The Economics of the Family

Chapter 6: Empirical issues for the collective model

Martin Browning
Department of Economics, Oxford University

Pierre-André Chiappori
Department of Economics, Columbia University

Yoram Weiss
Department of Economics, Tel Aviv University

October 2007

1 What are the objects of interest?

We have seen above that various approaches can be used to describe household behavior, from the unitary setting to noncooperative approaches and the collective model. Ultimately, the choice between these various frameworks will rely on three types of properties. First, general methodological principles may favor one approach over the others. For instance, one can argue that the unitary framework is not totally faithful to methodological individualism, a cornerstone of micro theory that postulates that individuals, not groups, are the ultimate decision makers. A second requirement is the model's ability to generate testable predictions upon behavior, that can be taken to data using standard techniques. Standard consumer theory fares pretty well in this respect. Utility maximization under linear budget constraint has been showed to engender strong predictions (income pooling, Slutsky symmetry and negativeness), and adequate methodologies have been developed for testing these properties. Finally, a crucial criterion is the fruitfulness of the approach, particularly in terms of normative analysis and policy recommendations. A remarkable feature of standard consumer theory is that individual preferences can be uniquely recovered from demand functions (if these satisfy the Slutsky conditions); it is therefore possible to analyze welfare issues from the sole knowledge of observed behavior. A more general requirement is that the model be *identifiable*, that is, that it should be possible to recover the underlying structure from observed behavior.

The first line of argument, concerning methodological individualism, has been evoked earlier. In this chapter, we concentrate on the remaining two aspects, namely testability and identifiability of preferences and processes from observed behavior. Most of the existing knowledge concerns the cooperative framework, and especially the collective model, although the predictions generated by non cooperative settings will be briefly evoked in the last Section. The testability requirement, *per se*, is not problematic. The idea that a model should generate predictions that can be taken to data belongs to the foundations of economics. Identifiability is more complex; it may be useful to define more precisely what is meant by 'recovering the underlying structure'. The structure, in our case, is defined by the (strictly convex) preferences of individuals in the group and the decision process. In the collective setting, because of the efficiency assumption, the decision process is fully summarized (for any particular cardinalisation of individual utilities) by the Pareto weight corresponding to the outcome at stake. The structure thus consists in a set of individual preferences (with a particular cardinalisation) and a Pareto weight.

The structure cannot be directly observed; instead we observe the group's aggregate behavior. In practice, the 'observation' of, say, a demand function is a complex process, that entails specific difficulties. For instance, one never observes a (continuous) function, but only a finite number of values on the function's graph. These values are measured with some errors, which raises problems of statistical inference. In some cases, the data are cross-sectional, in the sense that different groups are observed in different situations; specific assumptions have to be made on the nature and the form of (observed and unobserved) heterogeneity between the groups. Even when the same group is observed in different contexts (panel data), other assumptions are needed on the dynamics of the situation. For example, on the way past behavior influences present choices. All these issues lay at the core of what is usually called the *identification* problem.

A second and different aspect relates to what has been called the *identifiability* problem, which can be defined as follows: when is it the case that the *(hypothetically) perfect knowledge* of a smooth demand function uniquely defines the underlying structure within a given class? In other words, non identifiability does not result from the econometrician's inability to exactly recover the form of demand functions - say, because only noisy estimates of the parameters can be obtained, or even because the functional form itself (and the stochastic structure added to it) have been arbitrarily chosen. These econometric questions have, at least to some extent, econometric or statistical answers. For instance, confidence intervals can be computed for the parameters (and become negligible when the sample size grows); the relevance of the functional form can be checked using specification tests; etc. The non identifiability problem has a different nature: even if a *perfect* fit to *ideal* data was feasible, it might still be impossible to recover the underlying structure from observed behavior.

In the case of *individual* behavior, as analyzed by standard consumer theory, identifiability is an old but crucial result. Indeed, it has been known for more than a century that an individual demand function uniquely identifies the underlying preferences. Usual as this property may have become, it remains one of the strongest results in microeconomic theory. It implies, for instance, that assessments about individual well-being can unambiguously be made from the sole observation of demand behavior - a fact that opens the way to all of applied welfare economics. It is thus natural to ask whether this classical identifiability property can be extended to more general approaches.[1]

[1] Note, however, that only one utility is identifiable in the standard case. A 'unitary'

Finally, it should be remembered that identifiability is only a necessary condition for identification. If different structures are observationally equivalent, there is no hope that observed behavior will help to distinguish between them - only ad hoc functional form restrictions can do that. Since observationally equivalent models may have very different welfare implications, non identifiability severely limits our ability to formulate reliable normative judgments: any normative recommendation based on a particular structural model is unreliable, since it is ultimately based on the purely arbitrary choice of one underlying structural model among many. Still, whether an *identifiable* model is econometrically *identified* depends on the stochastic structure representing the various statistical issues (measurement errors, unobserved heterogeneity,...) discussed above. After all, the abundant empirical literature on consumer behavior, while dealing with a model that is always identifiable, has convinced us that identification crucially depends on the nature of available data.

