Coherency and regularity of demand systems with equality and inequality constraints
van Soest, A.H.O.; Kapteyn, A.J.; Kooreman, P.

Publication date: 1994

Coherency and Regularity of Demand Systems with Equality and Inequality Constraints

by
Arthur van Soest, Arie Kapteyn and Peter Kooreman


Reprint Series no. 150
CENTER FOR ECONOMIC RESEARCH

Board
Harry Barkema
Helmut Bester
Eric van Damme, Chairman
Frank van der Duyn Schouten
Jeffrey James

Management
Jeffrey James (Director of Graduate Studies)
Arie Kapteyn (Scientific Director)
Marie-Louise Kemperman (Managing Director)

Scientific Council
Anton Barten
Eduard Bomhoff
Willem Buiter
Jacques Drèze
Jack Kleijnen
Theo van de Klundert
Jean-Jacques Laffont
Merton Miller
Piet Moerland
Philippe Naert
Pieter Ruys
Université Catholique de Louvain
Erasmus University Rotterdam
Yale University
Université Catholique de Louvain
Tilburg University
Tilburg University
Université des Sciences Sociales de Toulouse
University of Chicago
Tilburg University
Nijenrode University
Tilburg University

Residential Fellows
Hans Bloemen
Lans Bovenberg
Hans Carlsson
Jay Pil Choi
Jan Magnus
Andrew Mountford
Bezalel Peleg
Mark Steel
Frank Verboven
Oscar Volij
Karl-Erik Wärneryd
CentER
CentER, Erasmus University Rotterdam
Gothenburg University and Lund University
Columbia University
CentER, LSE
CentER
Hebrew University of Jerusalem
CentER/Department of Econometrics, Tilburg University
CentER
Hebrew University of Jerusalem
Stockholm School of Economics

Address : P.O. Box 90153, 5000 LE Tilburg, The Netherlands
Phone : +31 13 663102
Telex : 52426 kub nl
Telefax : +31 13 663066
E-mail : center@kub.nl

ISSN 0924-7874

1994
Coherency and regularity of demand systems with equality and inequality constraints

Arthur van Soest and Arie Kapteyn
Tilburg University, Tilburg, The Netherlands

Peter Kooreman
Wageningen University, Wageningen, The Netherlands

Received December 1989, final version received February 1992

In neoclassical structural models dealing with rationing, corner solutions, or nonlinear budget constraints, utility theory is more crucial than in traditional demand systems. If in these models negativity of the Slutsky matrix is violated, the models will in general not be coherent, in the sense that endogenous variables are not unambiguously determined. We show that not imposing coherency may yield inconsistent estimators. A general framework is sketched which allows for the analysis of the relation between coherency, regularity of preferences, and the stochastic specification of the model. We discuss sufficient conditions for regularity which can be imposed in estimation.

1. Introduction

Empirical researchers in the field of demand theory are becoming increasingly aware of the tight structure that may be imposed on their models by neoclassical theory. In the somewhat older literature on demand systems a typical approach would be to choose a particular representation of preferences and derive the corresponding demand functions. After introducing error terms, the system would next be estimated. In the estimation, restrictions from neoclassical theory might or might not be imposed. In either case authors often have tested the various Slutsky conditions for their particular empirical specification, with mixed results [e.g., Barten (1977)]. As noted by McElroy (1987) consistency of the model with neoclassical theory has mostly been studied for the systematic
part of the demand equations only, with a rather cavalier treatment of the error structure. Her own work is a notable exception in this respect.

Whether or not authors severely test neoclassical restrictions for their data, it seems fair to say that in a standard demand system the empirical specification is rather loosely connected with underlying theory. If the estimation results turn out to be inconsistent with a utility maximization hypothesis, the empirical model can still yield an adequate description of the data.

In models dealing with for example rationing, corner solutions, or nonlinear budget constraints, utility theory plays a more crucial role than in traditional demand systems. If regularity conditions are violated, then these models will in general not be coherent, in the sense that the endogenous variables are not determined unambiguously by the model, or, in other words, the reduced form is not well-defined. See, e.g., Heckman (1978), who refers to coherency as ‘the principal assumption’.

In the literature, coherency of some specific models has been analysed before. For example, Ransom (1987a) has noted that the demand system based on the quadratic utility function with nonnegativity constraints introduced by Wales and Woodland (1983) is coherent if the parameters satisfy certain conditions, which are closely related to global concavity of the corresponding expenditure function (i.e., ‘negativity’). Van Soest and Kooreman (1990) obtain a similar result for the indirect translog demand system with nonnegativity constraints introduced by Lee and Pitt (1986). Hausman (1985) and MaCurdy et al. (1990) note the importance of negativity for coherency in individual labour supply models with kinked budget constraints. They also show that imposing negativity implies a priori restrictions on income and wage elasticities of labour supply. Coherency of simultaneous linear models with inequalities has been analysed by Gourieroux et al. (1980). In their models, coherency conditions have the form of restrictions on parameter values only. In some of the models that we consider, conditions not only depend on fixed parameters, but also on the possible values of the exogenous variables and of the error terms.

In empirical applications, one possible approach is to ignore coherency conditions in estimating the model and check afterwards for which observations coherency or regularity conditions are satisfied. The first goal of this paper is to point out that this practice may lead to inconsistent estimators, at least if maximum likelihood is used. Let $\Theta$ be the space of parameters generating coherent models and let $\Theta^*$ be an extension of $\Theta$, including parameters which do not yield a coherent model. In section 2, we present an example of a bivariate Probit model, for which the likelihood function (defined on $\Theta$ only) has a natural extension to $\Theta^*$. If coherency is ignored, this extended function will be maximized on $\Theta^*$. We show that this yields an inconsistent estimator with probability limit outside $\Theta$, even though the true parameter vector belongs to $\Theta$. 
The intuitive explanation is that the sum of 'probabilities' of events, which are mutually exclusive if the model is coherent, may exceed one for parameter values outside $\Theta$.

Given the fact that without imposing coherency ML estimation is inappropriate, the problem in empirical work is to formulate necessary and sufficient conditions for coherency in a particular model. In a well-defined neoclassical model, imposition of all the regularity conditions from demand theory is sufficient for coherency. Most of the specifications of preference structures considered in the literature only satisfy regularity conditions locally, i.e., in some subset of quantity or price space. For many flexible systems, the relationship between the parameter values and this subset is far from obvious. See, e.g., Barnett and Lee (1985) and Barnett (1983). For other systems, such as the generalized McFadden cost function proposed by Diewert and Wales (1987), explicit expressions of demand functions or conditional demand functions (necessary in case of binding constraints) cannot be obtained.

The second goal of the paper is to provide a framework which can be used to impose regularity conditions in some ‘large enough’ region of quantity or price space. We discuss this first for the case of a standard, an inverse, and a conditional demand system (section 3). The regularity conditions imply that the set of feasible (fixed and random) parameters cannot be too large. On the other hand, an extra condition is introduced which implies that the parameter space must be large enough, since the model must be able to explain certain features of the data. This is particularly relevant if measurement or optimization errors are excluded. As an example, we discuss how to impose the conditions for the quadratic direct utility system.

Imposing coherency becomes even more crucial in models characterized by endogenously switching regimes, considered in section 4, of which the kinked budget constraint labour supply model is the most familiar example. Due to nonlinearities the coherency conditions of Gourieroux et al. (1980) do not apply. We propose to impose regularity conditions quite similar to those introduced in section 3, which are sufficient for coherency, and illustrate with an example. Section 5 concludes.

2. Incoherency and ML estimation

Although the requirement that a model should be coherent may appear self-evident, one may still ask whether imposition of coherency conditions is strictly necessary. After all, given that the data-generating process is coherent, one might hope that parameter estimates automatically converge to values which satisfy coherency conditions. The example below shows that this is not the case.
Consider the following simultaneous Probit model [see, e.g., Schmidt (1981)]:

\[ y_1^* = \beta_1 x + \gamma_1 y_2 + \epsilon_1, \]
\[ y_2^* = \beta_2 x + \gamma_2 y_1 + \epsilon_2, \]
\[ y_i = 1 \text{ if } y_i^* > 0, \]
\[ = 0 \text{ if } y_i^* \leq 0, \quad i = 1, 2. \]  

(1)

Here \( x \) is an (observable) exogenous variable, \( y_1^* \) and \( y_2^* \) are latent endogenous variables, \( y_1 \) and \( y_2 \) are observed, and the error terms \( \epsilon_1 \) and \( \epsilon_2 \) are normal with mean zero, unit variance, and zero covariance. \( \theta = (\beta_1, \beta_2, \gamma_1, \gamma_2) \in \Theta^* = \mathbb{R}^4 \) is a vector of unknown parameters.

