Opening the Black Box of Intrahousehold Decision Making: Theory and Nonparametric Empirical Tests of General Collective Consumption Models

Laurens Cherchye

CentER, Tilburg University, and University of Leuven

Bram De Rock

Université Libre de Bruxelles, ECARES, and ECORE

Frederic Vermeulen

Netspar, CentER, Tilburg University

We provide "revealed preference" tests of general collective consumption models that account for public consumption and externalities within the household. We further propose a novel approach to model special cases of this model, which imply alternative assumptions regarding the sharing rule. Our application uses the panel data from the Russia Longitudinal Monitoring Survey. We find that the general model, together with a large class of special cases, cannot be rejected. By contrast, we do reject the standard unitary model. Since our tests are entirely nonparametric, this provides strong evidence in favor of models focusing on intrahousehold decision making.

We are grateful to editor Monika Piazzesi, three anonymous referees, Martin Browning, Pierre-André Chiappori, Ian Crawford, André Decoster, and Frans Spinnewyn for helpful comments and suggestions, which substantially improved the paper. We also thank seminar participants in Alicante, Amsterdam, Dublin, Leuven, London, Tilburg, Toulouse, and at the 2006 European meeting of the Econometric Society in Vienna for useful discussions. Laurens Cherchye gratefully acknowledges financial support from the Research Fund K.U.Leuven through the grant STRT1/08/004. Frederic Vermeulen gratefully acknowledges financial support from the Netherlands Organisation for Scientific Research through a Vidi grant.

[*Journal of Political Economy,* 2009, vol. 117, no. 6]

I. Introduction

Microeconomists are increasingly taking an interest in intrahousehold decision-making models for describing household consumption behavior. This interest is fed by the methodological argument that *individuals*, and not *households*, have preferences. As a consequence, the standard unitary assumption that a household consisting of several individuals behaves as if it were a single decision maker seems overly restrictive. This methodological argument is further supported by the growing empirical evidence that the unitary model for household consumption behavior indeed does not provide an adequate description of observed multiperson household behavior (see, e.g., Fortin and Lacroix 1997; Browning and Chiappori 1998; Vermeulen 2005). Given this, the so-called *collective* model for household consumption behavior (after Chiappori [1988, 1992]) has become growingly popular for analyzing intrahousehold decision-making processes. The collective model explicitly recognizes that the household consists of multiple decision makers (household members) with their own rational preferences. It only assumes that the household consumption decisions are Pareto-efficient outcomes of an intrahousehold allocation/bargaining process.

This paper adopts a *nonparametric* approach for analyzing collective consumption behavior. Specifically, in what follows, nonparametric analysis stands for "revealed preference" analysis in the tradition of, among others, Afriat (1967) and Varian (1982). This nonparametric approach is to be contrasted with the more standard parametric approach. In terms of the collective consumption model, such a parametric approach implicitly relies on nonverifiable assumptions regarding the functional structure of preferences and the intrahousehold bargaining process. Standard parametric tests check not only a model's theoretical implications but also the ad hoc functional specification that is assumed. A rejection of a behavioral model may thus well be due to misspecification rather than to an inadequate theory as such. By definition, nonparametric tests of consumption models do not assume any functional specification regarding the household consumption process. They directly test the theoretical models on the raw quantity and price data by using revealed preference axioms.

In this paper, we conduct a nonparametric empirical assessment of the *general collective consumption model* introduced by Browning and Chiappori (1998). This model considers general individual preferences that allow for public consumption and consumption externalities within the household. These features are particularly attractive in a household context: many goods are (partially) publicly consumed (e.g., rent and car use), and it seems natural to account for consumption externalities (e.g., related to clothing; this also includes nonegoistic/altruistic preferences). In addition, the general model starts from the minimalistic assumptions that the empirical analyst cannot determine which goods are privately and/or publicly consumed within the household and that there is no information on the intrahousehold allocation of the different goods. Indeed, minimal information is usually available in real-life applications of household consumption models (including our own application). Browning and Chiappori derived testable implications of this general model for the case that starts from a parametric specification of the household members' preferences and the intrahousehold bargaining process; these results were further elaborated by Chiappori and Ekeland (2006).

Browning and Chiappori also provided a parametric empirical assessment of the unitary model and the general collective consumption model for a time series of cross sections constructed on the basis of Canadian household budget surveys. First, they empirically evaluated the unitary model for one-person households (singles) and for two-person households (couples). They found that the unitary model is rejected for couples but not for singles. This suggests that there is something wrong with the preference aggregation assumptions that underlie the unitary approach, that is, that multiperson households behave as single decision makers. Next, Browning and Chiappori observed that the general collective consumption model cannot be rejected for couples. Thus, they concluded that the collective model effectively constitutes a more promising alternative for modeling the behavior of multiperson households. Still, these conclusions are based on parametric tests and thus crucially depend on the functional form that is used for representing the preferences and the intrahousehold bargaining process.

This directly suggests a nonparametric analysis as a natural complement of Browning and Chiappori's parametric analysis. Nonparametric tests of collective consumption models have been very scarce up to now. In fact, the few existing studies focus on the restrictive setting of labor supply behavior of egoistic individuals. This setting implies 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. For example, Snyder (2000) and Cherchye and Vermeulen (2008) conducted empirical tests of the nonparametric conditions derived by Chiappori (1988) for the labor supply model with egoistic household members.

General collective consumption models with public goods and/or externalities have not yet been tested nonparametrically on real-life data. A first objective of this paper is to fill this gap. More specifically, we provide a first empirical application of the nonparametric "revealed preference" conditions for data consistency with the general collective consumption model that have been derived by Cherchye, De Rock, and Vermeulen (2007). In doing so, we consider a data set that is drawn from the Russia Longitudinal Monitoring Survey (RLMS), a nationally representative survey of Russian households that was designed to evaluate the impact of Russian reforms on the economic well-being of households and individuals. The RLMS contains a lot of socioeconomic information such as detailed expenditures, incomes, assets, and health.[1] 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. As such, the RLMS is one of the few surveys that enables constructing a very detailed panel of household consumption. This panel structure of the RLMS is particularly interesting because it permits nonparametric tests without having to assume that preferences are homogeneous across similar individuals. Moreover, although our sample covers a time series of only eight observations for each household, there is enough relative price variation over time to test behavioral models in a meaningful way.

As for the practical implementation of the nonparametric tests of the general collective model, an important consideration concerns the computational burden of these tests. Interestingly, as we will show, some basic theoretical insights can considerably enhance the computational efficiency in practical applications. The derivation and application of efficiency-enhancing testing mechanisms constitute a second objective of this study.

If the general collective model cannot be rejected, a natural further step consists in testing more restrictive versions of the collective model. Evidently, such a more restrictive model implies a higher probability of rejection. Usually, restrictions on the general collective model are defined in terms of individual preferences or the observability of certain intrahousehold 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 approach to model restricted versions of the general collective model, which does not require specific assumptions regarding individual preferences or observability of intrahousehold allocations. Specifically, we consider the possibility of including alternative assumptions regarding the *sharing rule* that applies to each household. This sharing rule defines the within-household distribution of the household budget, so reflecting the intrahousehold bargaining power of the different household members. A third objective of this paper is to nonparametrically test plausible but more restrictive versions of the general collective consumption model,

[1] For more details on the RLMS data, we refer the reader to http://www.cpc.unc.edu/rlms/.

which are defined in terms of specific assumptions regarding the sharing rule that underlies the observed household consumption behavior.

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. The question remains how powerful the theoretical implications are in real-life applications. Therefore, in addition to the nonparametric 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. Bronars (1987) introduced power assessment tools for the nonparametric test of the unitary model, and Andreoni and Harbaugh (2006) provide a survey of nonparametric power assessment tools that are currently available.

The remainder of this study is structured as follows. Section II sets the stage by recapturing nonparametric revealed preference tests of the unitary model. It also introduces the RLMS data for our empirical exercises and includes the corresponding test results for the unitary model (for singles and couples). Section III introduces a necessary condition for data consistency with the general collective consumption model. In addition, it proposes efficiency-enhancing mechanisms for practical application of the necessity tests and presents the results for these necessity tests applied to our RLMS data set. Section IV institutes the sharing rule–based approach for modeling restricted versions of the general collective consumption model. Section V subsequently contains the empirical results for alternative specifications of the collective consumption model, with a special focus on the power of the different specifications. Section VI presents conclusions.

II. Nonparametric Tests of the Unitary Model

As indicated in the introduction, the unitary model treats the household as if it were a single decision maker. This section presents a necessary and sufficient nonparametric condition for data consistency with the unitary model. In addition, we introduce the RLMS data that are used in our empirical exercises. For these data, we also discuss the test results for the unitary model applied to one-person households (singles) and to two-person households (couples).