The main properties of the collective model have been described in the previous chapter. However, which empirical test can actually be performed obviously depends on the nature of available data. Three different contexts can be distinguished. In the first context, individual demand can be estimated as a function of income and possibly distribution factors; this approach is relevant when no price variation is observed, for instance because data are cross-sectional and prices are constant over the sample. We then allow that we also observe price variation so that we can estimate a complete demand system. The analysis of labor supply raises specific issues that are considered in the third section. The final half of this chapter presents a review of empirical analysis using non-unitary models (including the results of applying the tests of the first half of the chapter). We conclude the chapter with an account of intra-household allocation based on two Danish data sets that were specifically designed to address research issues concerning intrahousehold allocation.

2 Data without price variation.

2.1 Necessary and sufficient conditions for a collective model.

In this section we consider testing and identification in the absence of price variations. Let x denote the household's total expenditures, and let $\mathbf{z}$ be a

framework in which agents are each characterized by their own utility function, still the household behaves as a single decision unit (say, because Pareto weights are constant) is typically *not* identifiable (see below).

K-vector of distribution factors. Recall that distribution factors, by definition, influence neither preferences nor the budget constraint. In a unitary setting, they have no impact on demand. In the collective framework, on the contrary, household behavior can be described by a program of the following form:

$$\max \mu(x, \mathbf{z}) u^a(\mathbf{g}) + (1 - \mu(x, \mathbf{z})) u^b(\mathbf{g}) \quad (1)$$
$$\text{subject to } \mathbf{e}'\mathbf{g} \leq x$$

where $\mathbf{g}$ is the vector $(\mathbf{q}^a, \mathbf{q}^b, \mathbf{Q})$ and quantities are normalised so that the price vector is a vector of ones, $\mathbf{e}$. The resulting vector of *collective demand functions* can be written $\mathbf{g} = \tilde{g}(x, \mu(x, \mathbf{z}))$ with a corresponding observable demand functions $\hat{\mathbf{g}}(x, \mathbf{z})$. An alternative demand formulation which is useful for empirical work can be formulated if there is at least one good (good j, say) that is strictly monotone in one distribution factor (z_1, say); that is, $g_j(x, \mathbf{z})$ is strictly monotone in z_1. This demand function can be inverted on the first factor to give:

$$z_1 = \zeta(x, \mathbf{z}_{-1}, g_j)$$

where $\mathbf{z}_{-1}$ is the vector of distribution factors without the first element. Now substitute this into the demand for good i:

$$g_i = \hat{g}_i(x, z_1, \mathbf{z}_{-1}) = \hat{g}_i[x, \zeta(x, \mathbf{z}_{-1}, g_j), \mathbf{z}_{-1}] = \theta_i^j(x, \mathbf{z}_{-1}, g_j).$$

Thus the demand for good i can be written as a function of total expenditure, all distribution factors but the first and the demand for good j. To distinguish this conditioning from the more conventional conditional demands used in the literature, we shall refer to them as *z-conditional* demands. Note that, in the unitary setting, there are no distribution factors, so that z-conditional demands are not defined in this case.

We now address the issue of what restrictions a collective model impose upon observable demands. Bourguignon, Browning and Chiappori (2005) provide a complete characterization of these conditions. Specifically they prove that the following, equivalent conditions are necessary consequences of the collective model:

1. there exist real valued functions $\tilde{g}_1, \ldots\ldots, \tilde{g}_n$ and μ such that :

$$\hat{g}_j(x, \mathbf{z}) = \tilde{g}_i[x, \mu(x, \mathbf{z})] \qquad \forall i = 1, \ldots\ldots, n \quad (2)$$

2. household demand functions satisfy the *proportionality condition*:

$$\frac{\partial \hat{g}_i/\partial z_k}{\partial \hat{g}_j/\partial z_k} = \frac{\partial \hat{g}_i/\partial z_1}{\partial \hat{g}_j/\partial z_1} \qquad \forall i = 1,..,n;\; j = 1,..,n;\; k = 2,..,K \tag{3}$$

3. there exists at least one good j such that $\partial \hat{g}_j/\partial z_1 \neq 0$ and the z-conditional demands satisfy:

$$\frac{\partial \theta_i^j(x, \mathbf{z}_{-1}, q_j)}{\partial z_k} = 0 \qquad \forall i \neq j, i = 1,..,n; k = 2,..,K \tag{4}$$

Conversely, these conditions are also sufficient: if they are satisfied for the observable demands $\hat{g}(x, \mathbf{z})$, then one can find utility functions and Pareto weights which rationalise the observed demands. An important implication of these conditions is that in the absence of price variation, proportionality is the *only* testable implication of the collective model. This means that if we have only one distribution factor, then we can never reject the hypothesis of collective rationality. Any extra restrictions for a collective model require that *additional assumptions* be made on the form of individual preferences. For instance, restrictions exist even for a single distribution factor when some goods are private and/or consumed exclusively by one member of the household.

