share, but also as a minimum requirement for both individuals which prevents the dissolution of the couple.

Formally, this class of restrictions can be represented by the following maximization programme:

Definition 6. *For $\alpha \in [0, 0.5]$, a pair of utility functions U^1 and U^2 provides an α-restricted collective rationalization (α-CR-2) of the observed set S if there exist T combinations of two vectors $\mathbf{q}_j^1$ and $\mathbf{q}_j^2$, both $\in \mathbb{R}_+^n$, and T combinations of Lindahl prices $\boldsymbol{\pi}_j^1, \boldsymbol{\pi}_j^2$ and $\boldsymbol{\pi}_j^H$, all $\in \mathbb{R}_{++}^n$, such that:*

$$\begin{aligned}
(i)\ \mathbf{0}^n &\leq \mathbf{q}_j^m \leq \mathbf{q}_j;\ m = 1, 2;\\
(ii)\ \mathbf{0}^n &\leq \mathbf{q}_j - \mathbf{q}_j^1 - \mathbf{q}_j^2;\\
(iii)\ \alpha(\mathbf{p}_j'\mathbf{q}_j) &\leq \boldsymbol{\pi}_j'\widehat{\mathbf{q}}_j \leq (1-\alpha)(\mathbf{p}_j'\mathbf{q}_j);\\
(iv)\ U^1(\mathbf{q}_j^1, \mathbf{q}_j^2, \mathbf{q}_j - \mathbf{q}_j^1 - \mathbf{q}_j^2) &\geq U^1(\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^H)\\
\text{for all } (\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^H) &\in (\mathbb{R}_+^n)^3 \text{ with } \boldsymbol{\pi}_j'\widehat{\mathbf{z}} \leq \boldsymbol{\pi}_j'\widehat{\mathbf{q}}_j;\\
(v)\ U^2(\mathbf{q}_j^1, \mathbf{q}_j^2, \mathbf{q}_j - \mathbf{q}_j^1 - \mathbf{q}_j^2) &\geq U^2(\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^H)\\
\text{for all } (\mathbf{z}^1, \mathbf{z}^2, \mathbf{z}^H) &\in (\mathbb{R}_+^n)^3 \text{ with } (\widehat{\mathbf{p}}_j - \boldsymbol{\pi}_j)'\widehat{\mathbf{z}} \leq (\widehat{\mathbf{p}}_j - \boldsymbol{\pi}_j)'\widehat{\mathbf{q}}_j;
\end{aligned}$$

where $\widehat{\mathbf{z}} = (\mathbf{z}^{1\prime}, \mathbf{z}^{2\prime}, \mathbf{z}^{H\prime})'$.

The interpretation is directly analogous to that of Definition 1. The mere difference is the sharing rule restriction that is included in *(iii)*. As indicated above, such a restriction may be motivated by subsistence or dissolution-preventing arguments in practical applications. Still, it is worth stressing that it also encompasses a multitude of other special cases, including the more standard restrictions that are defined with respect to (possibly egoistic) individual preferences or the observability of certain intra-household allocations. (Compare with Section 6 in Cherchye *et alii* (2004), which discusses special cases of the general collective model.)

More generally, one may use alternative sharing rule restrictions for different household observations, where these restrictions may vary depending on the household and the specific situation under consideration. We will not explicitly consider such variants in this study, but our following discussion is easily extended to include such cases. Such extensions may, *e.g.*, be worthwhile to consider in applications where additional prior information regarding the intra-household process is available, or when sharing rule recovery forms a main purpose of the testing exercise.

4.2. Sufficiency conditions for collective rationality

Contrary to Section 2, we exclusively focus on sufficient collective rationality conditions in the following. Consistency with the sufficiency consistency condition (for particular α) means that there *certainly* exists at least one specification of the intra-household allocation that guarantees consistency of observed behavior with collective rationality as defined in Definition 6.[9] Building further on Cherchye *et alii* (2004), we get the following sufficiency condition for an α-*CR-2* of the data:[10]

Proposition 2. *A* sufficient condition *for the existence of utility functions U^1 and U^2 that provide an α-*CR-2 *of the observed set S is that there exists a partitioning N_1, N_2 ($N_1 \cup N_2 = \{1, ..., T\}$; $N_1 \cap N_2 = \emptyset$) with ($j \in N_1 \Rightarrow \boldsymbol{\pi}_i' \widehat{\mathbf{q}}_j = \alpha \mathbf{p}_i' \mathbf{q}_j \ \forall i \in \{1, ..., T\}$) and ($j \in N_2 \Rightarrow \boldsymbol{\pi}_i' \widehat{\mathbf{q}}_j = (1-\alpha) \mathbf{p}_i' \mathbf{q}_j \ \forall i \in \{1, ..., T\}$), such that both individual household members meet the corresponding* GARP *conditions.*