This model is coherent if and only if \( \gamma_1 \gamma_2 = 0 \) [cf. Schmidt (1981)]. The probabilities of the four different outcomes are

\[ \Pr[y_1 = 0, y_2 = 0] = \Phi(-\beta_1 x) \Phi(-\beta_2 x), \]
\[ \Pr[y_1 = 0, y_2 = 1] = \Phi(-\beta_1 x - \gamma_1) \Phi(\beta_2 x), \]
\[ \Pr[y_1 = 1, y_2 = 0] = \Phi(\beta_1 x) \Phi(-\beta_2 x - \gamma_2), \]
\[ \Pr[y_1 = 1, y_2 = 1] = \Phi(\beta_1 x + \gamma_1) \Phi(\beta_2 x + \gamma_2). \]  

(2)

Here \( \Phi \) is the standard normal distribution function. The expressions on the right-hand sides of (2) can still be computed if \( \gamma_1 \gamma_2 \neq 0 \). Their sum equals one if and only if \( \gamma_1 \gamma_2 = 0 \). Thus incoherency can loosely be interpreted as 'probabilities do not sum to one'.

Let \( \Theta = \{(\beta_1, \beta_2, \gamma_1, \gamma_2); \gamma_1 \gamma_2 = 0\} \). If the model is coherent, i.e., \( \theta \in \Theta \), then (2) implies that the log-likelihood of a random sample \((y_1, x_1), \ldots, (y_N, x_N)\) can be written as

\[
L(\Theta) = \sum_{i \in \Theta_0} \log [\Phi(-\beta_1 x_i) \Phi(-\beta_2 x_i)] \\
+ \sum_{i \in \Theta_1} \log [\Phi(-\beta_1 x_i - \gamma_1) \Phi(\beta_2 x_i)] \\
+ \sum_{i \in \Theta_0} \log [\Phi(\beta_1 x_i) \Phi(-\beta_2 x_i - \gamma_2)] \\
+ \sum_{i \in \Theta_1} \log [\Phi(\beta_1 x_i + \gamma_1) \Phi(\beta_2 x_i + \gamma_2)].
\]  

(3)
Here \( I_{ij} = \{t \in \{1, \ldots, N\}; y_{1t} = i, y_{2t} = j\} \). If \( \emptyset \notin \Theta \), the expression in (3) can still be computed, although it is no longer a log-likelihood. Thus, (3) also defines a natural extension of the likelihood function from \( \Theta \) to \( \Theta^* \).

If the model is not coherent, the likelihood function is not defined. Thus, strictly speaking, ML estimation requires a priori imposition of coherency conditions. Still, in many examples the likelihood function has a natural extension to the set of parameters for which the model is not coherent. In this example, the extension is given by (3). If coherency is not imposed, parameters will be estimated by maximizing (3) over \( \Theta^* \). In the appendix we show for a specific example, computing the expectation of (3), that the resulting estimator is inconsistent. Whereas the true parameter vector is an element of \( \Theta \), the probability limit of the estimator is not.

This result shows that 'maximum likelihood' estimation is not appropriate if coherency is not guaranteed for all values in the parameter space on which the likelihood function is to be maximized. Even though the model is coherent for the true parameter values, the ML technique will yield inconsistent estimators and standard methods of statistical inference will, in large enough samples, lead to the conclusion that the model is not coherent.

One can argue that the example of the bivariate probit model is not appropriate, since in this example coherency is well-known to be an issue and coherency conditions are easy to impose. Thus one will not easily fail to impose coherency conditions in this specific example. In many, more complicated, models however, it is much less clear that coherency might be a problem, and it is also less clear how coherency conditions should be imposed. In such cases the temptation exists to estimate the model without paying attention to coherency. The example above demonstrates that such an approach may lead to inconsistent estimates and incorrect statistical inference.

A common practice is to estimate the parameters by maximum likelihood without imposing coherency or regularity conditions, and then count the number of observations for which regularity is violated, given the parameter estimates. If there are observations for which regularity does not hold, this leads to the conclusion that not everyone behaves according to the neoclassical assumptions underlying the model. The example shows that this conclusion is inappropriate if violation of regularity conditions also involves violation of coherency conditions. This argument however works only in one direction: If, given the parameter estimates, regularity conditions and thus coherency conditions are satisfied, then there is no reason to mistrust the estimates. In this case, although coherency was not a priori imposed, maximization of the (extended) likelihood over the extended parameter space yields the same results as maximization of the (actual) likelihood over the set of parameters for which the model is coherent.
3. Demand systems with fixed regimes

In this section we describe a general framework to study coherency in a standard demand system, an inverse demand system, and a conditional demand system. We impose consistency with utility-maximizing behaviour and conditions which guarantee that the dual approach is appropriate, i.e., Shephard's lemma and Roy's identity can be applied in order to derive the standard demand functions, starting from the expenditure function or the indirect utility function, respectively. We first introduce some notation and standard regularity conditions. Next we consider restrictions on the parameter space that can be imposed in estimation to ensure that the estimated system satisfies the regularity conditions. Finally, another condition, called 'external coherency', is introduced, which states that the model must be able to explain enough features of the data.

3.1. Regularity conditions

We assume that each individual maximizes utility subject to a linear budget constraint. Topics such as rationing, nonnegativity constraints, and kinked budget constraints are discussed below. We start from an indirect utility function \( u = v_\theta(p, y) \), \((p, y) \in \tilde{V}_\theta \subset \mathbb{R}^n \times \mathbb{R}_+ \), where \( p = (p_1, ..., p_n)' \) is a vector of prices of \( n \) commodities, \( y \) is income (or total expenditures on the \( n \) commodities), \( u \) is the utility level, and \( \theta \in \Theta \subset \mathbb{R}^m \) is a vector of (fixed or random) parameters.

Standard regularity conditions for \( \theta \in \Theta \) are [see, e.g., Barten and Böhm (1982)]:

A.1. \( v_\theta \) is homogeneous of degree zero: for all \((p, y) \in \tilde{V}_\theta \) and \( \lambda \in \mathbb{R}_+ \), \( (\lambda p, \lambda y) \in \tilde{V}_\theta \) and \( v_\theta(\lambda p, \lambda y) = v_\theta(p, y) \).

A.2. \( v_\theta \) is twice continuously differentiable with respect to prices and income and for all \((p, y) \in \tilde{V}_\theta \), \((\partial v_\theta/\partial y)(p, y) > 0 \).

Assumption A.2 implies that \( v_\theta \) is strictly increasing in \( y \) and allows for the introduction of the expenditure or cost function \( e_\theta \) on the set \( \tilde{E}_\theta = \{(p, v_\theta(p, y)) ; (p, y) \in \tilde{V}_\theta \} \). \( e_\theta \) is implicitly defined by

\[
v_\theta(p, e_\theta(p, u)) = u, \quad (p, u) \in \tilde{E}_\theta.
\]

The dual approach is only consistent with utility-maximizing behaviour if 'strict' concavity is guaranteed. More precisely: \( e_\theta \) is said to be regular at given
If the \( n \times n \) matrix \( (\partial^2 e_s/\partial p \partial p')(p, u) \) is negative semi-definite and of rank \( n - 1 \), \( e_s \) is said to be regular at \( (p, y) \in \tilde{V}_s \) if \( e_s \) is regular at \( (p, v_s(p, y)) \). The third regularity condition can now be stated as:

A.3. \( v_s \) is regular at all \( (p, y) \in \tilde{V}_s \).

In what follows we work with a convex subset \( V_s \) of \( \tilde{V}_s \), and we assume that A.1–A.3 are satisfied on \( V_s \). \( V_s \) is referred to as the regular set in \( (p, y) \)-space. Marshallian (uncompensated) demand functions are denoted by

\[
q = F_s(p, y), \quad (p, y) \in V_s,
\]

where \( q = (q_1, \ldots, q_n)' \) is a vector of (not necessarily nonnegative) quantities and the components of the vector-valued function \( F_s \), defined on \( V_s \), are, according to Roy’s identity, given by

\[
F_{s,i}(p, y) = -\frac{\partial v_s/\partial p_i}{\partial v_s/\partial y}(p, y), \quad i = 1, \ldots, n.
\]

The regular set in \( q \)-space, \( Q_s \subset \mathbb{R}^n \), is defined as

\[
Q_s = \{F_s(p, y); (p, y) \in V_s\}.
\]

The assumptions A.1–A.3 together with the convexity of \( V_s \) imply that \( F_s \) is homogeneous of degree zero and one-to-one from \( \{(p, 1) \in V_s\} \) onto \( Q_s \) [cf. Gale and Nikaido (1965)].