The intuition for this result relates to the discussion provided in earlier chapters. Again, the basic idea is that, by definition, distribution factors do not influence the Pareto set. They may affect consumption, but only through their effect upon the *location* of the final outcome on the Pareto frontier - or, equivalently, upon the *respective weighting* of each member's utility that is implicit in this location. The key point is that this effect is one-dimensional. This explains why restrictions appear only in the case where there are more than one distribution factors. Whatever the number of such factors, they can only influence consumption through a single, real-valued function μ. Conditions (2) and (3) are direct translations of this remark. By the same token, if we compute q_i as a z^1-conditional function of $(x, \tilde{g}_j, \mathbf{z}_{-1})$, it should not depend on $\mathbf{z}_{-1}$. The reason is that, for any given value of x, whenever distribution factors $(z_1, \mathbf{z}_{-1})$ contain some information that is relevant for intra-household allocation (hence for household behavior), this information is one-dimensional and can be *fully* summarized by the value of $\tilde{g}_j$. Once we condition on $\tilde{g}_j$, z_{-1} then becomes irrelevant. This is the meaning of condition (4).

This result provides two distinct ways of testing for efficiency. Condition (3) leads to tests of cross-equation restrictions in a system of unconditional demand equations. An alternative method, implied by (4), tests for exclusion restrictions in a conditional demand framework. Empirically, the latter is likely to be more powerful for at least two reasons. First we can employ single equation methods (or even non-parametric methods). Second, single equation exclusion tests are more robust than tests of the equality of parameters across equations. Finally, note that the tests generalize easily to a Beckerian framework where domestic goods produced by the household are taken into account. Adding a domestic production function to go from the market inputs to the goods actually consumed by household members and taking into account the allocation of domestic labor do not modify the above tests on household demands for market goods.

As discussed in Chapter 3, the bargaining version of the collective model has attracted lot of attention. A bargaining framework should be expected to impose *additional* restrictions to those discussed above. Indeed, an easy test can be described as follows. Assume that some distribution factors, which are part of a K'-sub-vector $\mathbf{z}'$, are known to be positively correlated with member b's threat point, while others, constituting a K''-sub-vector $\mathbf{z}''$, are known to favor a. Then in program (1) μ should decrease with distribution factors in $\mathbf{z}''$ and increase with those in $\mathbf{z}'$. This property can readily be tested; it implies that,

$$\frac{\partial \hat{g}_i/\partial z'_k}{\partial \hat{g}_i/\partial z''_m} = \frac{\partial \hat{g}_j/\partial z'_k}{\partial \hat{g}_j/\partial z''_m} \leq 0 \text{ for } i,j = 1,..,n; k = 1,...,K'; m = 1,...,K''$$

Should one be willing to go further and assume, for instance, that only the ratio z'_1/z''_2 of distribution factors matters, then we have in addition:

$$\frac{\partial \hat{g}_i}{\partial Log(z'_1)} + \frac{\partial \hat{g}_i}{\partial Log(z''_2)} = 0 \qquad \forall i = 1,..,n$$

This is simple to test and easy to interpret.

2.2 Identifiability.

A more difficult issue regards identification: when is it possible to recover the underlying structure from the sole observation of household behavior? Note, first, that the nature of the data strongly limits what can be recovered. For instance, one cannot hope to identify utility functions in the absence of price variations. 'Identifiability', in this context, essentially means recovering individual Engel curves (that is, demand as a function of income) and the decision process, as summarized by the Pareto weights or (in the private good case) by the sharing rule. When is it possible to recover that underlying structure from observable household behavior?

The following gives some useful mathematical result in our context. Suppose we have a smooth unknown function $f(x,y)$ with non-zero partials f_x and f_y. Suppose first that we observe:

$$g^1(x,y) = f_x(x,y) \text{ and } g^2(x,y) = f_y(x,y) \tag{5}$$

If $f(.)$ is twice continuously differentiable, these two functions must satisfy $g^1_y(x,y) = g^2_x(x,y)$. If this symmetry condition is satisfied then $f(.)$ is identifiable up to an additive constant. That is, any solution is of the form $f(x,y) + k$ where k is an arbitrary constant. Suppose now that rather than observing the partials themselves we only observe their ratio:

$$g(x,y) = \frac{f_y}{f_x} \tag{6}$$

Given $g(x,y)$, $f(x,y)$ is identifiable 'up to a strictly monotone transformation'. That is, we can recover some $\bar{f}(x,y)$ such that any solution is of the form $f(x,y) = G\left(\bar{f}(x,y)\right)$ where $G(.)$ is an arbitrary strictly monotone function.