One interpretation of this sufficiency condition is that it requires the set S can be transformed in two sets S_1 and S_2 that both satisfy *GARP*. More specifically, $S_1 = \{(\alpha_j \mathbf{q}_j; \mathbf{p}_j),\ j = 1, ..., T\}$ and $S_2 = \{((1-\alpha_j)\mathbf{q}_j; \mathbf{p}_j),\ j = 1, ..., T\}$ where $\alpha_j = \alpha$ if $j \in N_1$ (*i.e.*, individual 1 receives the share α) and $\alpha_j = (1-\alpha)$ if $j \in N_2$ (*i.e.*, individual 1 receives the share $(1-\alpha)$). For example, assume that α is equal to 0.3. In terms of Definition 6, this means that each individual member gets at least 30 percent of the total household means. A sufficient condition for such a collective rationalization to be possible is data consistency with the member-specific *GARP* conditions when the two household members receive either 30 or 70 percent of the total household means. However, the specific value may vary depending on the specific observation. Consequently, for some observations an individual may receive a share of 70 percent, whereas it may amount to only 30 percent in other situations.

The non-parametric test for an α-*CR-2* first transforms the observed consumption bundles $\mathbf{q}_j$ ($j = 1, ..., T$) to $\alpha_j \mathbf{q}_j$ and subsequently tests the standard *GARP* condition on the resulting sets S_1 and S_2. The intuition behind the result is that both individuals should maximize their utility subject to the shares that are allocated to them, and that their choices should be consistent across the observations, independently of the fact whether they received the share α or $(1-\alpha)$. Of course, since intra-household allocations are assumed Pareto-efficient, the above requirements should be simultaneously satisfied for both individuals.

A few observations are in order with respect to the α-*CR-2* restrictions. First, if α equals 0.5, then the above restricted collective model reduces to the standard *unitary model*. If all consumption bundles are multiplied by 0.5, then standard revealed preference sets are obtained, which are the immediate ingredients for the *GARP* test.

[9] One could equally derive a necessity condition for data consistency with the collective model for particular α. The derivation is easily analogous to that of Proposition 3 in Cherchye *et alii* (2004), which obtains the necessary rationality condition for the general collective consumption model. Still, given our main focus on sufficiency conditions in the following, we abstract from explicitly including this result here.

[10] The explanation following the proposition provides the intuition of the result. A formal proof is obtained along directly similar lines as the proof of Proposition 4 in Cherchye *et alii*, which establishes a sufficient rationality condition for the general collective consumption model.

The unitary model is thus a first limit case of the α-*CR-2* model. More generally, this result implies that if individuals in couples *always* share the total resources equally, then this intra-household decision-making process can never be distinguished from the unitary approach.

A second limit case is the so-called *situation-dependent totalitarianism* situation, which is described in Proposition 4 of Cherchye *et alii* (2004). This model can be rationalized by setting α equal to zero, implying that an individual either has own expenditures equal to the household's total resources, or has no own expenditures at all. This model can also be viewed as an extension of the unitary model. In that interpretation, the unitary case implies that one household member has the full decision power within the household: the observed consumption bundle is assumed to result from the maximization of the preferences of the 'totalitarian' decision-maker subject to the household budget constraint. The situation-dependent totalitarianism model then implies *two* separate decision-makers in a couple. Each of these two decision-makers is responsible for only a (disjunct) subset of the observed consumption choices in S. Consequently, the sufficient condition for a 0-*CR-2* implies that there should exist an allocation of all observations in a couple's data set to at most two (disjunct) subsets that individually meet the *GARP* condition. Note that in this case *GARP* is to be checked on a limited number of observations (contained in each subset), whereas in specifications with $\alpha \in]0, 0.5]$, the sets S_1 and S_2 both consist of T observations.