3.2. Parameterization and restrictions in the parameter space

Preference variation across individuals (or households) can be incorporated in the parameter vector \( \vartheta \). For each individual \( t \), we write

\[
\vartheta_t = g_t(\psi, \eta_t).
\]

Here \( \psi \) is a vector (or matrix) of fixed parameters (the same for all individuals) chosen from a set \( \Psi \), and the vectors \( \eta_t \) are independent drawings from some probability distribution which does not depend on \( t \). The (vector-valued) function \( g_t \) may depend on \( t \), e.g., through a vector \( x_t \) of observed individual characteristics. The most common example is \( \vartheta_t = g_t(\psi, \eta) = \psi x_t + \eta_t \), where \( \psi \) is a matrix of appropriate size. Thus, systematic preference variation is allowed for if \( g_t \) depends on \( t \), and the \( \eta_t \)'s reflect random variation of preferences.

In estimating the system of demand equations, the following conditions may be imposed on \( \Psi \) and/or on the support \( \Omega \) of the distribution of the \( \eta_t \)'s. The
conditions must guarantee regularity of preferences in all relevant points of price and/or quantity space. Here ‘relevant points’ include observed points, but may also include, e.g., points for which model simulations are performed. Which condition is appropriate depends on the type of model to be estimated. In each case, the conditions are sufficient but in general not necessary for coherency.

Condition B.1 is appropriate if the model of interest is a standard demand system, i.e., Marshallian demand functions are estimated and the linear budget constraint is the only binding constraint in the model:

\[ \text{B.1 [regularity in a minimal subset of \((p, y)\)-space]. For all } t, \text{ for all } \psi \in \Psi, \text{ and for all } \eta \in \Omega: V_{\theta, (\psi, \eta)} \supseteq V_{\min}. \]

This condition states that for all parameter values (and thus for all possible individual preference structures) the model must be able to explain behaviour for at least some minimal subset, \( V_{\min} \), of \((p, y)\)-space. This subset must contain all observed \((p, y)\)'s in the sample, in order to guarantee that the demand system is defined and regular at each data point.\(^1\) If the model is used for simulations with values of \((p, y)\) outside the sample range of \((p, y)\), then \( V_{\min} \) must contain the extra \((p, y)\) values also. Condition B.1 implies that \( \Theta \) cannot be too large; otherwise there might be values of \( \psi \) or \( \eta \), such that at some points of \( V_{\min} \), conditions A.1–A.3 are not satisfied. \( V_{\min} \) can, e.g., be some rectangle in \((p, y)\)-space.

Condition B.2 is the counterpart of B.1 in quantity space. It can be used, for example, if the model of interest is an inverse demand system, i.e., consists of inverse Marshallian demand functions. See, e.g., Anderson (1980) for some theoretical properties, and Barten and Bettendorf (1989) for an empirical application. In empirical practice, there are not many situations in which estimation of an inverse demand system is relevant. Condition B.2, however, also appears to be important in models with switching regimes and inequality constraints. See section 4. The condition states that, for given (fixed and random) parameter values, certain quantity vectors must be optimal for some prices and income. As with B.1 this implies that the parameter space cannot be too large.

\(^1\) In some cases, it may be useful to allow \( V_{\min} \) to depend on \( t \). For example, it may be the case that certain values of \((p, y)\) are only observed for individuals with specific characteristics \( x_t \).
B.2 (regularity in a minimal subset of q-space). For all \( t \), for all \( \psi \in \Psi \), and for all \( \eta \in \Omega \): \( Q_{g_t(\psi, \eta)} \supseteq Q_{\min} \).

Condition B.2 is illustrated in fig. 2. For given parameter values \( g_t(\psi, \eta) \), the quantity space consists of three parts: the area where the direct utility function is not defined (because shadow prices do not exist) \( (QN) \), the area where indifference curves exist but are not convex \( (QI) \), and the regular area \( Q_{g_t(\psi, \eta)} \). The condition states that parameters have to be restricted such that \( Q_{\min} \) is contained in \( Q_{g_t(\psi, \eta)} \).

Conditions B.1 and B.2 are similar in the sense that they both define an area in q-space where indifference curves must be convex. However, since restrictions are imposed a priori, i.e., before parameters are estimated, it is not possible to tell which point in q-space corresponds to a given \((p, y)\). Thus, if we choose a particular \( V_{\min} \) and estimate the parameters imposing B.1, it may turn out that indifference curves are convex in an area quite different from the \( Q_{\min} \) we had in mind. Similarly, a choice of \( Q_{\min} \) and imposition of B.2 may actually imply concavity on an area in \((p, y)\)-space different from \( V_{\min} \). This point is illustrated in the example at the end of this section.

Before we introduce a condition similar to B.1 and B.2 which is useful in a conditional demand system, we first present a simple example to show why an explicit condition for this case is necessary. The example shows that it is not sufficient to impose B.1 and/or B.2, since it is necessary to take explicit account of the way in which \((p, y)\)-space and q-space are related.
3.3. An example with rationing

Consider the following Gorman polar form expenditure function for three goods, defined for $p_i > 0$ and $a_i > 0$ $(i = 1, 2, 3)$:

$$e(u, p_1, p_2, p_3) = -\frac{1}{2}(p_2^2/p_3) \exp(p_1/p_3) - p_3 \exp(p_2/p_3)$$

$$+ \sum_{i=1}^{3} a_i p_i + u p_3.$$ (4)

The $2 \times 2$ submatrix of second-order derivatives with respect to $p_1$ and $p_2$ is

$$-\frac{1}{p_3} \begin{bmatrix} \frac{1}{2} v_1^2 \exp(v_1) & v_2 \exp(v_1) \\ v_2 \exp(v_1) & \exp(v_1) + \exp(v_2) \end{bmatrix},$$

where $v_1 = p_1/p_3$ and $v_2 = p_2/p_3$. This matrix is negative definite for $v_1 < v_2$. The demand functions for goods 1 and 2 are

$$q_1 = -\frac{1}{2} v_1^2 \exp(v_1) + a_1,$$ (5a)

$$q_2 = -v_2 \exp(v_1) - \exp(v_2) + a_2.$$ (5b)
Demand for goods 1 and 2 does not depend on income. Suppose now good 1 is rationed at \( q_1 = \bar{q}_1 \). Then \( q_2 \) can be obtained by first solving \( v_1 \) from (5a), for given \( v_2 \) and \( q_1 = \bar{q}_1 \), and then inserting the solution (\( \bar{v}_1 \), say) in (5b). Let us assume that \( \bar{q}_1 = -1 + a_1 \). This value is feasible; it is generated by, for instance, \( (v_1, v_2) = (\log 2, 1) \), which satisfies negativity, since \( v_1 < v_2 \).

Now assume however that \( \bar{q}_1 = -1 + a_1 \) and \( v_2 = \frac{1}{2} \). Then \( \bar{v}_1 = \log 8 > v_2 \). Hence, there is no shadow price \( \bar{v}_1 \) for which the expenditure function is concave, even though \( \bar{q}_1 \) and \( v_2 \) are both feasible. It is the combination \( \bar{q}_1 = -1 + a_1 \) and \( v_2 = \frac{1}{2} \) which creates the problem.

This example shows that the relationship between negativity and existence of a feasible solution of the rationing problem is not straightforward. It is not enough to know the regions in \((p, y)\)- and \(q\)-space where concavity holds. In case of rationing, part of the price vector and part of the quantity vector are given. Conditional and inverse conditional demand functions are necessary to determine the quantity vector and the price vector corresponding to the given mixed price and quantity vector, and are therefore needed to determine whether the given mixed vector is feasible.

Since the example above shows that regularity in a conditional demand system cannot be stated in terms of B.1 and B.2, we need an alternative condition (B.3). First, we need some more notation.

Conditional demand means utility maximization subject to both the budget constraint and an a priori given set of equality constraints on quantities. We assume that \( q_1, \ldots, q_k \) are constrained, whereas the other quantities can be chosen freely. The number \( k \) and the order of the goods may vary across individuals. The individual solves the problem

\[
\max_{q_{II}} U_\delta(\tilde{q}_I, q_{II}) \quad \text{s.t.} \quad y = p'_I \tilde{q}_I + p'_{II} q_{II},
\]

Here \( q = (q'_I, q_{II})' \), \( p = (p'_I, p'_{II})' \), and the constraints are \( q_I = \tilde{q}_I \), where \( q_I = (q_1, \ldots, q_k)' \), etc.