In general, when $f(.)$ has more than two arguments, for example $f(x_1,...,x_n)$, assume that we observe $m < n-1$ ratios of partials, say those involving the $m+1$ first partials of f : $\frac{f_2}{f_1},...,\frac{f_{m+1}}{f_1}$. Then f is identifiable up to a function of the other variables; that is, we can identify some $\bar{f}(x_1,...,x_{m+1})$ such that any solution is of the form

$$f(x_1,...,x_n) = G\left(\bar{f}(x_1,...,x_{m+1}), x_{m+2}...,x_n\right)$$

where $G(.)$ is an arbitrary function. In particular:

- if we observe only one ratio of partials, say $g(x_1,...,x_n) = f_1/f_2$, then $f(.)$ is identifiable up to a function of the other variables $(x_3,...,x_n)$.
- if we observe all ratios of partials, then $f(.)$ is identifiable up to an arbitrary, strictly monotone transformation.

Note, as well, that whenever we observe more than one ratio of partials, testable restrictions are generated. These generalise the standard cross-derivative conditions.

We can now consider the identifiability problem in this context. Note, first, that even in the most general case (no identifying restriction beyond efficiency), some (but by no means all) of the structure can be recovered from the observation of demand functions. To see why, note that any solution to (1) must be of the form:

$$\hat{\mathbf{g}}(x, \mathbf{z}) = \tilde{\mathbf{g}}(x, \mu(x, \mathbf{z})) \tag{7}$$

It follows that:

$$\frac{\partial \hat{g}_i / \partial z_k}{\partial \hat{g}_i / \partial z_1} = \frac{\partial \mu / \partial z_k}{\partial \mu / \partial z_1} = \frac{\mu_k}{\mu_1} = \kappa_k \tag{8}$$

for all i and k. The left hand side expression is potentially observable so that we can identify the ratio of partials of $\mu(x, \mathbf{z})$ with respect to distribution factors. Note that since the right hand side does not depend on the good, the ratio on the left hand side must be the same for all goods; this is the proportionality condition. Given the ratio of partials of the Pareto weight, we can recover μ up to some function of x. That is, we can recover a particular Pareto weight $\bar{\mu}$ such the true Pareto weight μ must be of the form:

$$\mu(x, \mathbf{z}) = m(x, \bar{\mu}(x, \mathbf{z})) \tag{9}$$

for some function m.

A ratio such as κ_k has a natural interpretation in terms of *power compensation.* Assume, for instance, that $\mu_1 > 0$ and $\mu_k < 0$ so that z_1 favors b while z_k serves a. If z_1 is increased by some infinitesimal quantity dz then $\kappa_k dz$ is the increase in z_k required to offset the change and maintain the same balance of power. Power compensations may be important for welfare analysis, whenever a 'shift of power' has to be compensated. The good news is that even in the most general version of the collective model, they can be directly recovered from observed demands. Furthermore, the proportionality condition (3) imposes that the estimation of the power compensation ratio does not depend on the particular commodity chosen.

An alternative and important interpretation of this result is that the model always behaves 'as if' there were only one factor, $\bar{\mu}$, influencing the individual's relative powers. Whatever the actual number of distribution factors, they always operate through the index $\bar{\mu}$ .Moreover, this index is identifiable. What is not identifiable in the general case is the exact impact of the index over the actual Pareto weights - an impact that will in general depend on the level of total expenditures.

2.3 Private consumption.

Although useful, recovery of the Pareto weight up to a strictly monotone function that also depends on total expenditure is far short of what is needed for all purposes. Is it possible to recover more? To achieve this, we need either better data or more theory restrictions. Existing results for the latter deal with the particular but useful case in which all commodities are privately consumed and preferences are either egoistic or caring.[2] As we have seen in chapter 5, efficiency is then equivalent to the existence of a sharing rule ρ in which a receives this amount and b receives $(x-\rho)$. Individual a solves**:**

$$\max v^a\left(\mathbf{q}^a\right) \text{ subject to } \mathbf{e}'\mathbf{q}^a = \rho$$

and similarly for b. It follows that the household demand for commodity i takes the form:

$$q_i\left(x,\mathbf{z}\right) = q_i^a\left(\rho\left(x,\mathbf{z}\right)\right) + q_i^b\left(x-\rho\left(x,\mathbf{z}\right)\right)$$

where q_i^a is a's demand for good i. The question, now, is: what can be said about q_a^i, q_b^i and ρ from the observation of household demands q_i for $i = 1, ...n$.

