Arbeidsnotater

STATISTISK SENTRALBYRÅ

Dronningensgt. 16, Dep, Oslo 1. Tlf.*(02) 41 38 20

WORKING PAPERS FROM THE CENTRAL BUREAU OF STATISTICS OF NORWAY

IO 78/4 24. February 1978

ESTIMATING CONSUMER DEMAND FUNCTIONS FROM HOUSEHOLD BUDGET SURVEYS FROM SEVERAL YEARS: ECONOMETRIC PROBLEMS AND METHODS RELATED TO INCOMPLETE CROSS-SECTION/TIME-SERIES DATA

By

Erik Biørn*)

*) I am grateful to Eilev S. Jansen for useful comments. The responsibility for remaining errors and shortcomings of course rests with me. Comments are welcome.

CONTENTS

Page

1. Introduction 1

2. The system of demand functions 3

2.1. General concepts 3

2.2. The structure of disturbances: preliminary remarks 4

2.3. Possible parametrizations of the budget share functions . 5

3. The single-equation (one commodity) model approach 10

3.1. The structure of disturbances: Basic assumptions 10

3.2. The categories of data 11

3.3. The disturbance variance/covariance matrix: CS/TS data with disjointed samples 15

3.4. The disturbance variance/covariance matrix: Complete CS/TS data 17

3.5. The disturbance variance/covarianse matrix: Rotation samples 20

4. The simultaneous equations (multi-commodity) model approach ... 23

4.1. The structure of disturbances: Basic assumptions 23

4.2. The sampling schemes 26

4.3. The disturbance variance/covariance matrix: Disjointed CS/TS samples 27

4.4. The disturbance variance/covariance matrix: Complete CS/TS data 28

4.5. The disturbance variance/covariance matrix: Rotation samples 31

4.6. Special cases: Comparison with the single equation model 32

5. Estimation 37

5.1. Preliminaries: The Full Information Maximum Likelihood (FIML) principle 37

5.2. Estimation in the single-equation model 38

5.3. Estimation in the multi-equation model 49

Appendix A. The disturbance variance/covariance matrix corresponding with rotation samples, when ordering the observations first by individual second by period. Addendum to section 3.5. 57

Appendix B. The disturbance variance/covariance matrix corresponding with complete CS/TS data when the ordering by commodity precedes the ordering by period and individual. Addendum to section 4.4. 61

Appendix C. Some useful properties of Kronecker products 65

References 69

1. Introduction

The data commonly applied when estimating complete systems of consumer demand functions are time series of aggregate household expenditures (per capita) at current and constant prices. Since the consumer demand function, as derived from classical utility theory, is in essence a micro concept, this approach inevitably raises well-known problems of aggregation. Micro data from household budget surveys have proved very useful in analysing the effect on the **consumption** of different commodities resulting from changes in income, family size, age, etc. However, as such data are usually collected during a fairly short time span, e.g. one year, prices and other variables with essentially time specific variation, show virtually no variation, and we cannot estimate their effect with an acceptable degree of precision.

In this situation, the idea of combining data from several household budget surveys naturally comes to mind. Provided that the different surveys do not differ substantially with respect to the sampling plan, the definition of variables, the length of the period in which each household is under observation etc., this approach may be fruitful. Unfortunately, such conformity is not always satisfied in practice.[1)] Recently, however, attempts at a more or less continuous registration of household expenditures according to a unified plan have been made in several countries. The Central Bureau of Statistics of Norway, for instance, is continuously collecting expenditure reports on a sample survey basis from the year 1974 on. This sheds a new light on many problems related to empirical consumer demand analysis, and should accordingly induce econometricians working in this field to look their models and methods over again.

From the outset, we should make the following clear: It is practically very difficult - and probably impossible - to obtain household expenditure data which conform with those dealt with in the standard models for analysing 'complete cross-section/time series data'.[2)] We cannot expect all - not even the majority of - the households selected for observation in one year to accept participating in the next five or ten years as well. As the reporting of consumption expenditures is rather time-consuming - at

---

1) The design of the Norwegian Surveys of Consumer Expenditure of 1958, 1967, and 1973, for instance, differed significantly in several respects.