A. *A Necessary and Sufficient Condition for the Unitary Model*

Suppose that we observe T household consumption choices of n-valued quantity bundles. For each observation j the vector $\mathbf{q}_j \in \mathbb{R}^n_+$ denotes the chosen quantities under the prices $\mathbf{p}_j \in \mathbb{R}^n_{++}$, and $S = \{(\mathbf{p}_j;\ \mathbf{q}_j),\ j = 1, \dots, T\}$ represents the set of all observations for the household under study. The observed household choices are consistent with the unitary model if there exists a single utility function U that *rationalizes* the set of observations S, in the sense of the following definition. Throughout, we will focus on utility functions that are continuous, monotonically increasing, and concave.

DEFINITION 1. Let S be a set of observations. A utility function U provides a unitary rationalization of S if for each $\mathbf{q}_j$

$$U(\mathbf{q}_j) \geq U(\mathbf{z})$$

for all $\mathbf{z} \in \mathbb{R}^n_+$ such that $\mathbf{p}_j'\mathbf{z} \leq \mathbf{p}_j'\mathbf{q}_j$.

In other words, a unitary rationalization requires that each bundle $\mathbf{q}_j$ is utility maximizing subject to the budget constraint. A core result within the nonparametric approach to the unitary model is that a unitary rationalization of the data is possible if and only if the observed set S is consistent with the generalized axiom of revealed preference (GARP).

DEFINITION 2. Let $S = \{(\mathbf{p}_j;\ \mathbf{q}_j);\ j = 1, \dots, T\}$ be a set of observations. The set S satisfies GARP if there exist relations R_0 and R that meet the following rules:

i. if $\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j$, then $\mathbf{q}_i\ R_0\ \mathbf{q}_j$;
ii. if $\mathbf{q}_i\ R_0\ \mathbf{q}_k, \mathbf{q}_k\ R_0\ \mathbf{q}_l, \dots, \mathbf{q}_z\ R_0\ \mathbf{q}_j$ for some (possibly empty) sequence $(k, l, \dots, z)$, then $\mathbf{q}_i\ R\ \mathbf{q}_j$;
iii. if $\mathbf{q}_i\ R\ \mathbf{q}_j$, then $\mathbf{p}_j'\mathbf{q}_j \leq \mathbf{p}_j'\mathbf{q}_i$.

In words, the bundle $\mathbf{q}_i$ is "directly revealed preferred" over the bundle $\mathbf{q}_j$ (i.e., $\mathbf{q}_i\ R_0\ \mathbf{q}_j$) if $\mathbf{q}_i$ was chosen when $\mathbf{q}_j$ was equally attainable (i.e., $\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j$); see rule i. Next, the revealed preference relation R exploits transitivity of preferences; see rule ii. Finally, rule iii imposes that the bundle $\mathbf{q}_j$ cannot be more expensive than any revealed preferred bundle $\mathbf{q}_i$.

We thus have the following result (Varian 1982).

PROPOSITION 1. Let S be a set of observations. There exists a utility function U that provides a unitary rationalization of S if and only if S satisfies GARP.

The GARP condition provides the basis for a test of *unitary rationality*; it expresses the idea that the bundle $\mathbf{q}_j$ is utility maximizing subject to its budget constraint if and only if it is expenditure minimizing over the revealed preferred set of bundles $\mathbf{q}_i$ (i.e., $\mathbf{p}_j'\mathbf{q}_j \leq \mathbf{p}_j'\mathbf{q}_i$ whenever $\mathbf{q}_i\ R\ \mathbf{q}_j$). For each observation, testing GARP consistency proceeds in two steps: the first step starts from the R_0 relations to subsequently construct the R relations (using Warshall's algorithm; see Varian 1982), and the second step checks the expenditure minimization condition.

B. Data

As indicated in the introduction, our data are drawn from the RLMS. The RLMS data collection started in 1992 and was held in two distinct phases: phase I covers the period from 1992 to 1994, and phase II started in 1994. Because phase I and phase II have a different sample design, we take our data from phase II. Specifically, our data set covers the time period between 1994 and 2003 (rounds V–XII). Because the RLMS does not contain data for the years 1997 and 1999, this implies eight observations per household.

Our empirical study considers two samples of households: the first sample contains households that are adult singles, and the second sample contains adult couples. No household contains other persons such as children and/or siblings. We select households in which all members are employed to mitigate the issue of the nonseparability between consumption and leisure (see Browning and Meghir 1991). Finally, we consider only households that were observed in all the available rounds of phase II of the RLMS. This results in a sample of 108 singles and a sample of 148 couples.

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 percent for the selected sample. For a given census region, we obtain prices by averaging recorded prices across all households in that region. Some of the goods that we use are aggregate goods. 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., the Stone price index).

In anticipation of the empirical results, it should be stressed that we apply the nonparametric collective rationality test to each separate household, which implies that each household's quantity and price observations form a separate set S with $T = 8$ and $n = 21$. The fact that we test the unitary model (and, in the next sections, the collective model) for each household separately avoids possibly controversial preference homogeneity assumptions across different households.

C. *Empirical Results*

As a first step, we assess the empirical performance of the unitary model for singles and for couples. By construction, the theoretical predictions of the unitary approach and the collective approach coincide for singles. But, in general, they differ for couples. Thus, if we do not reject the unitary model for singles but we reject the unitary model for couples, then this suggests that the aggregation assumptions that underlie the unitary modeling of couples' behavior are potentially harmful. In addition, given that definition 1 assumes the same utility function for all observed consumption choices, nonrejection of the unitary model for singles provides empirical support for the assumption that individual preferences remain constant over the period under study; in the following sections, we will maintain this constant preferences assumption when interpreting our test results for the collective model.

Table 1 gives the test results. We find that all singles pass the unitary rationality test, and thus, we cannot reject the assumption of constant individual preferences for these data. By contrast, no less than about 21 percent of the couples fail the unitary rationality test. For these couples, it is impossible to construct a single utility function that rationalizes the observed behavior, or these couples do not seem to behave as single decision makers. In our opinion, this finding questions the harmless nature of the aggregation assumption in the unitary approach to modeling couples' behavior. In turn, this suggests the collective approach as potentially more fruitful for rationalizing this behavior. This will be investigated in the next sections.

Before doing so, we provide a somewhat more detailed investigation of the unitary GARP violations for our sample of couples. Specifically, figure 1 presents continuous subperiods of the data that satisfy GARP. We observe 10 different patterns. The first pattern corresponds to the 117 couples that are consistent with GARP when taking the eight period observations together; patterns 2–10 correspond to the remaining 31 GARP violating couples. For example, for pattern 2 (which applies to seven couples in our sample) the chronological sequence runs into a

TABLE 1
Unitary Test Results

	Frequency	Percentage
Singles:		
GARP rejected	0	.00
GARP not rejected	108	100.00
Couples:		
GARP rejected	31	20.95
GARP not rejected	117	79.05

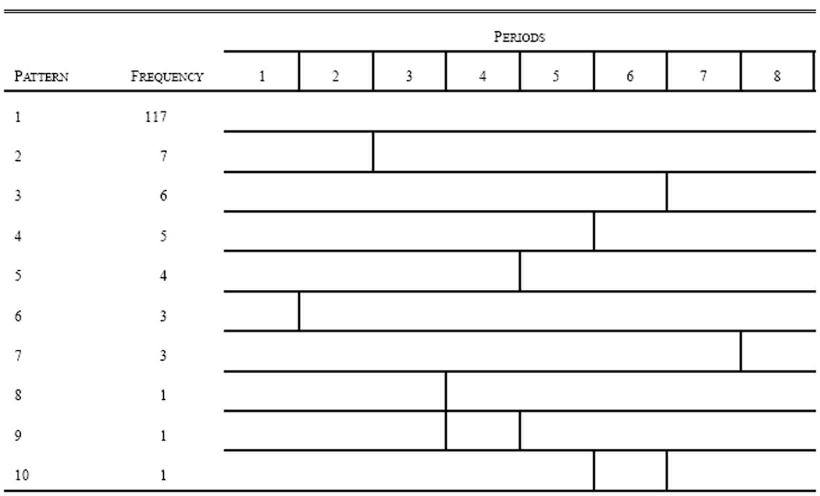

Fig. 1.—Continuous periods that satisfy GARP for couples

violation of GARP when period observation 3 is added. For these couples, we can divide the entire period into two continuous subperiods that are consistent with GARP: the first subperiod captures observations 1 and 2, and the second subperiod captures observations 3–8. An analogous interpretation applies to the remaining patterns in figure 1.