Note, first, that (8) has an equivalent in this context; indeed, (**??**) implies that:

$$\frac{\partial g_i/\partial z_k}{\partial g_i/\partial z_1} = \frac{\partial \rho/\partial z_k}{\partial \rho/\partial z_1} = \frac{\rho_k}{\rho_1}$$

for all k. Again, the potential observability of the left hand side ratios means that we can recover the sharing rule up to an arbitrary monotone function of total expenditures x. In other words, we can recover some $\bar{\rho}\left(x,\mathbf{z}\right)$ such that the true sharing rule must be of the form $\rho\left(x,\mathbf{z}\right) = G\left(\bar{\rho}\left(x,\mathbf{z}\right),x\right)$ for some mapping G. And as above, instead of analyzing the impact of each distribution factor independently, we may just consider the impact of the 'index' $\bar{\rho}$. Technically, thus, we can always consider the case of a unique distribution factor; no loss of generality results.

2.4 Assignability.

Suppose now that we have extra information on the *assignability* of particular goods to the two partners.[3] For example, suppose we observe the individual consumption of the first good and can estimate $\hat{g}_1^a\left(x,z\right)$ and $\hat{g}_1^b\left(x,z\right)$.

[2] Note, however, that the results below remain valid in the presence of public goods, provided that the sharing rule is taken to be conditional (as described in subsection 5.2 of Chapter 4)..

[3] When there is no price variation we can also regard an assignable good as two *exclusive* goods - each one being consumed by one partner exclsuively. Hence the results of this subsection also apply to exclusive goods.

Assuming, without loss of generality (as argued above), that there is only one distribution factor we have:

$$\hat{g}_1^a(x,z) = \tilde{g}_1^a(\rho(x,z)) \tag{10}$$

so that

$$\frac{\partial \hat{g}_1^a/\partial x}{\partial \hat{g}_1^a/\partial z} = \frac{\rho_x}{\rho_z}$$

Similarly, for b, we have:

$$\frac{\partial \hat{g}_1^b/\partial x}{\partial \hat{g}_1^b/\partial z} = -\frac{1-\rho_x}{\rho_z}$$

Thus the two ratios ρ_x/ρ_1 and $(1-\rho_x)/\rho_1$ are identifiable. There is a unique solution to these two equations for (ρ_x, ρ_z) if and only if we have:

$$\Gamma = \frac{\partial \hat{g}_1^a}{\partial x}\frac{\partial \hat{g}_1^b}{\partial z} - \frac{\partial \hat{g}_1^b}{\partial x}\frac{\partial \hat{g}_1^a}{\partial z} \neq 0 \tag{11}$$

If this condition holds, we can identify the partials of ρ:

$$\begin{aligned}\rho_x &= \frac{1}{\Gamma}\frac{\partial \hat{g}_1^a}{\partial x}\frac{\partial \hat{g}_1^b}{\partial z} \\ \rho_z &= \frac{1}{\Gamma}\frac{\partial \hat{g}_1^a}{\partial z}\frac{\partial \hat{g}_1^b}{\partial z}\end{aligned} \tag{12}$$

By the result before (6), knowing the partials allows us to identify the function itself , up to an additive constant: $\rho = \rho(x,z) + k$. Thus we can learn everything about the sharing rule from observing the assignment of a single good, *except its location*.

Moreover, new restrictions are generated, since

$$\frac{\partial}{\partial z}\left(\frac{\partial \rho}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial \rho}{\partial z}\right)$$

Finally, what about individual demands? First, for any value of the constant, (10) and (**??**) identify individual demands for commodity 1. Consider, now, commodity $i > 1$. From:

$$\hat{g}_i(x,\mathbf{z}) = \tilde{g}_i^a(\rho(x,\mathbf{z})) + \tilde{g}_i^b(x - \rho(x,\mathbf{z})) \tag{13}$$

we have (denoting the derivative of $\tilde{g}_1^a(.)$ by $\tilde{g}_1^{a\prime}$):

$$\begin{aligned}\frac{\partial \hat{g}_i}{\partial x} &= \tilde{g}_i^{a\prime}\rho_x + \tilde{g}_i^{b\prime}(1-\rho_x) \\ \frac{\partial \hat{g}_i}{\partial z} &= \left(\tilde{g}_i^{a\prime} - \tilde{g}_i^{b\prime}\right)\rho_z\end{aligned} \tag{14}$$

Since the left hand side is observed and we have (ρ_x, ρ_z) we invert (so long as $\rho_z \neq 0$) and identify $\tilde{g}_a^i$ and $\tilde{g}_b^i$ up to an additive constant. We conclude that the presence of an assignable good is sufficient to identify (up to additive constant) the sharing rule and individual demands for each commodity, including the non assignable ones.