A final observation concerns the fact that the sufficiency collective rationality conditions (in Proposition 2) are generally much easier to test than the necessity conditions (in Proposition 1). Specifically, they require checking at most 2^T alternative specifications of the sets N_1 and N_2, which is much below the maximum number of 3^{T^2} configurations in the necessity tests. Again, further efficiency gains may be realized by various refinements of the testing algorithm (including trimming and subsection testing). For the sake of compactness, we abstract from a detailed discussion here, but the treatment is analogous to that in Section 2.[11] Also, our own application, including the computation of the power measures (which imply 1000 iterations for each household and for the different α-specifications under consideration), does not utilize such efficiency-enhancing strategies. Nevertheless, our different exercises required little computation time (*e.g.*, for a given α the power assessment for the whole sample of all households only took a couple of minutes for a standard PC configuration).

[11] For example, it is easily verified that the efficiency enhancing procedures in Section 2 may equally be applied when testing consistency with the situation-dependent totalitarianism model.

5. Empirical application of the α-restricted tests for collective rationality

This section presents the results for α-restricted collective rationality tests when applied to our RLMS data set. As a main focus will be on the power of the alternative collective rationality models, we first outline our procedure for the power assessment.[12]

5.1. Power assessment method

Generally, a power analysis evaluates the probability of detecting an alternative hypothesis to the model under study. Bronars (1987) first defined power measures for the unitary model. His alternative hypothesis was based on Becker's (1962) notion of irrational behavior, which states that households randomly choose consumption bundles that exhaust the available budget. Bronars' power measures then capture the probability of rejecting the *GARP* condition for such randomly drawn consumption bundles from the observed budget hyperplanes. Our power assessment basically extends Bronars' (unitary) procedure for the collective rationality tests, except from some modifications that specifically relate to the nature of our RLMS data.

At least two data features impact on the power assessment. First, as Bronars has illustrated, power measures crucially depend on the degree of relative price variation in the data. For example, if budget hyperplanes do not intersect for a particular data set, then the unitary model can never be rejected for this data. Second, and more specific to our application, the power assessments should account for the presence of zero expenditures in the data. Generally, this is an important feature of microdata on detailed consumption, which is a particularly relevant consideration for the RLMS (where the data for each survey round refer to the consumption in a single week).

It should be noted that our focus on nondurables mitigates the zero expenditure problem to some extent. In addition, given the relative importance of food in the Russian consumption, the issue of zero expenditures on detailed food items due to infrequency of purchase is probably less important than in OECD countries.[13] Still, we do believe it is important to explicitly take up the presence of zero expenditures in our power assessment. In fact, without explicit correction, randomly drawing commodity

[12] As we concentrate on sufficient conditions for α-restricted collective rationality, our power estimates may also be interpreted as 'upper bounds' for the power of necessary and sufficient (α-restricted) collective rationality conditions. One could similarly conceive 'lower bounds' power measures starting from operational necessity conditions for collective rationality. Like before, these lower bounds will lie closer to the upper bounds for α closer to their maximum value of 0.5.

[13] Also our evaluation of the collective rationality conditions at the individual household level alleviates the potential problem of zero expenditures: if there are no expenditures on a given commodity in all eight rounds, then this household simply has a smaller consumption set than a household that has expenditures on all the commodities.

bundles from a household's budget constraint obtains a zero probability of simulating zero consumption of a certain item. Clearly, such a simulation does not match reality if zero expenditures are effectively observed.

Given all this, we use a power assessment procedure that starts from Becker's (1962) irrational behavior, but takes into account the observed zero expenditures. More specifically, we first calculate per household h and per commodity i the proportion of strictly positive expenditures in the eight household observations. Let us denote this proportion by z_{hi}. The drawings of household-specific irrational commodity bundles then proceeds as follows. First, per commodity i and per time period t we draw a random number from the uniform distribution between 0 and 1. If this commodity- and time-specific number is greater than z_{hi}, then the number v_{hit} is set equal to zero. In the opposite case, the number v_{hit} is the result of a new drawing from the uniform distribution (between 0 and 1). Subsequently, the budget share w_{hit} for household h of commodity i at time t is defined as $(v_{hit}/\sum_i v_{hit})$. Finally, the random/irrational quantity bundle for household h at time t is obtained by multiplying the thus obtained vector of budget shares by the observed expenditure level (of household h at time t), and dividing the different components of the resulting vector by the corresponding components of the observed commodity price vector (for household h at time t).

For each household and per RLMS-round, 1000 random consumption bundles are constructed in the way just described. The advantage of the procedure is that it results in an expected proportion of zero expenditures that complies with the observed proportion. Moreover, if a household does not have any expenditures on a particular commodity in all eight rounds of the RLMS, then it will never be randomly allocated a consumption bundle with strictly positive expenditures on that commodity.