Interestingly, the collective model provides an intuitive interpretation for the patterns reported in figure 1. We recall that the collective model describes observed household consumption behavior as the outcome of an intrahousehold bargaining process between rational individuals. In this respect, we note that the empirical restrictions of the unitary model coincide with those of the collective model if the bargaining power is kept constant over all observed consumption choices (i.e., the members' utility functions can be aggregated into a single household utility function; see also Browning and Chiappori 1998). Thus, when we maintain the assumption of constant preferences (based on our results for singles; see above), observed GARP violations reveal a shift in the bargaining power within the household. For example, the collective model can rationalize the behavior of the couples corresponding to pattern 2 in figure 1 by allowing a bargaining power shift between periods 6 and 7. In our empirical exercise in Section V, we will return in more detail to this particular collective rationalization of the unitary GARP violations.

III. Nonparametric Tests of the General Collective Model

This section presents the general collective consumption model and introduces a necessary condition for data consistency with this model that can be tested on the basis of the available price and quantity data. We also provide an efficient procedure for testing the necessary condition. Finally, we demonstrate the testing procedure for our RLMS data set.

A. *A Necessary Condition for the Collective Model*

A fundamental characteristic of the collective approach is that each individual household member has her or his own rational preferences, which can be represented by individual utility functions. The observed household consumption is then regarded as the Pareto-efficient outcome of a within-household bargaining/allocation process. We next define collective rationality for two-member (1 and 2) households. Essentially, our following discussion (including propositions 2 and 3 below) will recapture the main arguments of Cherchye et al. (2007).

As before, we start from the set $S = \{(\mathbf{p}_j; \mathbf{q}_j);\ j = 1, \ldots, T\}$ of observed consumption choices. The general model makes the distinction between, on the one hand, publicly consumed quantities and, and the other hand, privately consumed quantities for each individual household member. Given this, for observed aggregate quantities $\mathbf{q}_j$ we define *feasible personalized quantities* $\hat{\mathbf{q}}_j$ as

$$\hat{\mathbf{q}}_j = (\mathfrak{q}_j^1, \mathfrak{q}_j^2, \mathfrak{q}_j^h) \quad \text{with } \mathfrak{q}_j^1, \mathfrak{q}_j^2, \mathfrak{q}_j^h \in \mathbb{R}_+^n \text{ and } \mathfrak{q}_j^1 + \mathfrak{q}_j^2 + \mathfrak{q}_j^h = \mathbf{q}_j. \quad (1)$$

Each $\hat{\mathbf{q}}_j$ captures a feasible decomposition of the aggregate quantities $\mathbf{q}_j$ into private quantities for the two household members ($\mathfrak{q}_j^1$ and $\mathfrak{q}_j^2$) and public quantities ($\mathfrak{q}_j^h$). Importantly, we focus on “feasible” personalized quantities in our empirical analysis because we assume that the “true” private and public quantities are not observed by the empirical analyst. In addition, because we account for externalities and public consumption, the household members’ utility functions U^1 and U^2 have the (unobserved) private and public quantities, and not the (observed) aggregate quantities, as arguments.

The observed household choices are consistent with the collective model if there exists a pair of utility functions U^1 and U^2 such that the observed household consumption can be represented as the Pareto-efficient outcome of some within-household bargaining process. If this is the case, then a *collective rationalization* of the set S is possible.

Definition 3. Let S be a set of observations. A pair of utility functions U^1 and U^2 provide a collective rationalization of S if for each $\mathbf{q}_j$ there exist feasible personalized quantities $\hat{\mathbf{q}}_j$ and $\mu_j \in \mathbb{R}_{++}^n$ such that

$$U^1(\hat{\mathbf{q}}_j) + \mu_j U^2(\hat{\mathbf{q}}_j) \geq U^1(\hat{\mathbf{z}}) + \mu_j U^2(\hat{\mathbf{z}})$$

for all $\hat{\mathbf{z}} = (\mathfrak{z}^1, \mathfrak{z}^2, \mathfrak{z}^h)$ with $\mathfrak{z}^1, \mathfrak{z}^2, \mathfrak{z}^h \in \mathbb{R}_+^n$ and $\mathbf{p}_j'(\mathfrak{z}^1 + \mathfrak{z}^2 + \mathfrak{z}^h) \leq \mathbf{p}_j'\mathbf{q}_j$.

Thus, a collective rationalization of the set S is possible if and only if there exist, for each observation j, feasible personalized quantities $\hat{\mathbf{q}}_j$ that maximize a weighted sum of household member utilities U^1 and U^2 for the given household budget $\mathbf{p}_j'\mathbf{q}_j$. In this definition, the Pareto weights μ_j 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. Importantly, 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. We remark that the value of μ_j depends on j, which indicates that the bargaining power can vary across different observations.

We next establish nonparametric conditions for a collective rationalization of a given set S. To do so, we first define *feasible personalized prices* $(\hat{\mathbf{p}}_j^1, \hat{\mathbf{p}}_j^2)$ for observed aggregate prices $\mathbf{p}_j$, as follows:

$$\hat{\mathbf{p}}_j^1 = (\mathfrak{p}_j^1, \mathfrak{p}_j^2, \mathfrak{p}_j^h) \quad \text{and} \quad \hat{\mathbf{p}}_j^2 = (\mathbf{p}_j - \mathfrak{p}_j^1, \mathbf{p}_j - \mathfrak{p}_j^2, \mathbf{p}_j - \mathfrak{p}_j^h) \tag{2}$$

with $\mathfrak{p}_j^1, \mathfrak{p}_j^2, \mathfrak{p}_j^h \in \mathbb{R}_+^n$ and $\mathfrak{p}_j^c \leq \mathbf{p}_j$ $(c = 1, 2, h)$. This concept complements the concept of feasible personalized quantities in (1): $\hat{\mathbf{p}}_j^1$ and $\hat{\mathbf{p}}_j^2$ capture the fraction of the price for the personalized quantities $\hat{\mathbf{q}}_j$ that is borne by, respectively, member 1 and member 2; $\mathfrak{p}_j^1$ and $\mathfrak{p}_j^2$ pertain to private quantities and $\mathfrak{p}_j^h$ to public quantities. These prices can also be interpreted as Lindahl prices: they reflect the willingness to pay of each individual household member for the quantities that are consumed. If there were no externalities, which corresponds to the collective model with egoistic preferences that was mentioned in the introduction, then $\mathfrak{p}_j^1 = \mathbf{p}_j$ and $\mathfrak{p}_j^2 = \mathbf{0}$. In the general model under consideration, however, we can have $\mathfrak{p}_j^1 \neq \mathbf{p}_j$ and $\mathfrak{p}_j^2 \neq \mathbf{0}$.

On the basis of (1) and (2), we define a *set of feasible personalized prices and quantities*

$$\hat{S} = \{(\hat{\mathbf{p}}_j^1, \hat{\mathbf{p}}_j^2; \hat{\mathbf{q}}_j);\ j = 1, \ldots, T\}.$$

Using this concept, we obtain the following nonparametric condition for a collective rationalization of the set S, which provides a *decentralized* representation of collectively rational behavior.

PROPOSITION 2. Let S be a set of observations. There exists a pair of utility functions U^1 and U^2 that provide a collective rationalization of S if and only if there exists a set of feasible personalized prices and quantities $\hat{S}$ such that the corresponding sets $\{(\hat{\mathbf{p}}_j^1, \hat{\mathbf{q}}_j);\ j = 1, \ldots, T\}$ and $\{(\hat{\mathbf{p}}_j^2, \hat{\mathbf{q}}_j);\ j = 1, \ldots, T\}$ simultaneously satisfy GARP.

Hence, while unitary rationality requires GARP consistency of the observed set S (proposition 1), *collective rationality* requires GARP consistency at the level of each individual member m ($m = 1, 2$) in terms of feasible personalized prices $\hat{\mathbf{p}}_j^m$ and quantities $\hat{\mathbf{q}}_j$. This GARP consistency requirement for each member m complies with the required member-specific utility function U^m in the collective rationalization condition in definition 3. Unfortunately, this decentralized representation of collective rationality does not directly yield a test for collective rationality that has practical usefulness, because we observe only the set S and not some set $\hat{S}$. Clearly, in general, a given set S can define infinitely many feasible sets $\hat{S}$.

Given this, our following analysis will focus on a necessary condition for collective rationality that uses solely the available price and quantity data captured by the set S. Essentially, the condition imposes empirical restrictions on *hypothetical* preference relations H_0^m and H^m, which capture "feasible" specifications of the individual preference relations given the information that is revealed by the set S: $\mathbf{q}_i\ H^m\ \mathbf{q}_j$ ($\mathbf{q}_i\ H_0^m\ \mathbf{q}_j$) means that we "hypothesize" that member m (directly) prefers the quantities $\mathbf{q}_i$ over the quantities $\mathbf{q}_j$. We note that, while—of course—the "true" member-specific preferences are expressed in terms of (unobserved) privately and publicly consumed quantities, we define the hypothetical relations H_0^m and H^m in terms of the (observed) aggregate quantities, because we focus on observable implications of the collective consumption model.