2) For instance, the error components models of Balestra and Nerlove [3] and others.

least if it is based on detailed book-keeping, as in Norway - the data-collecting agency will hardly be able to persuade households to participate more that once, or at most twice. It is worth mentioning that the proportion of non-respondents in the Norwegian household budget surveys is as high as 30 per cent, even when the households are asked to report only once (and even though the book-keeping period of these households is only two weeks).

Thus, the data we can obtain in practice, constitute what may be properly called 'incomplete cross-section/time-series data': different households as well as different periods (years) are represented, but the sample of households changes over time. The purpose of this paper is to discuss problems relating to model formulation and estimation in such situations. In order to put our results in perspective, reference will also be made to the (hypothetic) case where complete cross-section/time-series data exist. Our plan is to try to implement, by means of Norwegian data, some of the theoretical results obtained in the paper at a later stage.

Chapter 2 discusses the model formulation in rather general terms, with one section devoted to the classification of variables, one section devoted to the decomposition of the disturbances and one section dealing with the parametrization of the demand functions. Chapter 3 concentrates on the special case where only one commodity group is under consideration. The disturbance variance-covariance matrix of the model are discussed separately for three different sampling schemes: disjointed samples, complete cross-section/time-series data, and samples where the households "rotate". In chapter 4, these results are generalized to a complete demand model. Finally, chapter 5 is devoted to estimation, taking the Full Information Maximum Likelihood principle as the point of departure.

2. The system of demand functions

2.1. General concepts

Assume that the consumption commodities are divided into N groups, and let the vector x symbolize the quantities consumed (expenditures at constant prices) by a 'typical' household to be explained by our model. The set of exogenous (explanatory) variables can be separated into three subvectors:

- q: a vector of individual (household specific) variables; i.e. variables showing variation across individuals (households), but (for practical purposes) no variation over time,
- s: a vector of time specific variables; i.e. variables showing variation over time, but (for practical purposes) no variation across individuals,
- z: a vector of combined variables; i.e. variables showing simultaneous variation across individuals and over time.

The q vector may, for instance, include the year of birth and the sex of the head of household, his (or her) socioeconomic group, the geographic location of the household etc., provided that the same household member is considered the head in all the periods of observation and that his socioeconomic group and the location of the household do not change over time. If the latter assumptions seem too restrictive, the relevant variables should be included into the z vector. Examples of variables belonging to the s vector may be: consumer prices (disregarding geographic price differences), tax rates, indicators of the general economic situation etc. Of course, different individuals may evaluate the economic outlook differently and have different expectations about the future. Variables of this kind, provided they can be quantified, should be included into the z vector. Finally, the z vector may represent the income and wealth position of the household, (and, possibly, lagged values of these variables), the stock of durables, the number of household members, the week(s) in each distinct period (e.g. year) during which the consumption expenditures are reported [1] etc.

Let subscript i denote the commodity number (no. of the element of the x vector), and let h symbolize the number of the individual (household) and t the number of the period. The system of demand equations can then be written in the following general form:

1) The variable 'week(s) of reporting' may be "individualized", i.e. transferred from z to q, by adopting an observation (sample) plan in which each household reports in the same week(s) in all the periods (years) it participates in the investigation.

(2.1) $x_{iht} = f_i'(q_h, s_t, z_{ht}) + \varepsilon_{iht}'$ (i = 1, ..., N),

where ε_{iht}' is a stochastic disturbance supposed to be uncorrelated with q_h, s_t and z_{ht}. We assume that the vectors of exogenous variables are properly specified, i.e. that the structural parts of the f_i' functions do not shift across households or over time.