Starting from \( v_2 \) and \( F_2 \) as introduced above, the maximization problem can be solved using shadow prices: Find \( \tilde{p}_I \in \mathbb{R}^k \), \( \tilde{y} \in \mathbb{R} \), and \( q_{II} \in \mathbb{R}^{n-k} \), such that

\[
\begin{align*}
\left\{ (\tilde{p}_I, p_{II}), \tilde{y} \right\} & \in V_2 \\
F_2((\tilde{p}_I, p_{II}), \tilde{y}) & = (\tilde{q}_I, q_{II}) \\
\tilde{y} & = y + (\tilde{p}_I - p_I)\tilde{q}_I
\end{align*}
\]

(6)

The optimal quantities, subject to \( q_I = \tilde{q}_I \), are then given by \( q_{II} \).
We can now formulate condition B.3:

B.3. Let, for each \( t \), \( VQ_t \) be a given subset of \( \{(p, y, q_t) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \} \). Then for all \( \psi \in \Psi \), for all \( \eta \in \Omega \), for each \( t \), and for all \( (p, y, q_t) \in VQ_t \), there must exist \( \tilde{p}_t \in \mathbb{R}^n \), \( \tilde{y} \in \mathbb{R} \), and \( q_{ht} \in \mathbb{R}^{n-2} \), such that

\[
\begin{align*}
((\tilde{p}_t, p_{ht}), \tilde{y}) &\in V_{g_t(\psi, \eta)} \\
F_{g_t(\psi, \eta)}((\tilde{p}_t, p_{ht}), \tilde{y}) &= (q_{ht}, q_{ht}) \\
\tilde{y} &= y + (\tilde{p}_t - p_t) q_t
\end{align*}
\]

Condition B.3 states that, for each parameter vector, the solution must have a regular solution. It is necessary to guarantee that the domain of the conditional demand functions is large enough and thus sufficient for coherency of the conditional demand system: it states that to each \( \psi \in \Psi \) and each \( \eta \in \Omega \), there corresponds at least one vector \( q_{ht} \) of endogenous variables, and together with concavity of \( \epsilon_9 \) and convexity of \( V_9 \) this implies coherency.

If there are no measurement errors on prices, income, or rationed quantities, then \( VQ_t \) must at least contain the observed \((p_t, y_t, q_t)\).

Conditions B.1, B.2, and B.3 are similar in the sense that they all impose restrictions in a subspace of \((p, y, q)\)-space which must hold for all parameters, and thus imply that the parameter space cannot be too large. Which conditions are useful not only depends on the model to be estimated, but also on what the model is used for. For example, if a conditional demand system is used for simulations in which quantity constraints are relaxed, then both B.3 and B.1 must be imposed.

B.1–B.3 imply that \( \Theta \) cannot be too large and thus limit the flexibility of the error structure of the model. However, particularly if optimization and measurement errors are excluded, the specification of the model must also ensure that the model can explain observed behaviour. In other words, likelihood contributions must be nonzero. This implies that the parameter space cannot be too small. For the case of a standard or inverse demand system, this condition can be formulated as follows:

B.4 ('external coherency'). Let, for all \( t \), \( VQ_t \) be a given subset of \( \{(p, y, q) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \}; p'q = y) \}. Then for all \( t \), for all \( \psi \in \Psi \), and for all \( (p, y, q) \in VQ_t \), there must be an \( \eta \in \Omega \) such that \( F_{g_t(\psi, \eta)}(p, y) = q \).

One can think of \( VQ_t \) as the set of prices, incomes, and quantities which may arise for observation \( t \). If no measurement or optimization errors are involved, \( VQ_t \) must at least contain the observed \((p, y, q)\)-vector for individual \( t \). In fact, \( VQ_t \) may then consist of one element only. B.4 states that random preferences \( \eta \) must guarantee so much flexibility that, for all \( \psi \in \Psi \) and at least one possible
value of \( \eta \), a given (observed) quantity vector is optimal for given prices and income. This motivates the term 'external coherency': The model has to be coherent with available data, in the sense that the likelihood contribution of any given data point must be positive. Rather than by imposing this condition, this may also be achieved by explicit incorporation of measurement or optimization errors in the model.

Note the conflicting nature of B.1 and B.2 on the one hand and B.4 on the other hand. B.1 and B.2 may for instance imply such strong restrictions on \( \Omega \) that, for a given observed \((p, y)\) (and for given \(t\) and \(\psi\)), demand \(F_{g_1(\psi, \eta)}(p, y)\) equals the observed \(q\) for no \(\eta \in \Omega\), so that B.4 is not satisfied. This would be due to the fact that B.1 and B.2 imply that, for given \(\Psi\), \(\Omega\) cannot be too large, whereas B.4 implies that it cannot be too small.

For the case of a conditional demand system, an 'external coherency' condition quite similar to B.4 can be formulated:

B.5 ('external coherency'). Let \(VQ\), be a given subset of \(\{(p, y, q) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n; p_1' q = y\}\). Then for all \(\psi \in \Psi\) and \((p, y, q) \in VQ\), there must be at least one \(\eta \in \Omega\) such that there are \(\tilde{p}_1 \in \mathbb{R}^n\) and \(\tilde{y} \in \mathbb{R}\) with

\[
\begin{cases}
((\tilde{p}_1, p_{11}), \tilde{y}) \in V_{g_1(\psi, \eta)} \\
F_{g_1(\psi, \eta)}(\tilde{p}_1, p_{11}, \tilde{y}) = q \\
\tilde{y} = y + (\tilde{p}_1 - p_1)'q_{11}
\end{cases}
\]

Condition B.5 states that certain quantity vectors \(q_{11}\) can be optimal for given prices, income, and rationed quantities \(q_{1}\). If the model contains no optimization or measurement errors, \(VQ\), must at least contain the observed vector \((p_1, y_1, (\tilde{q}_{11}, q_{11}))\). It guarantees, similar to B.4, that each data point has a non-zero likelihood contribution. Note however that B.5 is weaker than B.4, since quantities \(\tilde{q}_{11}\) do not have to be rationalized.

Note the difference between B.5 and B.3: B.3 states that for each \(\eta \in \Omega\) the rationed utility maximization problem has some regular solution. B.5 states that to each observed optimum there must correspond some \(\eta \in \Omega\).

3.4. Example: Quadratic direct utility (QDU)

The direct utility function is given by

\[
U(q) = \gamma'q - \frac{1}{2}q'Bq,
\]

where \(\gamma = (\gamma_1, \ldots, \gamma_n)'\) and \(B\) is a positive definite \(n \times n\) matrix with entries \(\beta_{ij}\)
(i, j = 1, ..., n). U has a satiation point at \( q = B^{-1}y \), with corresponding utility level \( \frac{1}{2}y'B^{-1}y \). Marshallian demand is given by

\[
q = B^{-1}y - (p'B^{-1}p)^{-1}\{y'B^{-1}p - y\}B^{-1}p,
\]

and indirect utility is thus given by

\[
v(p, y) = \frac{1}{2}\{y'B^{-1}y - (p'B^{-1}p)^{-1}[y'B^{-1}p - y]^2\}.
\]

\( v \) is increasing in \( y \) as long as the satiation point is not in the budget set, i.e., as long as \( y < \gamma'B^{-1}p \). Homogeneity of degree zero is automatically guaranteed. The expenditure function is given by

\[
e_{B,y}(p, u) = \gamma'B^{-1}p - (p'B^{-1}p)^{1/2}[\gamma'B^{-1}p - 2u]^{1/2}.
\]

The Hessian of \( e_{B,y} \) is

\[
(\partial^2 e_{B,y}/\partial p \partial p')(p, u) = (p'B^{-1}p)^{1/2}[\gamma'B^{-1}p - 2u]^{1/2}
\]

\[
\times [(p'B^{-1}p)^{-1}(B^{-1}p)(B^{-1}p)' - B^{-1}].
\]

As expected, \( e_{B,y}(p, u) \) is only defined for \( u \leq \frac{1}{2}y'B^{-1}y \) (the satiation level) and \( p \neq 0 \). It is easy to show that, since \( B \) is positive definite, the matrix

\[
(p'B^{-1}p)^{-1}(B^{-1}p)(B^{-1}p)' - B^{-1}
\]

is negative semidefinite and of rank \( n - 1 \). Hence, the expenditure function is concave for all \( u \leq \frac{1}{2}y'B^{-1}y \).

In what follows, we assume that there is one commodity, say the \( n \)th, for which the price is always positive: \( p_n > 0 \). This suggests the following choice of \( V_{B,y} \):

\[
V_{B,y} = \{(p, y); y < \gamma'B^{-1}p, p_n > 0\}.
\]

We consider the following stochastic specification [cf. Ransom (1987b)]:

\[
\gamma_t = \gamma_{t,0} + \eta_t,
\]

where \( \gamma_{t,0} = (\gamma_{t,0}, \ldots, \gamma_{m,0})' \) is fixed and \( \eta_t = (\eta_{t1}, \ldots, \eta_{tn})' \) is random with \( \eta_{tn} = 0 \). The elaboration of B.1, B.2, and B.4 is now as follows.
B.1. \((y_{t,0} + \eta)'B^{-1}p - y > 0\) for all \(t, \eta \in \Omega\), and \((p, y) \in V_{\min}\).