The next result defines a number of rules in terms of these hypothetical relations that must be satisfied for a collective rationalization of the data to be possible. We discuss the intuition of the different rules directly after the proposition; this will make clear that the relations H_0^m and H^m extend the unitary revealed preference relations R_0 and R in definition 2 by exploiting the Pareto efficiency interpretation of collectively rational behavior.

PROPOSITION 3. Suppose that there exists a pair of utility functions U^1 and U^2 that provide a collective rationalization of the set of observations S. Then there exist hypothetical relations H_0^m and H^m for each member $m \in \{1, 2\}$ such that the following rules apply:

i. if $\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j$, then $\mathbf{q}_i\ H_0^1\ \mathbf{q}_j$ or $\mathbf{q}_i\ H_0^2\ \mathbf{q}_j$;
ii. if $\mathbf{q}_i\ H_0^m\ \mathbf{q}_k$, $\mathbf{q}_k\ H_0^m\ \mathbf{q}_l$, …, $\mathbf{q}_z\ H_0^m\ \mathbf{q}_j$ for some (possibly empty) sequence $(k, l, \ldots, z)$, then $\mathbf{q}_i\ H^m\ \mathbf{q}_j$;
iii. if $\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j$ and $\mathbf{q}_j\ H^m\ \mathbf{q}_i$, then $\mathbf{q}_i\ H_0^l\ \mathbf{q}_j$ (with $m \neq l$);
iv. if $\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'(\mathbf{q}_{j_1} + \mathbf{q}_{j_2})$ and $\mathbf{q}_{j_1}\ H^m\ \mathbf{q}_i$, then $\mathbf{q}_i\ H_0^l\ \mathbf{q}_{j_2}$ (with $m \neq l$);
v. if $\mathbf{q}_{i_1}\ H^1\ \mathbf{q}_j$ and $\mathbf{q}_{i_2}\ H^2\ \mathbf{q}_j$, then $\mathbf{p}_j'\mathbf{q}_j \leq \mathbf{p}_j'(\mathbf{q}_{i_1} + \mathbf{q}_{i_2})$;
vi. if $\mathbf{q}_i\ H^1\ \mathbf{q}_j$ and $\mathbf{q}_i\ H^2\ \mathbf{q}_j$, then $\mathbf{p}_j'\mathbf{q}_j \leq \mathbf{p}_j'\mathbf{q}_i$.

This necessary condition has a direct interpretation in terms of the Pareto efficiency requirement that underlies collective rationality. First, rule i states that if the household has chosen the bundle $\mathbf{q}_i$ when the bundle $\mathbf{q}_j$ was equally available ($\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j$), then we always have that at least one household member must prefer the former bundle to the latter (i.e., $\mathbf{q}_i\, H_0^1\, \mathbf{q}_j$ or $\mathbf{q}_i\, H_0^2\, \mathbf{q}_j$). Rule ii defines the transitive closures H^1 and H^2 of the relations H_0^1 and H_0^2; it imposes transitivity of the member-specific preferences.

The interpretation of the remaining rules iii–vi pertains to the very nature of the collective model. More specifically, the four rules relate to *rationality across household members* in terms of the hypothetical preference relations H_0^m and H^m. Rule iii expresses that, if member m prefers the bundle $\mathbf{q}_j$ over the bundle $\mathbf{q}_i$ for $\mathbf{q}_j$ not more expensive than $\mathbf{q}_i$, then the choice of the bundle $\mathbf{q}_i$ can be rationalized only if the other member l prefers $\mathbf{q}_i$ over $\mathbf{q}_j$. Indeed, if this last condition were not satisfied, then the bundle $\mathbf{q}_j$ (under the given prices $\mathbf{p}_i$ and outlay $\mathbf{p}_i'\mathbf{q}_i$) would imply a Pareto improvement over the chosen bundle $\mathbf{q}_i$. Similarly, the meaning of rule iv is that if $\mathbf{q}_i$ is more expensive than the sum of $\mathbf{q}_{j_1}$ and $\mathbf{q}_{j_2}$ and member m prefers $\mathbf{q}_{j_1}$ over $\mathbf{q}_i$, then the choice of the bundle $\mathbf{q}_i$ can be Pareto efficient only if the other member l prefers $\mathbf{q}_i$ over $\mathbf{q}_{j_2}$.

Rules i–iv define restrictions on the relations H_0^m and H^m. For a specification of these relations, rules v and vi define expenditure upper bounds. First, rule v complements rule iv: if members 1 and 2 prefer, respectively, $\mathbf{q}_{i_1}$ and $\mathbf{q}_{i_2}$ over $\mathbf{q}_j$, then the choice of the bundle $\mathbf{q}_j$ can be rationalized only if it is not more expensive than the sum of $\mathbf{q}_{i_1}$ and $\mathbf{q}_{i_2}$. Indeed, if this last condition were not met, then for the given prices $\mathbf{p}_j$ and outlay $\mathbf{p}_j'\mathbf{q}_j$, both members would be better off by buying the summed quantities $\mathbf{q}_{i_1} + \mathbf{q}_{i_2}$ rather than the chosen bundle $\mathbf{q}_j$, which of course conflicts with collective rationality. Finally, rule vi considers the special case in which both members prefer the same bundle $\mathbf{q}_i$ over $\mathbf{q}_j$; in that case, Pareto efficiency requires that, under the prices $\mathbf{p}_j$, the bundle $\mathbf{q}_j$ cannot be associated with a strictly higher expenditure level than $\mathbf{q}_i$.

As a final remark, it is interesting to note that the necessary condition in proposition 3 has a structure analogous to the that of the GARP condition in definition 2, which applies to the unitary model. Specifically, GARP states (in the unitary case) rationality conditions in terms of the preference information that is revealed by the observed price and quantity data (see the relations R and R_0 in definition 2). Essentially, the necessary condition in proposition 3 does the same, but now the revealed preference information is understood in terms of the collective model of household consumption and, thus, pertains to the individual household members (see the relations H^m and H_0^m in proposition 3).

B. Efficient Tests of the Necessary Condition

The efficiency-enhancing mechanisms that we present in this section are based on the algorithm described in detail in appendix G of Cherchye et al. (2007). Basically, that algorithm checks the necessary condition in proposition 3 for each possible configuration of the hypothetical relations. More formally, for any couple of observations (i, j) for which $\mathbf{p}_i'\mathbf{q}_i \geq \mathbf{p}_i'\mathbf{q}_j$, we must hypothesize $\mathbf{q}_i\ H_0^1\ \mathbf{q}_j$ or $\mathbf{q}_i\ H_0^2\ \mathbf{q}_j$; this implies three possible scenarios for each couple (i, j): $\mathbf{q}_i\ H_0^1\ \mathbf{q}_j$, $\mathbf{q}_i\ H_0^2\ \mathbf{q}_j$, and $(\mathbf{q}_i\ H_0^1\ \mathbf{q}_j \wedge \mathbf{q}_i\ H_0^2\ \mathbf{q}_j)$. For every combination of the respective scenarios, the algorithm constructs the transitive closures H^1 and H^2 (using Warshall's algorithm, as for the unitary GARP test) and subsequently checks consistency with rules iii–vi of proposition 3. The algorithm concludes consistency with the necessary condition in proposition 3 as soon as there is one construction of the relations H^1 and H^2 that simultaneously meets rules iii–vi of proposition 3.

Given all this, the test implies checking the necessary condition for at most 3^{T^2} configurations of the hypothetical relations. Although this is an extreme situation, also other more realistic situations may be computationally burdensome if there are many observations. This may be problematic even for present-day computers. We next present mechanisms that can considerably enhance the efficiency of the necessity tests. Essentially, starting from the unitary GARP test, these mechanisms construct mutually independent subsets of observations for which the necessary condition can be tested separately (using the above-mentioned algorithm).