The randomly constructed consumption bundles can now be used to estimate the power of the rationality tests associated with different collective consumption models. A power measure gives the probability that a particular collective rationality test detects such irrational (budget-exhausting) behavior.[14] Our empirical exercise specifically considers two power measures, which exploit the panel structure of our data set and provide useful complementary information. The first measure (labelled *Power 1*) captures the proportion of the 1000 random cases where Becker's irrational behavior is detected for at least one household in the sample. The underlying idea is that a behavioral model is rejected if not all households can be fit in its theoretical implications. However, it is well possible that an outlier-household completely determines this first power measure. Therefore, our second power measure (labelled *Power 2*) gives the average proportion of households where Becker's irrational behavior is detected across all (1000) randomly

[14] Remark that there may be some confusion of tongues when using the notion of *irrational behavior*. In our study, we use the term to refer to randomly drawn commodity bundles, and *not* to household behavior that cannot be fit in the unitary model, which may actually be consistent with a more general collective specification.

drawn scenarios. We may summarize that the *Power 1* measure captures the power of the model at the level of the sample as a whole, while the *Power 2* measure provides complementary information regarding the power of the model at the level of the individual households.

5.2. Empirical results

Table 5.1 summarizes the test results associated with the α-restricted collective consumption models. Before discussing these results in greater detail, recall that our analysis focuses on sufficiency tests for collective rationality. As mentioned before, consistency with these sufficiency conditions for particular α means that there exists at least one definition of the collective consumption model (corresponding to specific sharing rule restrictions) that rationalizes the observed behavior.

Table 5.1: Sufficiency test results

Model	Number of rejections	Power 1	Power 2
$\alpha = 0.5$ (unitary model)	31	100	12.64
$\alpha = 0.495$	19	100	8.41
$\alpha = 0.49$	16	100	6.09
$\alpha = 0.47$	5	100	2.81
$\alpha = 0.45$	1	100	1.70
$\alpha = 0.4$	0	100	0.75
$\alpha = 0.3$	0	100	0.21
$\alpha = 0.2$	0	99.70	0.10
$\alpha = 0.01$	0	99.70	0.06
$\alpha = 0.005$	0	99.70	0.06
$\alpha = 0$ (situation-dependent totalitarianism)	0	99.70	0.05

Note: Power measures are in percentages. Power 1 gives the proportion of randomly drawn data sets for which at least one household does not satisfy the tested condition. Power 2 gives the average proportion of couples that does not satisfy the tested condition across the randomly drawn data sets.

A first observation then pertains to the case where α equals 0.50, which states that the two members equally divide the household resources under all circumstances. As discussed before, the empirical implications of this collective model are indistinguishable from those of the unitary model. Given this, the 31 households that did not pass the *GARP* test (see our discussion of the necessity test results) can never meet the empirical conditions corresponding to this limit case of the collective consumption model. This also appears in Table 5.1.

Next, we find in the table that all couples meet the other ('extreme') situation-dependent totalitarianism condition (for $\alpha = 0$). This implies that there certainly exists a collective rationalization of the data for the general collective consumption model, which complies with not including sharing rule restrictions. Let us then regard to what extent this finding alters for alternative sharing rule constraints. Table 5.1 makes clear that lower α values result in less households not passing the associated rationality tests. For example, 19 couples do not satisfy the α-*CR-2* restrictions under $\alpha = 0.495$ (*i.e.*, the couple's members control either 49.5 or 50.5% of the total expenditures). This number steadily decreases towards zero for lower α: only a single couple violates the α-*CR-2* restrictions for $\alpha = 0.45$; and all households meet the sufficiency restrictions when α is at most 0.40.

These findings suggest that, even though the definition of the collective consumption models underlying the respective sufficient rationality conditions may seem restrictive to some, a wide range of such models is effectively able to describe the observed couples' consumption behavior. Interestingly, these favorable test results should not necessarily be attributed to a low power of the different α-*CR-2* models: the *Power 1* values are close to unity for all the models under evaluation. This indicates that, for random data sets that are constructed on the basis of Becker's (1962) notion of irrational behavior, there will (quasi always) be at least one household that violates the collective rationality restrictions. In our opinion, these high (sample level) power values provide strong support for our above empirical tests results, which pertain to the (sample level) validity of alternative specifications of the collective consumption model.