Let p_{jt} denote the price (index) of the j'th commodity in the t'th period and let

(2.2) $$y_{ht} = \sum_{i=1}^{N} p_{it}x_{iht}$$

represent the total consumption expenditure of the h'th household in the t'th period. The price vector $(p_{1t}, ..., p_{Nt})$ is a part of the "time specific" vector s_t, and y_{ht} is an element in the "combined" vector z_{ht}. Defining

(2.3) $$a_{iht} = \frac{p_{it}x_{iht}}{y_{ht}}$$ (i = 1, ..., N),

i.e. the budget share of the i'th commodity for the h'th individual in the t'th period, we respecify eqs. (2.1) as follows:

(2.4) $a_{iht} = f_i(q_h, s_t, z_{ht}) + \varepsilon_{iht}$ (i = 1, ..., N),

where $f_i(\cdot) = p_{it}f_i'(\cdot)/y_{ht}$, and $\varepsilon_{iht} = p_{it}\varepsilon_{iht}'/y_{ht}$.

It seems more plausible to assume that the variances of ε_{iht} are constant across individuals and over time than it is to make this assumption for ε_{iht}'. This rests on the fact that homoscedasticity of ε_{iht} (across h and t) (i) pays regard to the notion that the scope for variations in consumption habits is larger the higher is the real income, and (ii) ensures that a proportional change of prices and income does not affect the second order moments of the distribution of the disturbances.[2] Therefore, we shall stick to the formulation given in (2.4).

2.2. The structure of disturbances: preliminary remarks

The disturbances ε_{iht} represent the net effect of several non-observable variables not included in the argument list of f_i. One possible specification of the stochastic structure might be, for each value of i, to

2) See Biørn [4], pp. 4-7, for a further elaboration of this point.

let all ε_{iht} be independenly and identically distributed. However, recalling that the effects taken care of by the disturbances may be partly purely individual (e.g., "tastes", "habits" etc.), partly purely time specific (e.g., general expectations concerning the economic development) and partly combined effects[3], this does not appear as the preferable solution.

Rather, we shall take the 'error components approach' previously used by several authors in the context of single equation regression models (cf. e.g., Balestra [2], Balestra and Nerlove [3], Chetty [6], Kuh [9], Maddala [11], Nerlove [13], [14], and Wallace and Hussain [18]), specifying ε_{iht} as the sum of three components:

$$(2.5) \qquad \varepsilon_{iht} = u_{ih} + v_{it} + w_{iht} \qquad (i = 1, \ldots, N).$$

Here u_{ih} is the individual component, v_{it} is the time specific component and w_{iht} is the combined component respectively associated with commodity i.

The structure of the disturbances will be discussed in some detail in chapters 3 and 4, so we leave this problem here. Only one remark: If the parametric speeification of the demand functions conforms "exactly" with constrained utility maximization, the adding-up condition

$$(2.6) \qquad \sum_{i=1}^{N} f_i(q_h, s_t, z_{ht}) = 1$$

will be satisfied identically in q_h, s_t, and z_{ht}. Then, in view of (2.2)-(2.4), the distribution of the disturbances should obey

$$(2.7) \qquad \sum_{i=1}^{N} \varepsilon_{iht} = \sum_{i=1}^{N} (u_{ih} + v_{it} + w_{iht}) = 0 \quad \text{for all h and t.}$$

On the other hand, if the functional forms chosen imply satisfaction of (2.6) only approximately, then (2.7) does not represent an exact and absolute constraint.

2.3. Possible parametrizations of the budget share functions

The choice of parametric specification of the budget share functions f_i clearly deserves particular attention. Briefly stated, our problem is to establish functional forms that a) are sufficiently flexible to reflect adequately the variations in consumption pattern across the individual as

---

3) Compare the formally similar disaggregation of the vector of (the observable) exogenous variables.

well as the time "dimension", b) agree reasonably well with commonly accepted theory of consumer's demand, and c) permit econometric estimation with the Aitken Generalized Least Squares or the Maximum Likelihood methodology.

We shall not attempt to give a definite solution to this problem; that is partly an empirical matter, of course. At this stage, we confine ourselves to a list of selected functions which may be worth investigating. We concentrate on the parametrization of the income and price responses. Demographic and socioeconomic variables, as well as the possible effect of expectational variables will not be introduced at this stage.