We thus get a lot of mileage from the presence of one assignable (or two exclusive) goods. Can we do without? Surprisingly enough, the answer is positive. Bourguignon, Browning and Chiappori (2005) prove the following, very strong result: if we observe household demand (as a function of total expenditures x and a distribution factor z) for at least three commodities, then 'generically' we can recover individual demands and the sharing rule up to the same additive constants as before *and* (this is the only twist) up to a permutation of a and b. Identifiability, here, is only 'generic'. It is indeed possible to construct examples in which it does not hold, but these examples are not robust. For instance, if individual demands and the sharing rule are all linear, identification does not obtain. However, adding quadratic terms is sufficient to guarantee identification except maybe for very specific values of the coefficients.

3 Observing price responses.

3.1 Testing the collective model

We now turn to the situation in which we observe prices as well as income and distribution factors. Then strong tests are available; moreover, the model can be proved to be identifiable under reasonably mild conditions. Again, we consider a two person household for expositional convenience. Tests of the most general form of the collective model are based on the fundamental SNR1 condition demonstrated in Chapter 4. Namely, the Slutsky matrix S must be of the form:

$$S = \Sigma + R \tag{15}$$

where Σ is symmetric, negative and R is of rank at most one.

Direct tests of (15) are not straightforward, because the theorem simply says that *there exists* such a decomposition. To construct a testable condition of the symmetry of Σ, consider the matrix M defined as:

$$M = S - S'$$

where S' is the transpose of S. Now, since Σ is symmetric:

$$M = R - R'$$

and since R is of rank (at most) 1, M is of rank (at most) 2. This property is easy to test, using either standard rank tests or more specific approaches. Note, however, that five commodities (at least) are needed for that purpose. The reason is that neither M nor S are of full rank. Indeed, a standard result of consumer theory, stemming from homogeneity and adding up, states that

$$\boldsymbol{\pi}' S = S \boldsymbol{\pi} = \mathbf{0}$$

where $\boldsymbol{\pi}$ denotes the price vector. It follows that $M\boldsymbol{\pi} = \mathbf{0}$, and M cannot be invertible. Moreover, M is antisymmetric (equal to minus its transpose); hence its rank must be even. With four commodities, M is a 4×4, antisymmetric, non invertible matrix, so that its rank cannot exceed 2 anyway.

Negativeness of Σ, on the other hand, can be directly tested on the Slutsky matrix. Indeed, among the eigenvalues of S, one is zero (reflecting non invertibility; among the others, one (at most) can be positive. In practice, such test may however not be very powerful. Note that while symmetry of Σ cannot be tested from less than five goods, three are sufficient to test negativeness.

Finally, distribution factors can readily be introduced. Using equation ? in Chapter 4, Browning and Chiappori (1998) prove the following result. Take any distribution factor k, and compute the vector $v' = \left(\frac{\partial \hat{g}_1}{\partial z_k}, ..., \frac{\partial \hat{g}_1}{\partial z_k}\right)$. Then replacing any column (or any row) of M with v should not increase the rank.

3.2 Identifying the collective model

In the presence of price variation, the identifiability problem can be stated in full generality; indeed, when price effects are observable it may be possible to recover individual utilities. Clearly, identifying assumptions are necessary; in its most general version (with general preferences $u^a\left(\mathbf{q}^a, \mathbf{q}^b, \mathbf{Q}\right)$ and $u^b\left(\mathbf{q}^a, \mathbf{q}^b, \mathbf{Q}\right)$), there exists a continuum of different structural models generating the same demand function. For instance, Chiappori and Ekeland (2005) show that any function satisfying SNR1 can be generated as the Pareto efficient demand of a household in which all consumption is public, and also of an (obviously different) household in which all consumption is private. Therefore, we assume in this subsection that preferences are egoistic ($u^a\left(\mathbf{q}^a, \mathbf{Q}\right)$ and $u^b\left(\mathbf{q}^b, \mathbf{Q}\right)$), although our results have implications for caring preferences as well.

Even with egoistic preferences, however, the collective structure cannot in general be fully identified from demand data. To give a simple counterexample, assume for a moment that all goods are publicly consumed and consider two pairs of utility functions, $\left(u^a(\mathbf{Q}), u^b(\mathbf{Q})\right)$ and $\left(\tilde{u}^a(\mathbf{Q}), \tilde{u}^b(\mathbf{Q})\right)$ with

$$\begin{aligned} \tilde{u}^a &= F\left(u^a, u^b\right) \\ \tilde{u}^b &= G\left(u^a, u^b\right) \end{aligned}$$

for two arbitrary, incr**easing functions F and G. It is easy to check that any allocation that is Pareto efficient for $\left(\tilde{u}^a, \tilde{u}^b\right)$ must be Pareto efficient for $\left(u^a, u^b\right)$ as well; otherwise one could increase u^a and u^b without violating the budget constraint, but this would increase $\tilde{u}^a$ and $\tilde{u}^b$, a contradiction. It follo**ws that any demand that can be rationalized by $\left(\tilde{u}^a, \tilde{u}^b\right)$ can also be rationalized by $\left(u^a, u^b\right)$ (of course, with different Pareto weights), so that the two structures are empirically indistinguishable. Since F and G are arbitrary, we are facing a large degree of indeterminacy.