Unitary GARP Testing

The efficiency-enhancing testing mechanisms build further on the results for the unitary GARP test; see definition 2. More specifically, they concentrate on the GARP violating condition for a couple of observations (i, j), that is,

$$\mathbf{q}_i R\, \mathbf{q}_j \quad \text{and} \quad \mathbf{p}_j'\mathbf{q}_j > \mathbf{p}_j'\mathbf{q}_i. \tag{3}$$

The basic insight is that, if the GARP violating condition (3) is not met, the couple (i, j) cannot be involved (at the level of the individual household members) in a rejection of the necessary condition for a collective rationalization of the data. Specifically, in such a case each constellation of the member-specific hypothetical relations H^1 and H^2 consistent with rules i–iv in proposition 3 will never imply a violation of rules v and vi of proposition 3 that involves i and j. This is obtained by noting that $\mathbf{q}_i H^m \mathbf{q}_j$ ($m = 1$ or 2) only if $\mathbf{q}_i R\, \mathbf{q}_j$, which in turn entails $\mathbf{p}_j'\mathbf{q}_j \leq \mathbf{p}_j'\mathbf{q}_i$ given that the GARP violating condition (3) is not met. Indeed, by construction, the combination of $\mathbf{q}_i\ H^m\ \mathbf{q}_j$ with $\mathbf{p}_j'\mathbf{q}_j \le \mathbf{p}_j'\mathbf{q}_i$ is always consistent with rules v and vi of proposition 3. Clearly, the order of the GARP violating couple (i, j) is relevant: the empirical content of the condition $(\mathbf{q}_i\ R\ \mathbf{q}_j \wedge \mathbf{p}_j'\mathbf{q}_j > \mathbf{p}_j'\mathbf{q}_i)$ is different from that of $(\mathbf{q}_j\ R\ \mathbf{q}_i \wedge \mathbf{p}_i'\mathbf{q}_i > \mathbf{p}_i'\mathbf{q}_j)$.

Filtering

The first step of the procedure, which we call "filtering," drops from the original data set observations that are uninformative for the necessity test. Specifically, it uses the above insight to concentrate exclusively on couples of observations (i, j) that satisfy the GARP violating condition (3). Of course, given the construction of the member-specific hypothetical relations H^1 and H^2, we must also include the sequence(s) of observations k that lie between i and j ($\mathbf{q}_i\ R\ \mathbf{q}_k$ and $\mathbf{q}_k\ R\ \mathbf{q}_j$). More generally, for each couple (i, j) we define the set

$$\mathrm{Seq}(i, j) = \begin{cases} \{k | \mathbf{q}_i\, R\, \mathbf{q}_k \text{ and } \mathbf{q}_k\, R\, \mathbf{q}_j\} & \text{if (3)} \\ \varnothing & \text{if not (3).} \end{cases}$$

It follows from our above argument that we can concentrate on the union of the sets $\mathrm{Seq}(i, j)$; we further refer to that union as Useq. Intuitively, this means focusing on the couples of observations i and j (and the associated in-between observations k) that cannot be rationalized by the unitary model.

The observations that do not belong to some $\mathrm{Seq}(i, j)$ as defined above are uninformative 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 reach a verdict. In fact, GARP consistency of a particular sample means that all observations are dropped by construction. In that case, all observations are uninformative 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 (filtered) data set Useq into subsets that are mutually independent when testing the necessary condition. In this context, mutual independence indicates that for any two subsets, say $U\mathrm{seq}^1$ and $U\mathrm{seq}^2$ (with $\cup_{l=1,2} U\mathrm{seq}^l \subseteq U\mathrm{seq}$ and $\cap_{l=1,2} U\mathrm{seq}^l = \varnothing$), we have that $U\mathrm{seq}^1$ does not include observations that are implicated in some GARP violation contained in the subset $U\mathrm{seq}^2$, and vice versa. Formally, this means that for each combination of couples $(i_1, j_1) \in U\mathrm{seq}^1 \times U\mathrm{seq}^1$ and $(i_2, j_2) \in U\mathrm{seq}^2 \times U\mathrm{seq}^2$, we have $i_l, j_l \notin \mathrm{Seq}(i_m, j_m)$ $(l, m \in \{1, 2\}, l \neq m)$.

Indeed, an argument similar to one before implies that we may restrict to testing the necessary condition for a collective rationalization of the data at the level of the separate subsets rather than the (unpartitioned) union *U*seq. Again, this subset testing may considerably reduce the computational burden of the necessity test, especially when the number of mutually independent subsets gets large.

C. *Empirical Results*

We next test collective rationality for the 148 couples in the RLMS data set introduced in Section II. Cherchye et al. (2007) established that falsification of the general collective model requires (at least) three goods and three observations. Hence, given that each household is observed eight times and the commodity bundle consists of 21 goods, the general collective model is potentially rejectable.

Table 2 summarizes the empirical results for the necessity test of the general collective consumption model. It is clear from panel A of the table that all couples in our sample pass the necessity test. Thus, we cannot reject collective rationality for any of the 148 couples.

Panel B of table 2 reports the results of the filtering procedure. In

TABLE 2
Necessity Test Results

	Frequency	Percentage
	A. Necessity Test	
Collective rationality rejected	0	.00
Collective rationality not rejected	148	100.00
	B. Filtering Procedure	
Number of uninformative observations:		
0	0	.00
1	0	.00
2	0	.00
3	1	.68
4	1	.68
5	8	5.41
6	21	14.19
7	0	.00
8	117	79.05
	C. Subset Tests	
Number of subsets (of informative observations):		
0	117	79.05
1	30	20.27
2	1	.68

this respect, we recall from our discussion of table 1 that the consumption behavior of 117 couples (79 percent out of the 148 couples) can be described by means of a unitary model; that is, their sets of observed quantity-price bundles satisfy the GARP condition. Following our filtering procedure, all eight observations of these households are uninformative for the necessity test in the sense described above. Next, all households that do not pass the GARP test have at least three uninformative observations. In fact, most of these households have five or six uninformative observations, which considerably favors the efficiency of the necessity test algorithm. This indicates that the filtering procedure considerably enhances the efficiency of the testing algorithm in practical applications.

The results of the subset testing procedure are given in panel C of table 2. For most households that do not satisfy GARP, only a single subset can be created from the original data sets. In such cases, all informative observations are linked to each other via revealed preference relations, which makes a separate analysis of subsets impossible. For one household, we can distinguish two subsets for which the necessary conditions can be tested separately. It has five informative 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 nonrejection of its theoretical implications when tested on real data. Two arguments provide additional support for this conclusion. First, we performed the nonparametric tests at the individual household level, which excludes the interpretation of the GARP violations as revealing cross-sectional unobserved heterogeneity. Second, we recall that the unitary model was not rejected for singles (see table 1), which empirically motivates the assumption that individual preferences are constant, so that we can effectively interpret the unitary GARP violations for couples in terms of bargaining power shifts.

Another conclusion may be that the theoretical implications of the general collective consumption model are simply too generous to obtain violations from real-life data. This alternative interpretation motivates our next section, which discusses how far we can go in restricting the general model. The empirical assessment in Section V 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, the following analysis focuses on sufficient conditions for collective rationality, which naturally complements this first-step assessment of the necessary condition. In particular, while the results in table 2 imply that we cannot exclude a collective rationalization of the data, these further sufficiency results will reveal whether or not it is *certainly* feasible to define (restricted) collective consumption models that rationalize the observed couples' behavior.

IV. Restricting the General Collective Model: A New Approach

If the general collective model cannot be rejected, one can investigate the extent to which more restrictive models are equally plausible. This section proposes a novel way to define restrictions on 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 suggests the consumption models that underlie the sufficiency tests as direct collective extensions of the unitary model.

A. *Collective Rationality under Sharing Rule Restrictions*

We suggest an approach that is as general as possible with respect to the structure of individual preferences and the nonobservability of intrahousehold allocations. The approach is based on the decentralized representation of collective rationality in proposition 2. This representation implies that observed household consumption can also be conceived *as if* it results from a two-step allocation procedure, in the following sense. In the first step, individuals divide the household's total expenditures/income between each other. In the second step, each individual member m ($m = 1, 2$) maximizes her or his utility in terms of the privately and publicly consumed quantities for the given income share and personalized prices. This second step corresponds to the GARP condition for each member m in proposition 2, which effectively implies that each member behaves utility maximizing in terms of a member-specific function U^m.

In formal terms, we can restate the collective rationalization condition in definition 3 as follows: a pair of utility functions U^1 and U^2 provides a *collective rationalization* of S if for each $\mathbf{q}_j$ there exists a set of feasible personalized prices and quantities $\hat{S} = \{(\hat{\mathbf{p}}_j^1, \hat{\mathbf{p}}_j^2; \hat{\mathbf{q}}_j); j = 1, \ldots, T\}$ such that, for each individual member m,

$$U^m(\hat{\mathbf{q}}_j) \geq U^m(\hat{\mathbf{z}}) \quad \text{for all } \hat{\mathbf{z}} = (\mathfrak{z}^1, \mathfrak{z}^2, \mathfrak{z}^h) \tag{4}$$

with $\mathfrak{z}^1, \mathfrak{z}^2, \mathfrak{z}^h \in \mathbb{R}^n_+$ and $(\hat{\mathbf{p}}_j^m)'\hat{\mathbf{z}} \leq (\hat{\mathbf{p}}_j^m)'\hat{\mathbf{q}}_j$. This member-specific utility maximization condition corresponds to the second step of the two-step representation introduced above. As for the first step, the income share allocated to each member m corresponds to $\hat{\eta}_j^m = (\hat{\mathbf{p}}_j^m)'\hat{\mathbf{q}}_j/\mathbf{p}_j'\mathbf{q}_j$. Clearly, we have $\hat{\eta}_j^1 + \hat{\eta}_j^2 = 1$. Given that the shares $\hat{\eta}_j^1$ and $\hat{\eta}_j^2$ are expressed in terms of feasible personalized prices and quantities, we call them *feasible* income shares.