If, for instance, \(V_{\min}\) is defined as the rectangle \(\{(p, y) \in \mathbb{R}^n \times \mathbb{R}; 0 \leq p \leq \bar{p}, 0 \leq y \leq \bar{y}\}\), for given \(\bar{p}, \bar{y}\), then B.1 implies

\[
\min_{0 \leq p \leq \bar{p}} \eta' B^{-1} p \geq \bar{y} - \min_{0 \leq p \leq \bar{p}} \min_{1 \leq i \leq n} y_{i,0} B^{-1} p \quad \text{for all } \eta \in \Omega. \tag{8}
\]

B.2. Inversion of the demand system for given parameter values (including \(y\)) yields shadow prices and corresponding virtual income as a function of \(y\):

\(p = \lambda(y - Bq)\) and \(y = p'q\), where \(\lambda\) can be chosen arbitrarily. The solution \((p, y)\) is in \(V_{B,y}\) if and only if \(\lambda > 0\) and \(y_n - (Bq)_n > 0\). Thus, imposition of regularity in a given region \(Q_{\min}\) in \(q\)-space yields \(y_n - (Bq)_n > 0\) for all \(q \in Q_{\min}\). This can be achieved by restricting the values of fixed parameters only, since \(y_n\) is nonrandom. Truncation of the distribution of \(\eta\) is unnecessary. If, for instance, \(Q_{\min}\) is some rectangle, \(2^n\) linear inequality restrictions on the coefficients in \(B\) and \(y_{i,n,0}\) result for each individual \(t\), which simultaneously must be imposed in estimation.

B.4. Let \(V_{Q,t} = \{(p, y, q, \eta)\}\). Solving \(\eta\) from the demand functions yields

\[
\eta_{ik} = \frac{p_{ik}}{p_{in}} y_{i,n} - y_{i,0} + \sum_{j=1}^{n} \left\{ \beta_{ij} - \frac{p_{ij}}{p_{in}} \beta_{nj} \right\} q_{ij},
\]

and \(\Omega\) should be big enough to contain these \(\eta\)'s for all \(t\), and for all \(y_{i,0}\) and \(\beta\) in the parameter space.

To get some more feeling for these conditions, we look at a simple numerical example for two commodities. Let \(n = 2\), \(B\) the \(2 \times 2\) identity matrix, and \(y_{i,0} = y_0\), fixed and independent of \(t\).

B.1. Let \(V_{\min}\) be a rectangle in normalized \((p, y)\)-space, i.e., \(V_{\min} = \{(p, y); 0 < v_1 \leq p_1/y \leq v_u\} \) and \(0 < v_2 \leq p_2/y \leq v_u\). Since \(p_2 > 0\), we use the normalization \(p_2 = 1\). \(V_{\min}\) can now be written as \(\{(p_1, y); v_1 y \leq p_1 \leq v_u y\}\). A feasible \(y = (y_1, y_2)'\) has to satisfy \(y_1 p_1 + y_2 - y > 0\) for all \((p_1, y) \in V_{\min}\). Thus \(y\) is feasible if \(y_2 > v_1^{-1} - y_1, y_2 > v_u^{-1} - y_1 v_1/y_u,\) and \(y_2 > v_2^{-1} - y_1 v_1/y_u\).

Fig. 3 presents the feasible area \((FAI)\) in \((y_1, y_2)\)-space. In this example (i.e., for this choice of \(B\)) \(FAI\) is nonempty for every \(v_u \geq v_1 > 0\). For each feasible \(y\) it is possible to derive the regular area in \((p_1, y)\)-space: \(V_y = \{(p_1, y); y_1 p_1 + y_2 - y > 0\}\). The intersection of the \(V_y\)'s is the region in \((p_1, y)\)-space, where the indirect utility function is regular for all \(y \in FAI\): \(V = \bigcap_{y \in FAI} V_y\).
In fig. 4, $V$ and $V_{\text{min}}$ are sketched. Note that automatically $V_{\text{min}} \subset V$, but the figure shows that $V$ is larger than $V_{\text{min}}$. If $V$ instead of $V_{\text{min}}$ was chosen to begin with, the same region $FA1$ in $\gamma$-space would have been obtained.

**B.2.** For given $\gamma$, the regular region in $q$-space is $Q_\gamma = \{(q_1, q_2) \in \mathbb{R}^2; q_2 < \gamma_2\}$. Thus, regularity on some region $Q_{\text{min}} \subset \{(q_1, q_2) \in \mathbb{R}^2; q_2 \leq \bar{q}_2\}$ is guaranteed if $\gamma_2$ is larger than $\bar{q}_2$, and does not impose restrictions on $\gamma_1$. The feasible area in $\gamma$-space is given by $FA2 = \{(\gamma_1, \gamma_2); \gamma_2 > \bar{q}_2\}$. 

---

**Fig. 3.** The feasible area in $\gamma$-space.

**Fig. 4.** The minimal and the actual regular region in $(p_1, \gamma)$-space.
Note that B.2 cannot replace B.1: the region in \((p, y)\)-space where regularity is guaranteed for all \(y \in FA2\), is given by

\[
V' = \bigcap_{y \in FA2} V_{y} = \{(p_1, y); \gamma_1 p_1 + \gamma_2 - y > 0 \text{ for all } y \in FA2\}
\]

\[
= \{(p_1, y); p_1 = 0, y \leq \hat{q}_2\}.
\]

Thus, for any \(V_{min}\) containing only nonzero \(p_1\)'s, it is impossible to choose a rectangular \(Q_{min}\) such that imposition of B.2 on \(Q_{min}\) implies that B.1 holds on \(V_{min}\). Following the same argument, it can be shown that B.1 cannot replace B.2.

B.4. For fixed \(y_2 = y_{2,0}\) and given \(p_1, y, (p_2 = 1), q_1\) and \(q_2\) with \(y = p'q\), we must find a feasible solution for \(y_1\) from the demand system

\[
q_1 = \gamma_1 - (1 + p_1^2)^{-1} \{\gamma_1 p_1 + \gamma_2 - y\} p_1,
\]

\[
q_2 = \gamma_2 - (1 + p_1^2)^{-1} \{\gamma_1 p_1 + \gamma_2 - y\} p_2.
\]

This is a system of two linearly dependent equations in \(y_1\) with a unique solution \(y_1 = q_1 + p_1(y_2 - q_2)\). The solution is feasible if and only if

\[
(p_1 + 1)\gamma_2 > v_1^{-1} - q_1 + p_1 q_2,
\]

\[
(p_1 + v_1/v_2)\gamma_2 > v_1^{-1} - q_1 + p_1 q_2,
\]

\[
(p_1 + v_1/v_2)\gamma_2 > v_1^{-1} - q_1 + p_1 q_2.
\]

If sample prices \(p_1\) always exceed \(-v_1/v_2\), then it is possible to guarantee the existence of a feasible solution for all \((q_1, q_2, y)\) in the sample by restricting \(y_2\) to be large enough.

To analyse B.3 and B.5, we need the conditional demand equations. We assume that no rationing applies to the \(n\)th commodity. Solving

\[
\max_{q_{II}} I_{II}(\tilde{q}_{II}, q_{II}) = (y'_1, y'_{II}) \left[ \begin{array}{c} \tilde{q}_I \\ \tilde{q}_{II} \end{array} \right] - \frac{1}{2} (\tilde{q}_I, q'_{II}) \left[ \begin{array}{cc} B_{11} & B_{12} \\ B_{12} & B_{22} \end{array} \right] \left[ \begin{array}{c} \tilde{q}_I \\ q_{II} \end{array} \right]
\]

s.t. \(y = p'_I \tilde{q}_I + p'_{II} q_{II}\),
A. van Soest et al., Coherency and regularity of demand systems

yields

\[ q_{11} = B_{22}^{-1} (\gamma_{11} - B_{12} \bar{q}_1) - (p_{11} B_{22} B_{12})^{-1} \]

\[ \times [(\gamma_{11} - B_{12} \bar{q}_1) B_{22} p_{11} - y - p_{1} \bar{q}_1] B_{22} p_{11}, \]  

with obvious partitioning of \( y \) and \( B \). The solution is feasible if \((\bar{q}_1, q_{11})\) is in the regular area of \( q \)-space, i.e., if \( \bar{q}_1 B_{22} p_{11} - y \in \{ p_{1} - B_{12} B_{22} p_{11} \} \geq 0 \).

Note that (9) has the same functional form as Marshallian demand with \( y \) replaced by \( y - p_{1} \bar{q}_1, \gamma \) by \( (\gamma_{11} - B_{12} \bar{q}_1), B \) by \( B_{22} \), and \( p \) by \( p_{11} \).

The elaboration of conditions 6.3 and B.5 is as follows:

**B.3.** Let \( VQ_1^i = \{(p_i, y_i, q_{1i})\} \). Existence of a feasible solution for given \( \gamma = (\gamma_{11}, \gamma_{11}) \) means

\[ \gamma_{11} B_{22} p_{11} - y_i + \bar{q}_i (p_{11} - B_{12} B_{22} p_{11}) > 0. \]

Substitution of \( \gamma = \gamma_{i,0} + \eta \) restricts the set \( \Omega \) of possible \( \eta \)'s.