All the equation systems A-H below satisfy the homogeneity constraints of the static theory of choice; A-D satisfy the adding-up condition identically. Among these, only A-C satisfy the conditions of symmetry and negative definiteness of the Slutzky substitution matrix.[4)]

A. Linear Expenditure Functions à la Stone

This specification implies

$$(2.8) \qquad a_{iht} = c_i \frac{p_{it}}{y_{ht}} + b_i \left(1 - \sum_{j=1}^{N} \frac{p_{jt}}{y_{ht}} c_i\right)$$

$$(0<b_i<1, \ \Sigma_i b_i = 1).$$

B. A generalization of the Stone system along the lines of Fourgeaud and Nataf

Replace the constants c_i in (2.8) by

$$(2.9) \qquad c_i = t_i \left(\frac{p_{it}}{P_t}\right)^{\beta-1} C\left(\frac{y_{ht}}{P_t}\right),$$

where $P_t = \sum_{j=1}^{N} (t_j p_{jt}^{\beta})^{1/\beta}$ and $C(\cdot)$ is an unspecified function obeying certain constraints. (Cf. Johansen [8], Nasse [12].)

4) In the rest of this chapter, all disturbance terms are, for simplicity, omitted.

C. Carlevaro's generalization of the Stone system

Replace the constants b_i in (2.8) by

$$b_i = b_{i0} + b_{i1}\phi\left(\frac{y_{ht}-\Sigma_j p_{jt}c_j}{P_t}\right) \quad (\Sigma b_{i0} = 1,\ \Sigma b_{i1} = 0), \tag{2.10}$$

where $\phi(\cdot)$ and the price index function $P_t = P(p_{1t}, \ldots, p_{Nt})$ satisfy certain conditions. (Cf. Carlevaro [5].)
Notice, in passing, that (2.3), (2.8) and (2.10) imply

$$\frac{p_{it}(x_{iht}-c_i)}{y_{ht}-\Sigma_j p_{jt}c_j} = b_{i0} + b_{i1}\phi\left(\frac{y_{ht}-\Sigma_j p_{jt}c_j}{P_t}\right),$$

i.e., the 'supernumerary budget shares' are linear functions of a transformation $\phi(\cdot)$ of the 'supernumerary real income'.

D. Budget shares polynomials of the third degree in total real expenditure

We postulate[5]

$$a_{iht} = \alpha_i^D + \beta_i^D \frac{y_{ht}}{P_t} + \gamma_i^D\left(\frac{y_{ht}}{P_t}\right)^2 + \delta_i^D\left(\frac{y_{ht}}{P_t}\right)^3, \tag{2.11}$$

where P_t is a consumer price index homogeneous of the first degree, and $\Sigma_i\alpha_i^D = 1$, $\Sigma_i\beta_i^D = \Sigma_i\gamma_i^D = \Sigma_i\delta_i^D = 0$.

E. Expenditures at constant prices polynomials of the third degree in total real expenditure

This specification implies that x_{iht}, rather than a_{iht}, is a polynomial of the third degree in y_{ht}/P_t, i.e.,

$$a_{iht} = \alpha_i^E\frac{p_{it}}{y_{ht}} + \beta_i^E\frac{p_{it}}{P_t} + \gamma_i^E\frac{p_{it}y_{ht}}{P_t^2} + \delta_i^E\frac{p_{it}y_{ht}^2}{P_t^3}. \tag{2.12}$$

This parametrization does not, in contrast with A-D, satisfy the adding-up constraint (2.6) exactly. It will, however, hold reasonably well if in the period of interest prices change in such a way that

$$\Sigma_i\alpha_i^E p_{it} = \Sigma_i\gamma_i^E p_{it} = \Sigma_i\delta_i^E p_{it} = 0,$$

$$\Sigma_i\beta_i^E p_{it} = P_t,$$

are satisfied approximately.

---

5) Of course, this specification may be generalized to polynomials of any degree.