A negative result of this type has a simple meaning: additional identifying hypothesis are required. If there **are** at least four commodities, then Chiappori and Ekeland (2006) prove the following results:

- If each household member is the exclusive consumer of (at least) one good, then generically one can exactly recover the collective indirect utility of each member (up to an increasing transform). For any cardinalisation of these utilities, Pareto weights are recovered.

- If all commodities are publicly consumed, identifying collective indirect utilities is equivalent to identifying individual utilities. With private consumptions, on the contrary, any given pair of collective indirect utilities is compatible with a continuum of combinations of individual utilities and (conditional) sharing rules. However, all these are welfare equivalent, in the sense that they generate the same welfare conclusions. **For instance, if a given reform is found, using a specific combination, to increase the welfare of a while decreasing that of b, the same conclusion will hold for all combinations.**

- Finally, if there is at least one distribution factor , the exclusivity restriction can be relaxed and identifiability obtains either with one exclusive good overall (instead of one per member), or with one assignable good.

Again, the result is generic, in the sense that it holds for 'almost all' structures. An interesting remark is that (non-generic) exceptions include the case in which Pareto weights are constant; in such a case, the collective indirect utilities are not identifiable in general.[4] **Ironically, this corresponds to the Samuelson justification of the unitary setting, in which a single, price-independent welfare index is maximized. From an identification viewpoint, adopting a unitary framework is thus a very inappropriate choice, since it forbids the identification of individual welfares.**

The general conclusion is that welfare relevant structure is indeed identifiable in general, provided that one can observe one exclusive consumption per member (or one overall with a distribution factor). However, identifiability fails to obtain in a context in which the household behaves as a single decision maker.

4 The case of labor supply

4.1 Egoistic preferences and private consumption

A large part of the empirical literature on household behavior is devoted to labor supply. The theory has been presented in Section 4 of Chapter 4; here we concentrate on the empirical implications. Most empirical works consider the simple setting with egoistic preferences and private consumption studies in Subsection 4.3 of Chapter 4. In this framework, results have been established by Chiappori (1988, 1992) and Chiappori, Fortin and Lacroix (2002). Regarding testability, strong implications can be derived, even in this three-commodity setting. Even more remarkable is the fact that the observation of individual labor supplies, as functions of wages, non labor income and distribution factors, allows us to identify the sharing rule up to an additive constant. We start from the two leisure demand equations

$$L^a = \Lambda^a\left(w^a, \rho\right) \tag{16}$$

$$L^b = \Lambda^b\left(w^b, y-\rho\right) \tag{17}$$

where Λ^a denotes the Marshallian demand for leisure by person a and ρ is a's share of full income (see section ? in chapter 5). A simple derivation

[4] This case is 'non generic' in the sense that in the set of continuous functions, constant functions are non generic.

gives

$$\begin{aligned}
\frac{\partial L^a}{\partial w^b} &= \frac{\partial \Lambda^a}{\partial \rho}\frac{\partial \rho}{\partial w^b}, \\
\frac{\partial L^a}{\partial y} &= \frac{\partial \Lambda^a}{\partial \rho}\frac{\partial \rho}{\partial y} \\
\frac{\partial L^a}{\partial z} &= \frac{\partial \Lambda^a}{\partial \rho}\frac{\partial \rho}{\partial z}
\end{aligned} \tag{18}$$

Assuming $\partial L^a/\partial y \neq 0$ this gives:

$$\frac{\partial L^a/\partial z}{\partial L^a/\partial y} = \frac{\partial \rho/\partial z}{\partial \rho/\partial y} \tag{19}$$

Similar conditions obtain for b:

$$\begin{aligned}
\frac{\partial L^b}{\partial w^a} &= -\frac{\partial \Lambda^b}{\partial y^b}\frac{\partial \rho}{\partial w^a}, \\
\frac{\partial L^b}{\partial y} &= \frac{\partial \Lambda^b}{\partial y^b}\left(1-\frac{\partial \rho}{\partial y}\right) \\
\frac{\partial L^b}{\partial z} &= -\frac{\partial \Lambda^b}{\partial y^b}\frac{\partial \rho}{\partial z}
\end{aligned} \tag{20}$$

so that

$$\frac{\partial L^b/\partial z}{\partial L^b/\partial y} = -\frac{\partial \rho/\partial z}{1-\partial \rho/\partial y} \tag{21}$$