**B.5.** Let \( VQ_1 = \{(p_i, y_i, q_i)\} \), with \( p_{11} > 0 \) and \( y_{i} = p_i q_i \). Solving \( \gamma \) from (9) yields

\[ \gamma_{11} = B_{12} q_{11} + B_{22} q_{11} - \lambda_i p_{11}, \]  

where

\[ \lambda_i = \left\{ - \gamma_{i,0} + [B_{12} q_{11} + B_{22} q_{11}] \right\} / p_{11}. \]  

The solution is feasible if and only if \( \lambda_i > 0 \). \( \Omega \) has to be large enough to contain at least one value of \( \eta \) such that \( \gamma_i = \gamma_{i,0} + \eta \) is feasible and satisfies (10), with \( \lambda_i \) given by (11).

4. Demand systems with endogenously switching regimes

In this section we consider the problem of an individual who maximizes utility subject to a set of linear inequality constraints. Common examples are the case of nonnegativity constraints [see, e.g., Wales and Woodland (1983), Lee a.,d Pitt (1986), Ransom (1987a), Van Soest and Kooreman (1990)] and the kinked budget set in labour supply models [Hausman (1985), Moffitt (1986)]. Regularity properties of the flexible form system introduced by Hausman and Ruud (1984) and the way to impose them in practice are discussed in Kapteyn et al. (1990). In contrast to the discussion in the previous section, we now assume that
it is not known in advance which constraints are binding and which are not. In other words, the 'regime' is endogenous.

The utility maximization problem in its primal form can be written as

$$\max_{q} U(q) \quad \text{s.t.} \quad Rq \leq r. \quad (12)$$

Here \(k\) is the number of restrictions, including the budget constraint, \(R\) is a \(k \times n\) matrix, and \(r \in \mathbb{R}^k\). To simplify notation we assume that \(R\) and \(r\) are fixed, but this is not essential. \(R\) and \(r\) may for instance depend on the fixed and random parameters \(\theta\). In what follows, we assume that \(\{q \in \mathbb{R}^n; Rq \leq r\}\) is compact and nonempty. Specific choice of \(R\) and \(r\) yield the examples referred to above:

**Example a** - Nonnegativity constraints: \(q \geq 0\), budget constraint: \(p'q \leq y\). So

\[ k = n + 1, \quad R = (p, -I)', \quad r = (y, 0, \ldots, 0)' . \]

**Example b** - Kinked budget constraint: \(c \leq w_j h + y_j (j = 1, \ldots, m)\), time constraint: \(0 \leq h \leq T\) (\(h\) = working hours, \(T\) = time endowment, \(c\) = total income, \(w_j, y_j\) = marginal wage and virtual nonlabour income bracket \(j\)). Thus

\[ k = m + 2, \quad R' = \begin{bmatrix} -w_1 & \cdots & -w_m & -1 & 1 \\ 1 & \cdots & 1 & 0 & 0 \end{bmatrix}, \quad r = (y_1, \ldots, y_m, 0, T)' . \]

If the utility function is strictly quasi-concave and continuously differentiable on the convex set \(\{q \in \mathbb{R}^n; Rq \leq r\}\), then the solution of the maximization problem is unique and can be found from Kuhn–Tucker conditions: \(q\) is optimal if and only if there is some \(\lambda \in \mathbb{R}^k\), such that

\[ \lambda \geq 0, \]

\[ Rq \leq r, \]

\[ \lambda (Rq - r) = 0, \]

\[ (\partial U(q)/\partial q)(q) = R' \lambda . \]

If, in addition to the conditions mentioned above, nonsatiation is imposed, then (13) can be rewritten in terms of the corresponding (homogeneous of degree zero) demand system \(F_\theta(p, y)\), which has the properties

\[ (\partial U(q)/\partial q)(F_\theta(p, y)) = \mu p \quad \text{for some} \quad \mu > 0, \]
and
\[ p'F_s(p, y) = y. \]

Using these properties and substituting \( \lambda = \frac{\overline{\lambda}}{\mu} \), (13) can be written as
\[
\begin{align*}
\lambda & \geq 0, \\
Rq & \leq r,
\end{align*}
\]
and
\[ q = F_s(R'\lambda, r'\lambda). \]

(Since demand is homogeneous of degree zero and \( \lambda \neq 0 \), some normalization on \( \lambda \) may be added.) \( R'\lambda \) and \( r'\lambda \) can be interpreted as a vector of shadow prices and corresponding shadow income, respectively.

To illustrate its general nature, we elaborate (14) for the two examples given above.

**Example a (continued).** (14) yields \( \lambda \geq 0, \ p'q \leq y, \ -q \leq 0, \) and \( q = F_s(\{\lambda_1 \rho - (\lambda_2, \ldots, \lambda_{n+1})\}, \lambda_1 y). \)

If utility is increasing in at least one of the quantities, then the budget constraint is binding, so \( \lambda_1 > 0 \). We can then choose the normalization \( \lambda_1 = 1 \) and this yields, with \( \overline{\lambda} = (\lambda_2, \ldots, \lambda_{n+1})' \):
\[
\begin{align*}
\overline{\lambda} & \geq 0, \ p'q = y, \ q \geq 0, \ q = F_s(p - \overline{\lambda}, y). \\
\end{align*}
\]

This is the well-known result that shadow prices cannot exceed actual prices.

**Example b (continued).** (14) yields \( \lambda \geq 0, \ c \leq w_j h + y_j \ (j = 1, \ldots, m), \ -h \leq 0, \ h \leq T, \) and
\[
(h, c)' = F_g \left( \left( -\sum_{j=1}^{m} w_j \lambda_j - \lambda_{m+1} + \lambda_{m+2} \right), \sum_{j=1}^{m} \lambda_j y_j + T \lambda_{m+2} \right) .
\]

Assuming that utility increases with \( c, \lambda_1 + \cdots + \lambda_m \) is positive, and \( \lambda \) can be normalized such that \( \lambda_1 + \cdots + \lambda_m = 1 \). This yields
\[
\begin{align*}
\lambda & \geq 0, \ \lambda_1 + \cdots + \lambda_m = 1, \ c \leq w_j h + y_j \ (j = 1, \ldots, m), \ 0 \leq h \leq T, \\
(h, c)' & = F_g \left( \left( -\sum_{j=1}^{m} w_j \lambda_j - \lambda_{m+1} + \lambda_{m+2} \right), \sum_{j=1}^{m} \lambda_j y_j + T \lambda_{m+2} \right) .
\end{align*}
\]
If all tax brackets consist of more than a single point, then at most two restrictions can be binding at the same time and there can only be $2m + 1$ regimes: $m$ regimes with one binding constraint and $m + 1$ regimes with two binding constraints ($m - 1$ kink points and two corners).

In case of one binding constraint, say the $j$th ($j \in \{1, \ldots, m\}$), we have

$$(h, c)' = F_{\theta}((-w_j, 1)', y_j), \quad \lambda_j = 1.$$  

In case of a kink point, between brackets $j$ and $j + 1$ ($j \in \{1, \ldots, m - 1\}$), we have

$$(h, c)' = F_{\theta}((- w_j \lambda_j - w_{j+1} [1 - \lambda_j], 1)', y_j \lambda_j + y_{j+1} [1 - \lambda_j])$$

$$= F_{\theta}((- \tilde{w}, 1)', \tilde{y}).$$

where $0 < \lambda_j < 1$. This is a familiar result: The shadow wage $\tilde{w}$ lies between $w_j$ and $w_{j+1}$ and shadow income $\tilde{y}$ satisfies $\tilde{y} + \tilde{w} h = y_j + w_j h = y_{j+1} + w_{j+1} h$, where $h$ is the number of hours at the kink point. The corners yield similar results.

### 4.1. Regularity conditions

The way in which regularity conditions for the model introduced above are formulated depends on whether we use (12), (13), or (14). If we use (12) only, then all we need is:

C.1. For each $\theta$, $U_{\theta}(q)$ must be defined on $\{q \in \mathbb{R}^n: Rq \leq r\}$ and (12) must yield a unique solution.$^2$

If we start from (13) and do not rely on duality results, then it is necessary to impose conditions that guarantee both coherency of (13) and equivalence of (13) with (12). The latter is necessary to retain the micro-economic foundation of (13). It is then sufficient to impose:

C.2. For each $\theta$, $u_{\theta}$ must be continuously differentiable and strictly quasi-concave on $\{q \in \mathbb{R}^n; Rq \leq r\}$.