In fact, this two-step representation of the general collective model is directly similar to the well-known two-step representation of collective models with egoistic agents and without public consumption, which has formed the theoretical basis for many collective rationality tests (see, e.g., Chiappori 1988; Fortin and Lacroix 1997; Cherchye and Vermeulen 2008). An important difference is that in the general model the utility of each individual household member U^m does not depend only on her or his own private consumption, but also on the other member's private consumption and on the public consumption. This also makes that the budget constraint $(\hat{\mathbf{p}}_j^m)'\hat{\mathbf{z}} \leq (\hat{\mathbf{p}}_j^m)'\hat{\mathbf{q}}_j$ is defined in terms of (unobserved) personalized prices rather than (observed) aggregate prices.

The income shares of the different household members are determined by the *sharing rule*, which thus governs the within-household distribution of the household income. As such, the sharing rule reflects the bargaining power of the different household members in the household allocation process; we recall that in definition 3 this bargaining power is captured by the Pareto weights. (Browning, Chiappori, and Lewbel [2006] provide a formal discussion of the relation between a member's income share and her or his bargaining power.) It follows from the two-step representation that the sharing rule is determined before the actual consumption choices take place; this parallels the modeling of the bargaining power, with exogenously determined Pareto/bargaining weights, in the collective approach.

In what follows, we will focus on restricted collective consumption models that essentially constrain the feasible income shares $\hat{\eta}_j^1$ and $\hat{\eta}_j^2$; these restricted models capture specific assumptions regarding the sharing rule that underlies the observed household consumption behavior. In particular, we focus on a broad class of special cases of the general collective model that can be defined through alternative sharing rule restrictions of the form $\alpha \leq \hat{\eta}_j^m \leq 1 - \alpha$, which impose that each individual receives a share of at least $\alpha \in [0, 0.5]$. For example, α can then be interpreted as a minimum requirement for both individuals to prevent the dissolution of the couple.

To avoid any possible confusion, we stress that sharing rule restrictions do not imply any specific assumption regarding the true (unobserved) values of the personalized quantities or prices, but only regarding their product. More formally, it is easy to verify that, for any given share $\hat{\eta}_j^1$ and any given personalized quantities $\hat{\mathbf{q}}_j$ (or, conversely, $\hat{\mathbf{p}}_j^1$ and $\hat{\mathbf{p}}_j^2$), one can always find personalized prices $\hat{\mathbf{p}}_j^1$ and $\hat{\mathbf{p}}_j^2$ (or $\hat{\mathbf{q}}_j$) such that $\hat{\eta}_j^1 = (\hat{\mathbf{p}}_j^1)'\hat{\mathbf{q}}_j/\mathbf{p}_j'\mathbf{q}_j$ and $1 - \hat{\eta}_j^1 = (\hat{\mathbf{p}}_j^2)'\hat{\mathbf{q}}_j/\mathbf{p}_j'\mathbf{q}_j$. In other words, sharing rule restrictions do not exclude public consumption and externalities (or non-egoistic/altruistic preferences).

On the basis of (4), the condition for a collective rationalization of a set of observations S under the additional sharing rule restrictions $\alpha \leq \hat{\eta}_j^m \leq 1 - \alpha$ is defined as follows.

DEFINITION 4. Let S be a set of observations and $\alpha \in [0, 0.5]$. A pair of utility functions U^1 and U^2 provides an α-restricted collective rationalization (α-CR) of S if there exists a set of feasible personalized prices and quantities $\hat{S} = \{(\hat{\mathbf{p}}_j^1, \hat{\mathbf{p}}_j^2; \hat{\mathbf{q}}_j); j = 1, \ldots, T\}$ such that, for each individual member m ($m = 1, 2$),

$$U^m(\hat{\mathbf{q}}_j) \geq U^m(\hat{\mathbf{z}}) \quad \text{for all } \hat{\mathbf{z}} = (\mathfrak{z}^1, \mathfrak{z}^2, \mathfrak{z}^h)$$

with $\mathfrak{z}^1, \mathfrak{z}^2, \mathfrak{z}^h \in \mathbb{R}_+^n$ and $(\hat{\mathbf{p}}_j^m)'\hat{\mathbf{z}} \leq \hat{\eta}_j^m(\mathbf{p}_j'\mathbf{q}_j)$, for feasible income shares $\hat{\eta}_j^m = (\hat{\mathbf{p}}_j^m)'\hat{\mathbf{q}}_j/(\mathbf{p}_j'\mathbf{q}_j)$ that satisfy

$$\alpha \leq \hat{\eta}_j^m \leq 1 - \alpha.$$

B. *Sufficient Conditions for Collective Rationality*

Contrary to Section II, we focus exclusively on sufficient collective rationality conditions in the following. Consistency with the sufficient condition (for particular α) means that there *certainly* exists at least one specification of the intrahousehold allocation that guarantees consistency of observed behavior with collective rationality as defined in definition 4. The next result specifies the sufficient condition.[2]

PROPOSITION 4. Let S be a set of observations and $\alpha \in [0, 0.5]$. A pair of utility functions U^1 and U^2 provides an α-restricted collective rationalization (α-CR) of the observed set S if there exists a partitioning N_1, N_2 ($N_1 \cup N_2 = \{1, \ldots, T\}$; $N_1 \cap N_2 = \emptyset$) with ($j \in N_1 \Rightarrow (\hat{\mathbf{p}}_i^1)'\hat{\mathbf{q}}_j = \alpha\mathbf{p}_i'\mathbf{q}_j \ \forall i \in \{1, \ldots, T\}$) and ($j \in N_2 \Rightarrow (\hat{\mathbf{p}}_i^1)'\hat{\mathbf{q}}_j = (1 - \alpha)\mathbf{p}_i'\mathbf{q}_j \ \forall i \in \{1, \ldots, T\}$), such that both individual household members meet the corresponding GARP conditions.

[2] The explanation following the proposition provides the intuition of the result. A formal proof is obtained along lines directly similar to the proof of proposition 4 in Cherchye et al. (2007), which establishes a sufficient condition for the general collective consumption model.

This sufficient condition can be interpreted as follows. For $j \in N_1$ we impose $\hat{\eta}_j^1 = \alpha$ (and thus $\hat{\eta}_j^2 = 1 - \alpha$) by setting $(\hat{\mathbf{p}}_i^1)'\hat{\mathbf{q}}_j = \alpha\mathbf{p}_i'\mathbf{q}_j$ for all i. Alternatively, for $j \in N_2$ we impose $\hat{\eta}_j^1 = 1 - \alpha$ (and $\hat{\eta}_j^2 = \alpha$) by setting $(\hat{\mathbf{p}}_i^1)'\hat{\mathbf{q}}_j = (1 - \alpha)\mathbf{p}_i'\mathbf{q}_j$ for all i. Given this, the construction of the sets N_1 and N_2 corresponds to transforming the set S into two sets, S_1 and S_2, that must both satisfy GARP; intuitively, these sets S_1 and S_2 correspond to, respectively, member 1 and member 2 in the household. More specifically, $S_1 = \{(\mathbf{p}_j;\ \alpha_j\mathbf{q}_j),\ j = 1, \ldots, T\}$ and $S_2 = \{(\mathbf{p}_j;\ (1 - \alpha_j)\mathbf{q}_j),\ j = 1, \ldots, T\}$, 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 $1 - \alpha$). To give an example, assume that $\alpha = 0.3$. In terms of definition 4, this means that each individual member gets at least 30 percent of the total household income. 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 exactly 30 percent or 70 percent of the total household income. 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 nonparametric test for an α-CR first transforms the observed consumption bundles $\mathbf{q}_j$ ($j = 1, \ldots, T$) into $\alpha_j\mathbf{q}_j$ 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 must maximize their utility subject to the shares that are allocated to them and that their choices must be consistent across the observations, independently of the fact whether they received the share α or $1 - \alpha$. Of course, since intrahousehold allocations are assumed to be Pareto efficient, the GARP requirement must be simultaneously satisfied for both individuals.