F. Budget shares linear in total real expenditure and own real price

In this case, the budget share functions take the form

$$(2.13)\quad a_{iht} = \beta^F_{i0} + \beta^F_{i1}\frac{y_{ht}}{P_t} + \beta^F_{i2}\frac{p_{it}}{P_t}.$$

Neither in this case will the adding-up constraint be satisfied exactly, but it will hold approximately if

$$\Sigma_i\beta^F_{i0} \approx 1,\ \Sigma_i\beta^F_{i1} \approx 0,\ \Sigma_i\beta^F_{i2}p_{it} \approx 0,$$

or if $\Sigma_i\beta^F_{i0} \approx 0,\ \Sigma_i\beta^F_{i1} \approx 0,\ \Sigma_i\beta^F_{i2}p_{it} \approx P_t.$

G. Expenditures at constant prices linear in total real expenditure and own real price

Specification G differs from F in a similar way as E differs from D, i.e. we postulate

$$(2.14)\quad a_{iht} = \beta^G_{i0}\frac{p_{it}}{y_{ht}} + \beta^G_{i1}\frac{p_{it}}{P_t} + \beta^G_{i2}\frac{p_{it}^2}{y_{ht}P_t}.$$

Approximate satisfaction of the adding-up constraint in this case is ensured if

$$\Sigma_i\beta^G_{i0}p_{it} \approx 0,\ \Sigma_i\beta^G_{i1}p_t \approx P_t,\ \Sigma_i\beta^G_{i2}p_{it}^2 \approx 0.$$

H. Modification of specification G to permit cross-price effects

A drawback with specifications D-G is that cross-price responses are rather summarily represented. For some levels of aggregation this may be felt a serious lack of realism. A possible remedy might of course be to extend (2.13), or (2.14), to a full quadratic form in y_{ht}/P_t and $p_{jt}/P_t(j=1, \ldots, N)$, or in p_{it}/y_{ht} and $p_{jt}/P_t(j=1, \ldots, N)$, respectively. Such extensions would, however, imply a lavish increase in the number of coefficients, even for moderate values of N.

The following specification, proposed by Lybeck [10] in connection with aggregate time series data:

$$(2.15)\quad a_{iht} = \beta^H_{i0}\frac{p_{it}}{y_{ht}} + \beta^H_{i1}\frac{p_{it}}{P_t} + \sum_{j=1}^{N}\gamma^H_{ij}\frac{p_{it}}{P_t}\cdot\frac{p_{jt}}{P_t},$$

represent an intermediate solution. (Lybeck, however, presents and uses his equation with $P_t x_{iht}/y_{ht} = a_{iht} P_t/p_{it}$, i.e. the budget shares at constant prices, as left-hand variables.) In this case, approximate satisfaction of the adding-up constraint would be ensured if

$$\Sigma_i \beta^H_{i0} p_{it} \approx 0, \ \Sigma_i \beta^H_{i1} p_{it} \approx P_t, \ \Sigma_i \Sigma_j p_{it} \gamma^H_{ij} p_{jt} \approx 0.$$

3. The single-equation (one commodity) model approach

3.1 The structure of disturbances: Basic assumptions

In this chapter , our attention is devoted to one commodity only. Let ε_{ht} denote the disturbance of the demand function for this commodity relating to the h'th household and the t'th period. Assuming that ε_{ht} may be decomposed into three additive components, a household specific (individual) component u_h, a period (time) specific component v_t, and a combined component (a remainder) w_{ht}, we have

$$(3.1) \quad \varepsilon_{ht} = u_h + v_t + w_{ht} \qquad \text{for all h and t.}$$

(Notice that the symbols correspond to those used in ch. 2 with the commodity subscript i omitted.) All components are supposed to have zero expectations,

$$(3.2) \quad E(u_h) = E(v_t) = E(w_{ht}) \qquad \text{for all h and t,}$$

and to be mutually uncorrelated, with constant variances, i.e.,

$$\left.\begin{array}{ll} (3.3a) & E(u_h u_k) = \delta_{hk}\sigma_I^2 , \\ (3.3b) & E(v_t v_s) = \delta_{ts}\sigma_T^2 , \\ (3.3c) & E(w_{ht} w_{ks}) = \delta_{hk}\delta_{ts}\sigma_C^2 , \\ (3.3d) & E(u_h v_t) = E(u_h w_{ks}) = E(v_t w_{ks}) = 0, \end{array}\right\} \text{for all h,k,s and t,}$$

where δ_{hk} and δ_{ts} denote 'Kronecker deltas'