According to standard Lagrange theory, C.2 is sufficient for the equivalence of (12) with (13) and for coherency of (13). However, it is easy to see that it is not

$^2$ As in section 3, we assume that $\theta = q/(\psi, \eta)$, where $\psi$ is fixed and $\eta$ is random with support $\Omega$. 'For each $\theta$' thus should be interpreted as 'for each $\psi \in \Psi$ and for each $\eta \in \Omega$'.
necessary. Take, for example, the case of Example a above, with a direct utility function which is increasing in all its components and quasi-concave on the edge of the budget set, but not on the whole interior.

In order to be able to start from (14), we need an extra condition to guarantee the equivalence of (14) with (13). Substitution of $\lambda/\mu$ by $\lambda$ is possible if nonsatisfaction is satisfied. If $F_\theta$ is obtained from the indirect utility function or the expenditure function, an extra condition is necessary to guarantee that $U_\theta$ is defined on $\{q \in \mathbb{R}^n; Rq \leq r\}$. In terms of section 3, this means that $\{q \in \mathbb{R}^n; Rq \leq r\}$ must be contained in the regular set in $q$-space, $Q_\theta$. Thus, it is sufficient (but, again, not necessary) to impose the condition B.2, defined in section 3, with $Q_{\min} = (q \in \mathbb{R}^n; Rq \leq r)$:

\[ C.3. \text{For each } \theta, Q_\theta \supset \{q \in \mathbb{R}^n; Rq \leq r\}. \]

Note that C.3 is stronger than C.2, since nonsatiation on $\{q \in \mathbb{R}^n; Rq \leq r\}$ is imposed. The advantage of C.3 compared to C.2 is that, in principle, C.3 can also be used if no explicit specification of the direct utility function is available.

In practice, conditions C.1–C.3 often appear to be stronger than necessary. In Example a for instance, if it is a priori known that $u_\theta$ increases in all its arguments, the 'budget set' can also be defined as $\{q; p'q = y, q \geq 0\}$. This implies a different choice of $R$ and $r$. Imposing C.2 or C.3 on the smaller set is less restrictive than on the original budget set, and still sufficient to guarantee coherency and regularity in the optimum.

Another example of a case in which C.2 or C.3 seem unnecessarily restrictive is the indirect translog model with nonnegativity constraints [cf. Lee and Pitt (1986)]. In this case shadow prices corresponding to the optimal quantities $q$ are either actual prices (if $q_i > 0$), or (if $q_i = 0$) can be obtained from a system of linear equations [cf., e.g., Lee and Pitt (1986)]. Van Soest and Kooreman (1990) derive sufficient conditions for coherency of (14). These imply regularity of preferences at the optimum, but not necessarily on $\{q \in \mathbb{R}^n; q \geq 0\}$, and are thus weaker than C.3.

4.2. External coherency

The analogue of B.4 and B.5 for the case of endogenously switching regimes requires that, in absence of measurement or optimization errors, the error structure of the model must be rich enough to explain observed optimal behaviour. Starting from (14), the condition is given:

\[ C.4. \text{('external coherency'). Let } RQ_\eta \text{ be a given set of restrictions (including the budget constraint) and quantities that satisfy these restrictions. For all } t, \psi, \text{ and } (R, r, q) \in RQ_\eta, \text{ there must be some } \eta \in \Omega \text{ such that } \lambda \in \mathbb{R}^k \text{ exists with } \lambda \geq 0 \text{ and } q = F_{\theta(t, \psi, \eta)}(R'\lambda, r'\lambda). \]
Similar conditions can be formulated starting from (12) or (13). Essentially, C.4 is the same as B.4 and B.5. It implies that the support $\Omega$ of the random preference terms $\eta$ must be large enough. Operationalization of this condition for a given demand system may be difficult. Imposition of C.4 can be avoided by explicit incorporation of measurement or optimization errors, but this is often undesirable from an economics viewpoint.

Example: QDU

We illustrate conditions C.1–C.3 with the QDU example, introduced in section 3, with focus on nonnegativity constraints and kinked budget sets. The QDU direct utility function is defined in each point of $\mathbb{R}^n$ by (7). C.1 holds if (12) has a unique solution. We assume that the set $\{q \in \mathbb{R}^n; Rq \leq r\}$ is compact, so existence of the solution is guaranteed, and only uniqueness may be a problem. Sufficient, but not necessary, for uniqueness is strict quasi-concavity $u_{B,n}$, i.e., the matrix $B$ of fixed parameters must be positive definite. For the case of non-negativity constraints, Ransom (1987a) proves coherency of (13) directly, writing it as a linear complementarity problem.

The fact that the choice set may contain the satiation point $B^{-1}y$, in which case duality results are no longer valid, is no problem if we use (13) only. It does become one if we start from (14). For C.3 to hold, we need both that $B$ is positive definite and that $B^{-1}y \in \{q \notin \mathbb{R}^n; Rq \leq r\}$.

Sufficient conditions for this can be derived if assumptions similar to those in section 3 are made. In particular, assume that in the relevant region of $q$-space, utility increases with $q_n$, i.e., $\gamma_n - (Bq)_n > 0$. This leads to imposition of B.2 with $Q_{\min} = \{q \in \mathbb{R}^n; Rq \leq r\}$:

$$\gamma_n > \max_{q} \{(Bq)_n; Rq \leq r\}. \quad (15)$$

The maximum of the right-hand side of (15) can be found by linear programming.

In the special case of nonnegativity constraints, assuming that all prices are strictly positive, (15) yields

$$\gamma_n > y \max_{1 \leq j \leq n} (\beta_{nj}/p_j).$$

In case of a kinked budget constraint (15) yields

$$\gamma_n > \max_{0 \leq j \leq m} (\beta_{21} h_j + \beta_{22} c_j),$$
where \((h_j, c_j), j = 0, \ldots, n\), are the corners \((h_0, c_0) = (0, y_1)\) and \((h_m, c_m) = (T, w_m T + y_m)\) and the kink points \((h_j, c_j) = \left(\{y_{j+1} - y_j\}/\{w_j - w_{j+1}\}, w_j h_j + y_j\), \(j = 1, \ldots, m - 1\).

5. Conclusions

We have studied coherency and regularity properties of various static neo-classical models of consumer demand and labour supply. Emphasis has been placed on the relation between regularity properties of underlying preferences, in particular concavity and monotonicity, and coherency of the econometric model based on these preferences. In section 2, we have demonstrated the necessity of imposing coherency conditions in practice. An example shows that even though the true model is coherent, failure to impose appropriate conditions a priori may lead to inconsistent estimators of the parameters, at least if maximum likelihood is used. The estimates would then lead to the conclusion that the model is misspecified. This also shows that it is not possible to test whether the model is coherent or not.

If specifications of demand systems would be available which were tractable, flexible, and globally concave at the same time, then the treatment of coherency conditions would be straightforward. Given regularity of preferences, existence and uniqueness of the solution of the utility maximization problem follows from standard Lagrange theory, at least if the budget set (all constraints taken into account) is compact and convex. In general however, tractable, flexible systems only have local concavity properties. Coherency can then be guaranteed by imposing regularity conditions in some relevant region of price or quantity space. In section 3, a general framework is sketched which shows how regularity conditions can be imposed which are relevant in case of a standard, an inverse, or a conditional demand system. In all these cases it is a priori known which quantity restrictions are binding. Regularity conditions generally limit the range of possible realisations of the error terms (random preferences) in the model. Another condition, 'external coherency', is introduced, which states that, on the other hand, the presence of random terms must provide so much flexibility that the data can be described, i.e., likelihood contributions must be positive. The latter condition can be in conflict with the regularity conditions.

The example in section 3 shows how the restrictions can be imposed in practice. In this case, they can be dealt with in a rather tractable way, since explicit expressions are available for both standard and inverse demand functions. For specifications such as AIDS, the inverse demand equations cannot

\[\text{For example, the generalized McFadden cost function proposed by Diewert and Wales (1987) is second-order locally flexible and concave on the positive orthant of price space, but is not tractable, in the sense that it does not permit explicit expressions for Marshallian demand functions.}\]
be derived analytically. As a consequence, the conditions become very intricate
and imposing regularity is a cumbersome affair which will only be possible using
numerical tools; see Van Soest et al. (1990).

In section 4, we consider coherency and regularity of models with endogen-
ously switching regimes. In this type of models, coherency is more often a prob-
lem than in the models in section 3. Coherency conditions are derived for the
linear case by Gourieroux et al. (1980), but the models based on neoclassical
theory which we consider will hardly ever be linear. Instead of imposing
coherency directly, we suggest to impose regularity conditions in the relevant
area of quantity space. This will in general be sufficient for coherency but may be
very restrictive and conflicting with the external coherency requirement. Often it
is clear that less restrictive conditions will suffice, because in practice most of the
budget set is irrelevant for the individual anyway. Thus the utility function only
needs to be quasi-concave in that part of the budget set of which it is a priori
clear that it will contain the optimum, e.g., the edge of the budget set only. By
making the area where regularity conditions are imposed as small as possible,
maximum flexibility of the functional form is retained. At the same time this may
complicate the analysis since it is often cumbersome to specify in which area
regularity should be imposed.