For notational simplicity, let F^i denote the fraction $\frac{\partial L^i/\partial z}{\partial L^i/\partial y}, i = a, b$; note that F^i can in principle be observed (or estimated) as a function of $\left(w^a, w^b, y, z\right)$. Now, (19) and (21) can be solved in $\partial \rho/\partial z$ and $\partial \rho/\partial y$ so long as $F^b \neq F^a$:

$$\begin{aligned}
\frac{\partial \rho}{\partial y} &= \frac{F^b}{F^b - F^a} \\
\frac{\partial \rho}{\partial z} &= \frac{F^a F^b}{F^b - F^a}
\end{aligned}$$

The condition $F^b \neq F^a$ is almost always satisfied empirically. Indeed, in general an increase in non labor income benefits both members, so that both $\partial \rho/\partial y$ and $1 - \partial \rho/\partial y$ are positive; then F^b and F^a have opposite signs and cannot be equal.[5] We thus conclude that the partials of ρ with respect

[5].. unless they are both zero (i.e., unless the distribution factor has no impact on behavior): the result requires the existence of a meaningful distribution factor.

to income and distribution factor are identifiable. Moreover, strong testable restrictions are generated by the cross differential restrictions.

Finally, the first two equations of (19) and of (21) give respectively:

$$\begin{aligned}\frac{\partial \rho}{\partial w^b} &= \frac{\partial L^a/\partial w^b}{\partial L^a/\partial y}\frac{\partial \rho}{\partial y} = \frac{\partial L^a/\partial w^b}{\partial L^a/\partial y}\frac{F^b}{F^b - F^a} \text{ and} \\ \frac{\partial \rho}{\partial w^a} &= \frac{\partial L^b/\partial w^a}{\partial L^b/\partial y}\left(1 - \frac{\partial \rho}{\partial y}\right) = -\frac{\partial L^a/\partial w^b}{\partial L^a/\partial y}\frac{F^a}{F^b - F^a}\end{aligned}$$

The conclusion is thus that *the partial derivatives of the sharing rule can be exactly recovered* from the observation of the two labor supply functions. From the sole observation of labor supplies, one can recover the impact of wages, non labor income and distribution factors on the sharing rule. The sharing rule itself is identified up to an additive constant; however, this constant is welfare irrelevant, in the sense that changing the constant affects neither the comparative statics nor the welfare analysis derived from the model. Lastly, cross derivative restrictions generate additional testable predictions.

From the previous derivation, one might get the feeling that the presence of a distribution factor is needed for identifiability. This is actually not the case. The observation of individual labor supplies, as functions of wages and non labor income, are 'generically' sufficient to recover the sharing rule up to an additive constant (Chiappori 1988, 1992). However, identification is only generic in that case; moreover, it is arguably less robust, since it involves second derivatives of the labor supply functions.

4.2 Extensions

The model has been extended in various directions. First, while the assumption of a unique, Hicksian composite consumption good is standard in the labor supply literature, the model can address a more general framework. Chiappori and Ekeland (2006c) consider a model with two leisures and n consumption goods that are privately (but not exclusively) consumed by the members. The context is cross-sectional, in the sense that there is variation in wages but not in prices. They show that if one distribution factor (at least) is available, then it is possible to identify (again up to additive constants) not only the sharing rule but also the individual demands for all private commodities, as functions of wages and non-labor income. They conclude that in a collective model of consumption and labor supply estimated on cross sectional data, it is possible to recover the income and wage elasticities of individual demands for each good.

Secondly, the computations above rely on the assumption that labor supply is a continuous variable. This may fail to hold for two reasons. First, in some households one member may elect not to participate; in that case, the person's labor supply is at a corner solution equal to zero. Secondly, the structure of labor markets may put constraints on the number of hours supplied by individuals. For instance, the choice may be only between working part time, full time or not at all; then labor supply should rather be modelled as a discrete variable. Extensions of the previous model to such situations have been studied by Blundell et al. (2000) and Donni (2004).

Finally, the privateness assumption has also been criticized on two grounds. First, while some consumptions are indeed private, others are not. Children expenditures are a typical example of public goods within the household. Blundell, Chiappori and Meghir (2005) analyze a model similar to the previous one but for the consumption good, which is taken to be public. They show that, again, the model is identifiable from the observation of labor supply behavior. They show how their approach can be extended to household production under various specifications.

A second criticism concerns the private nature of individual leisure. It could indeed be argued that leisure is, to some extend, publicly consumed; after all, the utility I derive from my own free time may be higher when my spouse is available as well. The general insight, here, is that a model in which both members' leisure enter each individual utility is still identifiable, provided that some other commodities are exclusive (this is a consequence of the general identifiability results described in Section 2). Fong and Zhang (2001) analyze an interesting framework in which leisure is partly private and partly public; they show that one assignable good is sufficient for identification in the presence of a distribution factor.