At this point, it is worth discussing in some more detail the precise interpretation of the household allocation process for data that satisfy the sufficient condition. We note that the above construction of S_1 and S_2 does not exclude egoistic preferences (which means that externalities are absent). Specifically, if the data satisfy the sufficient condition, then we can always specify the feasible personalized prices $\mathfrak{p}_j^1 = \mathbf{p}_j$ and $\mathfrak{p}_j^2 = \mathbf{0}$ and the feasible personalized quantities $\mathfrak{q}_j^1 = \alpha_j\mathbf{q}_j$ and $\mathfrak{q}_j^2 = (1 - \alpha_j)\mathbf{q}_j$, to obtain a collective rationalization in the sense of definition 4. Still, we must emphasize that this specification of these prices and quantities should *not* be the *unique* one that obtains such a collective rationalization. The only valid conclusion is that, for data that satisfy the sufficient condition in proposition 4, this (egoistic) representation of the within-household allocation process always constitutes one possibility; but there may well be other (nonegoistic) representations that equally obtain a collective rationalization of the same data.

The sufficiency tests based on proposition 4 include some interesting limiting cases. First, if $\alpha = 0.5$, then the implications of the above restricted collective model reduce to those of the unitary model. Indeed, if all consumption bundles are multiplied by 0.5, then it is easily verified that the corresponding GARP tests for the individual members are formally equivalent to the unitary GARP test for the household. As such, we cannot distinguish empirically the 0.5-CR model from the unitary model. In other words, we obtain the unitary rationality test as a limiting case within the general class of α-CR tests.

Another limiting case is the test for the *situation-dependent dictatorship* model, which is described in proposition 4 of Cherchye et al. (2007). We obtain this test if we set α equal to zero. The interpretation of the corresponding collective consumption model is that, depending on the specific choice observation at hand, each individual household member either controls the full household income/expenditures or controls no income at all. As such, the couple has two "dictatorial" decision makers, who are each responsible for only a (disjoint) subset of the observed consumption choices in S. Consequently, the sufficient condition for a 0-CR requires that the observed set S can be partitioned into two subsets S_1 and S_2 that individually meet the GARP condition; the sets S_1 and S_2 then correspond to, respectively, member 1 and member 2 as the dictatorial decision makers. We note that in this case the sets S_1 and S_2 may contain fewer than T observations, whereas in specifications with $\alpha \in\,]0, 0.5]$, the sets S_1 and S_2 always consist of T observations.

Two additional remarks are in order with respect to the α-CR restrictions in proposition 4. First, one can conceive of alternative refinements of the sufficient condition. In this respect, we recall that this condition allows for sharing rule shifts between every two consecutive observations; in terms of the notation used in proposition 4, this means that for every two observations i and $i+1$ we can have $i \in N_1$ and $i+1 \in N_2$ (or, alternatively, $i \in N_2$ and $i+1 \in N_1$). Refinements of the sufficient condition (and corresponding test) can limit this flexibility for the sharing rule shifts. We will illustrate such an extension in our empirical application.

A final remark concerns the fact that the sufficient condition in proposition 4 is generally much easier to test than the necessary condition in proposition 3. Specifically, the sufficiency tests require checking at most 2^T alternative specifications of the sets S_1 and S_2, which is much below the maximum number of 3^{T^2} configurations in the necessity test. Again, further efficiency gains may be realized by efficiency-enhancing mechanisms such as filtering and subset testing. For the sake of compactness, we refrain from a detailed discussion here, but the treatment is analogous to the one in Section III. Also, our own application, including the computation of the power measures (which imply 1,000 iterations for each household and for the different α-specifications under consideration), does not utilize efficiency-enhancing mechanisms. Nevertheless, our different exercises required little computation time (e.g., for a given α the power assessment for the whole sample of all households took only a couple of minutes for a standard personal computer configuration).

V. Empirical Application of the α-CR Tests

This section presents the results for α-CR 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.

A. *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 to our collective rationality tests, except from some modifications that specifically relate to the nature of our RLMS data.

At least two data features affect 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 set. The results in Sections II and III show that there is enough price variation in our sample for such rejection. 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. Still, we do believe that it is important to explicitly take up the presence of zero expenditures in our power assessment. In fact, without explicit correction, randomly drawing commodity 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 drawing 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 zero and one. 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 zero and one). 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 for each RLMS round, 1,000 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. 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 (labeled *power 1*) captures the proportion of the 1,000 random cases in which 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 (labeled *power 2*) gives the average proportion of households in which Becker's irrational behavior is detected across all (1,000) randomly drawn scenarios. In summary, the power 1 measure captures the power of the model at the level of the sample as a whole, and the power 2 measure provides complementary information regarding the power of the model at the level of the individual households.

B. *Empirical Results*

Table 3 summarizes the test results associated with the α-CR models applied to the 148 couples in our sample. Before discussing these results in greater detail, we recall that our analysis focuses on sufficiency tests for collective rationality. As mentioned before, consistency with the sufficient conditions for a 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.

A first observation then pertains to the case in which $\alpha = 0.50$, which states that the two members divide the household income/expenditures equally 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 couples that did not pass the unitary GARP test (see our discussion in Sec. II) can never meet the empirical conditions corresponding to this limiting case of the collective consumption model. This also appears in table 3.

Next, we find in the table that all couples meet the (other extreme) situation-dependent dictatorship condition (for $\alpha = 0$). This implies that there certainly exists a collective rationalization of the data for the general collective consumption model. We recall that in Section III the necessary condition for collective rationality is satisfied for our sample of couples. Here we construct a specification of the intrahousehold

TABLE 3
SUFFICIENCY TEST RESULTS

Model	Number of Rejections	Power 1	Power 2
$\alpha = .5$	31	100.0	12.63
$\alpha = .495$	19	100.0	11.74
$\alpha = .49$	16	100.0	10.17
$\alpha = .47$	5	100.0	5.89
$\alpha = .45$	1	99.9	4.05
$\alpha = .4$	0	96.3	2.15
$\alpha = .3$	0	68.8	.77
$\alpha = .2$	0	38.3	.32
$\alpha = .01$	0	7.8	.06
$\alpha = 0$*	0	7.5	.05

* Situation-dependent dictatorship.

allocation process that is certainly consistent with collective rationality defined in definition 3. Given this, one can then investigate which extra restrictions can be added to this general model. More precisely, here we regard to what extent the above findings change for alternative sharing rule constraints. Table 3 makes clear that lower α values result in more couples passing the associated rationality tests. For example, 19 couples do not satisfy the α-CR restrictions in proposition 4 under $\alpha = 0.495$ (i.e., the couple's members receive either 49.5 percent or 50.5 percent of the total expenditures). This number steadily decreases toward zero for lower α: only a single couple violates the α-CR restrictions in proposition 4 for $\alpha = 0.45$, and all couples meet the sufficiency restrictions when α is not above 0.40.

These findings suggest that, even though the definition of the collective consumption models underlying the respective sufficient 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 models: the power 1 values are (quasi) 100 percent for all the models in which α is at least equal to 0.45; and the value equals no less than 96 percent for the model that uses α equal to 0.40, which cannot be rejected for any couple in our sample.

As discussed above, 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 more pronounced and that, in general, the values are rather low. Specifically, while the unitary model (which complies with $\alpha = 0.50$) is associated with a power 2 value of 12.63 percent, which means that on average about 13 percent of the couples do not satisfy the unitary restrictions when behaving randomly, the power 2 measure decreases rapidly when α becomes smaller. For example, when $\alpha = 0.40$, the power 2 value drops to only 2.15 percent, which means that irrational consumption behavior is detected for an average proportion of slightly more than 2 percent of the couples.

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 of concluding 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 intrahousehold allocation or the preferences of the individual household members) or forecasting issues (e.g., to predict household consumption in new price and income situations); see, for example, Varian (1982, 1983, 2006) and Blundell, Browning, and Crawford (2003, 2008) for nonparametric recovery and forecasting tools in the unitary setting.

It is interesting to have a closer 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 individuals, 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 (nonparametric testing) orientation of the current study, we believe that it is recommendable to abstract from a homogeneity assumption across different individuals to maximally avoid specification errors.

Another reason pertains to the fact that the general collective model of this paper assumes minimal information regarding the intrahousehold allocation. Given the specific nature of the α-CR tests, natural extensions assume (or, alternatively, test) a specific structure regarding possible shifts in the sharing rule over the observed household choices. Indeed, we recall that the above sufficiency tests allow for sharing rule shifts between every two consecutive observations. Therefore, as indicated in Section IV, one can refine the sufficiency tests by limiting the flexibility for the sharing rule shifts.