$(\delta_{hh} = \delta_{tt} = 1,\ \delta_{hk} = 0 \text{ for } k \neq h,\ \delta_{ts} = 0 \text{ for } s \neq t)$.

The subscripts I,T, and C symbolize "individual", "time specific" and "combined", respectively.

From assumptions (3.1) - (3.3) follows

$$(3.4) \quad E(\varepsilon_{ht}) = 0,$$

$$(3.5)\quad E(\varepsilon_{ht}\varepsilon_{ks}) = \delta_{hk}\sigma_I^2 + \delta_{ts}\sigma_T^2 + \delta_{hk}\delta_{ts}\sigma_C^2 .$$

The variances/covariances may be written, more explicitly, as

$$(3.5a)\quad E(\varepsilon_{ht}\varepsilon_{ks}) = \begin{cases} \sigma^2 & \text{for } k = h \ \& \ s = t, \\ \rho\sigma^2 & \text{for } k = h \ \& \ s \neq t, \\ \omega\sigma^2 & \text{for } k \neq h \ \& \ s = t, \\ 0 & \text{for } k \neq h \ \& \ s \neq t, \end{cases}$$

where

$$(3.6)\quad \begin{cases} \sigma^2 = \sigma_I^2 + \sigma_T^2 + \sigma_C^2, \\ \rho = \sigma_I^2/\sigma^2, \\ \omega = \sigma_T^2/\sigma^2. \end{cases}$$

3.2 The categories of data

Of course, we do not observe all households in the population in all periods; i.e. observations for all possible combinations of h and t do not exist in the sample at our disposal. Only selected (h,t) constellations are represented. It is useful to distinguish between the following three main categories of data:

A. Pure cross section (CS) data.

B. Pure time series (TS) data.

C. Combined cross section/time series (CS/TS) data.

In pure cross section data, all observations are taken from one period, i.e., observations for which $s \neq t$ do not exist. Assuming that the period in question has number 1, the first and second order moments of the disturbances are completely described by (cf.(3.4)-(3.5))

$$E(\varepsilon_{h1}) = 0,$$

$$E(\varepsilon_{h1}\varepsilon_{k1}) = \begin{cases} \sigma^2 & \text{for } k=h, \\ \omega\sigma^2 & \text{for } k\neq h, \end{cases}$$

We notice that all disturbances are correlated, since they have v_1 as a common stochastic component. At a first glance this seems to be in conflict with the specification commonly adopted when analysing cross section data. The solution is, of course, that the usual assumptions should be interpreted as conditional with respect to the value of the time specific component in period no. 1. From (3.1)-(3.3) follows

$$E(\varepsilon_{h1}|v_1) = v_1,$$

$$E(\varepsilon_{h1}\varepsilon_{k1}|v_1) = \begin{cases} \sigma_I^2 + \sigma_C^2 = (1-\omega)\sigma^2 & \text{for } k=h, \\ 0 & \text{for } k \neq h, \end{cases}$$

I.e., ε_{h1} and ε_{k1} ($k \neq h$) are uncorrelated in the distribution conditional on v_1. The non-zero expectation v_1 causes no problem, as this is an unidentifiable constant that cannot be distinguished from the constant term of the "structural" part of the equation; $E(\varepsilon_{h1}|v_1)$ may be set equal to zero without loss of generality.