The approach of imposing regularity conditions *a priori* to guarantee coherency
suggested in this paper, suffers from a number of drawbacks and complications.
First, the conditions given are generally sufficient but only in very specific cases
they are, in some sense, necessary. This is important since imposing unnecessary
conditions affects model flexibility. Moreover, budget set and parameters may
vary across individuals. Some conditions, like 'external coherency', suggest that
the parameter space should not be too small, whereas others imply that it cannot
be too large. These conditions may easily be conflicting. An issue related to the
previous points is that the stochastic specification tends to be difficult. In the
examples considered, the support of the random variables was often constrained
to a polyhedron. If, for instance, we would specify a normal distribution for the
random preferences, this would lead to complicated truncations.

Another implication of the analysis is that in most models with endogenous
regimes or corner solutions, analytical expressions for the parameter restrictions
implied by the regularity conditions can only be obtained if a closed form
expression of the direct utility function is available. This is rather clear from the
analysis in section 4, but also under exogenous rationing, conditions like B.2 or
B.3 require knowledge of shadow prices in a rationing point. Although in
principle shadow prices can be computed numerically whenever given in implicit
form, it is hard to see how conditions like B.2 or B.3 should be imposed when no
closed form expressions for shadow prices are available. And, of course, knowing
shadow prices corresponding to given quantities amounts to knowing the
direct utility function. As a result, many of the popular flexible forms like AIDS
or Indirect Translog cannot be used in general.
Altogether, the treatment of endogenous regimes or corner solutions often requires intricate procedures for the imposition of regularity conditions and severely limits the number of functional forms that can be considered. Despite these difficulties, it should be clear that without the imposition of regularity conditions one will often end up with a nonsensical model. Thus the choice appears to be between complexity and incoherency.

Appendix: Inconsistency of the ‘ML’ estimator

In the bivariate Probit model in section 2, assume that the true parameters are

\[ \beta_1 = 1, \quad \gamma_1 = -1, \quad \beta_2 = 0, \quad \gamma_2 = 0, \]

where \( \gamma_2 = 0 \) implies coherency. Assume that \( x \) is a dummy variable with \( P[x = 0] = P[x = 1] = \frac{1}{2} \). Inserting the true parameter values in (2) yields (subscript \( t \) is omitted)

\[
\begin{align*}
    \Pr[y_1 = 0, y_2 = 0 | x = 0] &= 0.250; \quad \Pr[y_1 = 0, y_2 = 0 | x = 1] = 0.079, \\
    \Pr[y_1 = 0, y_2 = 1 | x = 0] &= 0.421; \quad \Pr[y_1 = 0, y_2 = 1 | x = 1] = 0.250, \\
    \Pr[y_1 = 1, y_2 = 0 | x = 0] &= 0.250; \quad \Pr[y_1 = 1, y_2 = 0 | x = 1] = 0.421, \\
    \Pr[y_1 = 1, y_2 = 1 | x = 0] &= 0.079; \quad \Pr[y_1 = 1, y_2 = 1 | x = 1] = 0.250.
\end{align*}
\]

Let \( K(i, j, k) \) be the number of observations with \( y_1 = i, y_2 = j, \) and \( x = k \) \((i, j, k \in \{0, 1\})\). Note that, if the total number of observations is \( 2N \),

\[
\lim_{N \to \infty} \{ K(i, j, k) / N \} = \Pr[y_1 = i, y_2 = j | x = k], \quad i, j, k \in \{0, 1\}.
\]

The extended log-likelihood (3) can be rewritten as

\[
L(\beta_1, \beta_2, \gamma_1, \gamma_2) = NL_1(\beta_1, \gamma_2) + NL_2(\beta_2, \gamma_2),
\]

where

\[
L_1(\beta_1, \gamma_1) = \frac{1}{N} \left\{ \sum_{j, k = 0}^1 K(0, j, k) \log \Phi(-\beta_1 k - \gamma_1 j) \right. \\
+ \left. \sum_{j, k = 0}^1 K(1, j, k) \log \Phi(\beta_1 k + \gamma_1 j) \right\}
\]

and

\[
L_2(\beta_2, \gamma_2) = \frac{1}{N} \left\{ \sum_{i, k = 0}^1 K(i, 0, k) \log \Phi(-\beta_2 k - \gamma_2 i) \\
+ \sum_{i, k = 0}^1 K(i, 1, k) \log \Phi(\beta_2 k + \gamma_2 i) \right\}.
\]
$L$ can be maximized by maximizing $L_1(\beta_1, \gamma_1)$ and $L_2(\beta_2, \gamma_2)$ separately. This means that the two simultaneous Probit equations are treated as if they were separate Probit equations, i.e., maximizing $L$ over $\Theta^*$ yields a consistent estimator if $\gamma_2$ is exogenous in the first equation and $\gamma_1$ is exogenous in the second equation. Since the true $\gamma_2$ is 0, $\gamma_2$ is independent of $\varepsilon_1$. Thus the estimator for $(\beta_1, \gamma_1)$ is consistent. The estimators for $\beta_2$ and $\gamma_2$ however are inconsistent, as can be shown by a straightforward computation of their probability limits: In the limiting case ($N \to \infty$), we have

$$
\text{plim } L_2(\beta_2, \gamma_2) = \frac{1}{2} \left[ 0.250 \log \Phi(0) + 0.079 \log \Phi(-\beta_2) 
+ 0.250 \log \Phi(-\gamma_2) + 0.421 \log \Phi(-\beta_2 - \gamma_2) 
+ 0.421 \log \Phi(0) + 0.250 \log \Phi(\beta_2) 
+ 0.079 \log \Phi(\gamma_2) + 0.250 \log \Phi(\beta_2 + \gamma_2) \right].
$$

Since $L_2$ has the form of the log-likelihood of a Probit model, it is globally concave. The unique maximum can easily be found numerically. For $N \to \infty$, we thus obtain

$$
\text{plim } \hat{\beta}_2 = 0.5726 \neq 0 = \beta_2 \quad \text{and} \quad \text{plim } \hat{\gamma}_2 = -0.8405 \neq 0 = \gamma_2.
$$

Finally, note that if the restriction $\gamma_1 \gamma_2 = 0$ is imposed (which is necessary and sufficient for coherency), then the resulting estimator for $\beta_2$ is consistent.4

References


4 It is easily verified that the function $\text{plim } \{L(\beta_1, \beta_2, \gamma_1, \gamma_2)/N\}$ attains two local maxima on the set $\{(\beta_1, \beta_2, \gamma_1, \gamma_2) : \gamma_1, \gamma_2 = 0\}$: $\text{plim } \{L(0.443, 0.573, 0, -0.841)/N\} = -2.26$ and $\text{plim } \{L(1.0, -1.0)/N\} = -2.17$. The global maximum is thus attained for the true parameter values $\beta_1 = 1, \beta_2 = 0, \gamma_1 = -1, \gamma_2 = 0.$


Van Soest, A., P. Kooreman, and A. Kapteyn, 1990, Coherency and regularity of demand systems with equality and inequality constraints, CentER discussion paper 9001 (Tilburg University, Tilburg).


Reprint Series, CentER, Tilburg University, The Netherlands:


No. 5  Th. ten Raa and F. van der Ploeg, A statistical approach to the problem of negatives in input-output analysis, *Economic Modelling*, vol. 6, no. 1, 1989, pp. 2 - 19.


No. 8  Th. van de Klundert and F. van der Ploeg, Wage rigidity and capital mobility in an optimizing model of a small open economy, *De Economist*, vol. 137, nr. 1, 1989, pp. 47 - 75.


<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>Author(s)</td>
<td>Title</td>
<td>Journal/Media</td>
</tr>
<tr>
<td>------</td>
<td>---------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>130</td>
<td>R. Gradus and A. de Zeeuw</td>
<td>An employment game between government and firms</td>
<td>Optimal Control Applications &amp; Methods, vol. 13, no. 1, 1992</td>
</tr>
<tr>
<td>132</td>
<td>F. Groot, C. Withagen and A. de Zeeuw</td>
<td>Note on the open-loop Von Stackelberg equilibrium in the Cartel versus Fringe model</td>
<td>The Economic Journal, vol. 102, no. 415, 1992</td>
</tr>
<tr>
<td>136</td>
<td>H. G. Bloemen and A. Kapteyn</td>
<td>The joint estimation of a non-linear labour supply function and a wage equation using simulated response probabilities</td>
<td>Annales d’Économie et de Statistique, No. 29, 1993</td>
</tr>
<tr>
<td>140</td>
<td>S. Eijffinger and E. Schaling</td>
<td>Central bank independence in twelve industrial countries</td>
<td>Banca Nazionale del Lavoro Quarterly Review, No. 184, March 1993</td>
</tr>
</tbody>
</table>