For our application, we illustrate this possibility by using the information reported in figure 1. More specifically, we investigate the assumption that continuous subperiods of the data that satisfy GARP are characterized by the same sharing rule, and thus, we allow for sharing rule shifts only when we run into a GARP violation when adding observations to the chronological sequence. For example, for pattern 1 in figure 1, which corresponds to the 117 couples that are consistent with GARP when taking the eight period observations together, we assume that the income shares are the same in all observations. By contrast, for pattern 2 we allow for a (single) sharing rule shift between periods 2 and 3. And so on. In fact, adding this structure directly complies with the "collective rationalization" for the GARP violations in figure 1, which interprets such violations as revealing shifts in the bargaining power (and thus the sharing rule) within the household; see our discussion in Section II.

Table 4 gives the test results for this refined sufficient condition. We

TABLE 4
TEST RESULTS FOR A REFINED SUFFICIENT CONDITION

Model	Number of Rejections	Power 1	Power 2
$\alpha = .5$	31	100.0	12.63
$\alpha = .495$	29	100.0	13.30
$\alpha = .49$	26	100.0	13.38
$\alpha = .47$	16	100.0	12.81
$\alpha = .45$	14	100.0	12.66
$\alpha = .4$	9	100.0	11.57
$\alpha = .3$	3	100.0	11.60
$\alpha = .2$	1	100.0	11.11
$\alpha = .01$	0	100.0	10.46
$\alpha = 0$*	0	100.0	10.43

* Situation-dependent dictatorship.

consider the same α-values as in table 3, but as just explained, we limit the observations in which a sharing rule shift can take place. We find that all couples pass the refined collective rationality test for $\alpha = 0$ and $\alpha = 0.1$. In addition, we observe very few rejections of the collective rationality conditions for $\alpha = 0.2$, $\alpha = 0.3$, and, to a somewhat lesser extent, $\alpha = 0.4$. As before, the number of rejections increases when α increases. Interestingly, as compared to table 3, table 4 reports much higher power values, in particular for low α-values (including $\alpha = 0$ and $\alpha = 0.1$). More specifically, power 1 values are 100 percent for all models under consideration. In addition, power 2 values are everywhere close to the value of 12.63 percent that applies to the unitary model.

Generally, the results in table 4 demonstrate the usefulness of assuming additional sharing rule structure when starting from the basic α-CR condition in proposition 4. This practice obtains tests for the collective model that are as simple to implement as the basic sufficiency tests themselves. Importantly, these refined tests can be considerably more powerful than the original tests. We have illustrated this by a specific example that builds on the results for the unitary GARP tests to add intuitive sharing rule structure. But, of course, alternative refinements can include other sharing rule restrictions for specific households and consumption choice observations.

VI. Summary and Conclusions

We have presented a first empirical application of nonparametric collective rationality tests that account for public consumption and externalities within the household. Specifically, starting from the work of Cherchye et al. (2007), we analyzed the collective rationalization of couples that were drawn from the Russia Longitudinal Monitoring Survey. Interestingly, the panel structure of this data set allows us to nonparametrically test the collective consumption model without relying on preference homogeneity assumptions across similar individuals.

First, we conceived an efficient procedure to test the necessary condition for the general collective consumption model, which does not put any structure on the public consumption or the within-household externalities. This procedure includes a number of efficiency-enhancing mechanisms that can substantially lower the computational burden associated with the necessity test; these operational refinements build on basic theoretical insights regarding the revealed preference relationships for individual household members. Application of this necessity test obtains that collective rationality cannot be rejected for the RLMS data. In addition, it shows the practical usefulness of the efficiency-enhancing testing mechanisms.

Next, we have investigated sufficient conditions for collective rationality. We first developed a novel nonparametric framework for collective consumption models. This framework is based on the sharing rule concept, which defines the within-household distribution of the household income. The framework incorporates a wide range of special cases of the general collective consumption model, which incorporate alternative assumptions regarding the specification of the household-specific sharing rules. We then conceived operational sufficient conditions that enable testing such sharing rule assumptions. 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 the unitary GARP tests for simple transformations of the original data set, which makes them easy to implement.

Consistency with these sufficient conditions means that there exists at least one definition of the collective consumption model (satisfying specific sharing rule restrictions) that rationalizes the observed behavior. Using this, our empirical investigation obtained that a multitude of collective consumption models are 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 receives at least 40 percent of the total household income. By contrast, we obtained that the unitary model is not able to rationalize the observed couples' behavior, whereas it does well fit observed singles' behavior. Interestingly, these results are consistent with the results of Browning and Chiappori (1998) discussed in the introduction. We recall that Browning and Chiappori provided parametric tests of the alternative behavioral models and focused on a data set drawn from a time series of cross sections. Given that our tests are entirely nonparametric and because the panel structure of our data set avoids potentially distortive preference homogeneity assumptions across different individuals, this provides strong evidence in favor of models focusing on intrahousehold decision making.

Finally, we have analyzed the power of 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 above 95 percent for a large class of the collective rationality models that we evaluated. We conclude that the collective rationality tests are rather powerful at the sample level, which provides strong support for our above empirical findings. A second 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 that 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).

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. For instance, such restrictions can be conceived on the basis of additional prior information about the intrahousehold allocation process. As we indicated, it is easy to extend the proposed sufficiency tests by adding sharing rule restrictions that vary for different households and choice observations. This obtains refined sufficiency tests that are as easy to implement as the basic sufficiency tests themselves but can be substantially more powerful. We have illustrated this for our own empirical application by building on the unitary GARP tests to add structure to the collective rationality tests. These more powerful tests provided further empirical support for the collective approach to modeling the consumption behavior of multiperson households (in the case of couples).

References

Afriat, Sidney. 1967. "The Construction of Utility Functions from Expenditure Data." *Internat. Econ. Rev.* 8:67–77.

Andreoni, James, and William Harbaugh. 2006. "Power Indices for Revealed Preference Tests." Working Paper no. 2005-10 (version 2006), Dept. Econ., Univ. Wisconsin–Madison.

Becker, Gary. 1962. "Irrational Behavior and Economic Theory." *J.P.E.* 70:1–13.

Blundell, Richard, Martin Browning, and Ian Crawford. 2003. "Nonparametric Engel Curves and Revealed Preference." *Econometrica* 71:205–40.

———. 2008. “Best Nonparametric Bounds on Demand Responses.” *Econometrica* 76:1227–62.

Bronars, Stephen. 1987. “The Power of Nonparametric Tests of Preference Maximization.” *Econometrica* 55:693–98.

Browning, Martin, and Pierre-André Chiappori. 1998. “Efficient Intrahousehold Allocations: A General Characterization and Empirical Tests.” *Econometrica* 66:1241–78.

Browning, Martin, Pierre-André Chiappori, and Arthur Lewbel. 2006. “Estimating Consumption Economies of Scale, Adult Equivalence Scales, and Household Bargaining Power.” Working Paper no. WP588, Dept. Econ., Boston Coll.

Browning, Martin, and Costas Meghir. 1991. “The Effects of Male and Female Labor Supply on Commodity Demands.” *Econometrica* 59:925–51.

Cherchye, Laurens, Bram De Rock, and Frederic Vermeulen. 2007. “The Collective Model of Household Consumption: A Nonparametric Characterization.” *Econometrica* 75:553–74.

Cherchye, Laurens, and Frederic Vermeulen. 2008. “Nonparametric Analysis of Household Labour Supply: Goodness-of-Fit and Power of the Unitary and the Collective Model.” *Rev. Econ. and Statis.* 90:267–74.

Chiappori, Pierre-André. 1988. “Rational Household Labor Supply.” *Econometrica* 56:63–89.

———. 1992. “Collective Labor Supply and Welfare.” *J.P.E.* 100:437–67.

Chiappori, Pierre-André, and Ivar Ekeland. 2006. “The Micro Economics of Group Behavior: General Characterization.” *J. Econ. Theory* 130:1–26.

Fortin, Bernard, and Guy Lacroix. 1997. “A Test of the Neo-Classical and Collective Models of Labour Supply.” *Econ. J.* 107:933–55.

Snyder, Susan. 2000. “Nonparametric Testable Restrictions of Household Behavior.” *Southern Econ. J.* 67:171–85.

Varian, Hal. 1982. “The Nonparametric Approach to Demand Analysis.” *Econometrica* 50:945–73.

———. 1983. “Nonparametric Tests of Consumer Behavior.” *Rev. Econ. Studies* 50:99–110.

———. 2006. “Revealed Preference.” In *Samuelsonian Economics and the 21st Century*, edited by M. Szenberg, L. Ramrattan, and A. A. Gottesman. Oxford: Oxford Univ. Press.

Vermeulen, Frederic. 2005. “And the Winner Is . . .: An Empirical Evaluation of Unitary and Collective Labour Supply Models.” *Empirical Econ.* 30:711–34.