As discussed in the previous subsection, the measure *Power 2* reveals to what extent these high *Power 1* values are supported by generally high power at the level of the individual households. As for this second measure, we find that the variation across the different collective models is somewhat more pronounced and that, in general, the values are rather low. For example, the unitary model (*i.e.*, for $\alpha = 0.50$) is associated with a *Power 2* value of no more than 12.64 percent: on average, about 13% of the couples do not satisfy the implications when behaving randomly. This percentage further decreases for smaller α-values. For example, when α equals 40 percent, the *Power 2* value drops to only 0.75%, which means that irrational consumption behavior is detected for an average proportion of less than 1 percent of the households.

Given our specific purpose of testing alternative behavioral models, we attribute a relatively high weight to the favorable *Power 1* results. Indeed, the construction of that measure directly complies with our practice to conclude data consistency with a behavioral model only if *all households simultaneously* pass the associated rationality tests. Still, in some instances the *Power 2* results may seem more informative. For example, generally high power estimates at the level of individual households seem recommendable when addressing recovery questions (*e.g.* regarding the intra-household allocation or the preferences of the individual household members) or forecasting issues; see, *e.g.*, Varian (1982, 1983) for recovery and forecasting in the (unitary) non-parametric approach.

From that perspective, it may be interesting to have a look at the possible causes of the relatively low *Power 2* values. One reasonable explanation for these low values lies in the fact that we have only eight observations per household: we may generally expect higher power for larger samples. Moreover, we conduct our analysis at the level of individual households. Parametric applications usually assume that at least part of the preference parameters are similar across different households, which may result in a higher power to detect alternative hypotheses. Obviously, by its very nature this parametric treatment of household heterogeneity is subject to the same risk of specification error as the parametric rationality tests themselves. In view of the particular (non-parametric testing) orientation of the current study, we believe it is recommendable to abstract from a homogeneity assumption across different households, to maximally avoid specification errors. This gives our collective rationality tests as much as possible a genuinely non-parametric interpretation.

6. Summary and conclusions

This paper presents a first empirical application of non-parametric collective rationality tests that account for public consumption and externalities within the household. Specifically, starting from the work of Cherchye *et alii* (2004), we analyzed the collective rationalization of couples that were drawn from the Russia Longitudinal Monitoring Survey (RLMS). Interestingly, the panel structure of this data set allows us to non-parametrically test the collective consumption model without relying on preference homogeneity assumptions across similar individuals.

First, we conceived an algorithm for testing the necessity conditions for the most general collective consumption model, which does not put any structure on the public consumption or the within-household externalities. This algorithm includes a number of efficiency enhancing procedures that may substantially decrease the computational burden associated with the necessity tests; these operational refinements build on basic theoretical insights regarding the revealed preference relationships for individual household members. Application of these tests obtains that collective rationality cannot be rejected for the RLMS data. In addition, it shows the practical usefulness of the suggested efficiency enhancing testing strategies.

Next, we have investigated sufficiency conditions for collective rationality. We first developed a novel non-parametric framework for collective consumption models. This framework is based on the sharing rule concept, which defines the within-household distribution of the household means. Interestingly, the framework incorporates a wide range of special cases of the general collective consumption model (*e.g.*, pertaining to observability of the intra-household allocation of some commodities and specific assumptions regarding the individual preferences). We then conceived operational sufficiency conditions that enable testing alternative positions regarding the specification of the household-specific sharing rules. Interestingly, these sufficient conditions for collective rationality can be conceived as direct extensions of the standard unitary rationality conditions. Specifically, the associated collective tests imply unitary $GARP$ tests for simple transformations of the original data set.

Consistency with these sufficiency conditions means that there exists at least one definition of the collective consumption model (corresponding to specific sharing rule restrictions) that rationalizes the observed behavior. Using this, our empirical investigation obtained that a multitude of collective consumption models is able to describe the couples' consumption behavior in the RLMS data. For example, we found that there certainly exists a collective rationalization of each couple within the data set under the assumption that each household member is responsible for at least 40 percent of the total household means. By contrast, we obtained that the unitary model, which is empirically equivalent to assuming that each household member always gets 50 percent of the total means, is not able to rationalize the observed behavior.

Finally, we have analyzed the power of the alternative specifications of the collective model (which correspond to different sharing rule restrictions). A first power measure captures the probability of detecting irrational behavior of at least one household in the sample. This measure was very close to unity for all collective rationality models that we evaluated. We conclude that the collective rationality tests are rather powerful at the sample level, which provides a strong support for our above empirical findings.

A second, complementary power measure captures the average/expected proportion of households of which irrational behavior is detected. The values of this measure were rather low for all model specifications (including the unitary specification). We believe this result can at least partly be explained by the availability of only eight observations per household. In this respect, it is worth noting that our (necessity and sufficiency) tests also apply to larger data sets. Such larger data sets may entail higher power at the level of individual households (captured by our second power measure). More powerful tests at the level of individual households may especially be interesting if the ultimate objective of the analysis is not so much to test data consistency with the behavioral model (as in this study) but rather to recover more detailed information regarding the intra-household allocation and member-specific preferences, to subsequently forecast household behavior in new situations. See, *e.g.*, Varian (1982, 1983) and, more recently, Blundell, Browning and Crawford (2003a,b), for non-parametric recovery and forecasting tools in the unitary setting.

Apart from increasing the sample size, another potentially fruitful strategy for obtaining more powerful collective rationality tests uses more stringent household-specific sharing rule restrictions (rather than a common restriction for all households, as in our study). Such restrictions can, *e.g.*, be conceived on the basis of additional prior information about the intra-household allocation process. As we indicated, it is easy to extend the proposed testing tools for such sharing rule restrictions that vary for different households and according to the specific situation at hand.

References

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

[2] Becker, G. (1962), "Irrational behaviour and economic theory", *Journal of Political Economy*, 70, 1-13.

[3] Blundell, R., M. Browning and I. Crawford (2003a), "Nonparametric Engel curves and revealed preference", *Econometrica*, 71, 205-240.

[4] Blundell, R., M. Browning and I. Crawford (2003b), "Nonparametric bounds on demand responses", mimeo, Institute for Fiscal Studies, London.

[5] Bronars, S. (1987), "The power of nonparametric tests of preference maximization", *Econometrica*, 55, 693-698.

[6] Browning, M. and P.-A. Chiappori (1998), "Efficient intra-household allocations: a general characterization and empirical tests", *Econometrica*, 66, 1241-1278.

[7] Browning, M., P.-A. Chiappori and A. Lewbel (2004), "Estimating consumption economies of scale, adult equivalence scales, and household bargaining power", *CAM Working Paper*, 2003-12, version November 2004.

[8] Browning, M. and C. Meghir (1991), "The effects of male and female labor supply on commodity demands", *Econometrica*, 59, 925-951.

[9] Cherchye, L., B. De Rock and F. Vermeulen (2004), "The collective model of household consumption: a nonparametric characterization", *CentER Discussion Paper*, 2004-76, CentER, Tilburg University.

[10] Cherchye, L. and F. Vermeulen (2004), "Nonparametric analysis of household labour supply: goodness-of-fit and power of the unitary and the collective model", *CES-KU Leuven Public Economics Discussion Paper*, 03.02, version August 2004.

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

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

[13] Fortin, B. and G. Lacroix (1997), "A test of the neo-classical and collective models of labour supply", *Economic Journal*, 107, 933-955.

[14] Snyder, S. (2000), "Nonparametric testable restrictions of household behavior", *Southern Economic Journal*, 67, 171-185.

[15] Varian, H. (1982), "The nonparametric approach to demand analysis", *Econometrica*, 50, 945-973.

[16] Varian, H. (1983), "Non-parametric Tests of Consumer Behavior", *Review of Economic Studies*, 50, 99-110.

[17] Vermeulen, F. (2004), "And the winner is... An empirical evaluation of unitary and collective labour supply models", forthcoming in *Empirical Economics*.

Appendix

Table 6.1: Summary statistics

	Mean	Std. dev.
Budget shares		
Bread	0.103	0.141
Potatoes	0.010	0.054
Vegetables	0.018	0.055
Fruit	0.013	0.029
Meat	0.093	0.119
Dairy	0.047	0.063
Fat	0.025	0.049
Sugar	0.047	0.092
Eggs	0.011	0.022
Fish	0.016	0.039
Other food	0.017	0.041
Alcohol	0.014	0.041
Tobacco	0.016	0.058
Food outside the home	0.029	0.107
Clothing	0.073	0.158
Car fuel	0.054	0.123
Wood fuel	0.034	0.134
Gas fuel	0.022	0.072
Luxury goods	0.018	0.097
Services	0.191	0.222
Rent	0.146	0.170
Expenditures on nondurables	2578.30	3947.30

Note: Expenditures are in December 2003 Russian rubles per week (1 RUB = 0.03401 USD).