Correspondingly, in pure time series data, all observations relate to one individual, i.e., observations for which $k \neq h$ do not exist. If the individual in question has number 1, the structure of the disturbances in the sample may be described as (cf. (3.4)-(3.5))

$$E(\varepsilon_{1t}) = 0,$$

$$E(\varepsilon_{1t}\varepsilon_{1s}) = \begin{cases} \sigma^2 & \text{for } s=t, \\ \rho\sigma^2 & \text{for } s \neq t. \end{cases}$$

Also in this situation all disturbances are correlated, since they have the individual stochastic component u_1 in common. The corresponding first and second order moments conditional on the value of this component are

$$E(\varepsilon_{1t}|u_1) = u_1,$$

$$E(\varepsilon_{1t}\varepsilon_{1s}|u_1) = \begin{cases} \sigma_T^2 + \sigma_C^2 = (1-\rho)\sigma^2 & \text{for } s=t, \\ 0 & \text{for } s \neq t. \end{cases}$$

I.e., ε_{1t} and ε_{1s} ($s \neq t$) are uncorrelated in the distribution conditional on u_1.

A sample of combined cross section/time series data contains observations from different individuals as well as observations from different time periods. Assume that the observations relate to the periods 1,2,...,T respectively, and let the set I_t denote the members of the households selected from the population in the t'th period. Adopting this notation, we may distinguish between three types of combined cross section/time series data, which can be formally described as follows:

C 1. CS/TS data with disjointed samples:

Data in which the sets $I_t \cap I_s$ are empty for all t and $s \neq t$.

C 2. Complete CS/TS data:

Data in which $I_1 = I_2 = \ldots\ldots = I_T$.

C 3. Incomplete CS/TS data:

Data in which the sets $I_1, I_2, \ldots, I_T$ are not identical, and the sets $I_t \cap I_s$ are not empty for all t and $s \neq t$. "Rotation samples" i.e., samples for which $I_1 \cap I_2 \cap \ldots \cap I_T$ is empty, whereas $I_t \cap I_{t-1}$ is not empty, (t=2,...,T), constitute a particularly interesting subclass of incomplete CS/TS data.

Stated in words: In CS/TS data with disjointed samples, different individuals are selected for investigation each period; i.e., we find no observations for which k=h and $s \neq t$. In complete CS/TS data, the individuals selected in each of the T periods are identically the same. In incomplete CS/TS data, some, but not all, of the individuals selected in one period are included in the sample for one or more of the subsequent periods. In rotation samples, in particular, some of the individuals selected in period 1 are included also in period 2, while the remaining ones are replaced by a fresh sample drawn from the (updated) population. A subset of this sample is kept for investigation in period 3 and combined with a fresh sample from the (updated) population in period 2, etc. Thus, in complete as well as in incomplete CS/TS data sets, observations for which k=h and $s \neq t$ exist.

Obviously, the form of the variance/covariance matrix of the complete sample vector of disturbances does strongly depend on the choice of sampling plan. We have already pointed out the differences in this respect between CS data, TS data, and CS/TS data. In sections 3.3-3.5, we shall discuss the covariance structure within the CS/TS class, assuming, for simplicity, that all data sets include H observations (individual household reports) from each of the periods 1,2,...,T. I.e., the total number of observations is HT. Assume, further, that the individuals are numbered consecutively from no. 1 onwards, and that the number of individuals in the population by far exceeds HT.

We shall consider the following three sampling schemes:

CS/TS data with disjointed samples:

$I_1 = \{1,2,\ldots,H\}$,

$I_2 = \{H + 1, H + 2,\ldots,2H\}$,

:

:

$I_t = \{(t-1) H + 1, (t-1) H + 2, \ldots,tH\}$,

:

:

$I_T = \{(T-1) H + 1, (T-1) H + 2,\ldots,TH\}$.

Complete CS/TS data:

$I_t = \{1,2,\ldots,H\} \quad (t=1,2,\ldots,T)$.

Rotation samples with one half of the individuals "in rotation" each period:

$I_1 = \{1,2,\ldots,H\}$,

$I_2 = \{\frac{H}{2} + 1,\frac{H}{2} + 2,\ldots,\frac{3H}{2}\}$,

:

:

$I_t = \{(t-1) \frac{H}{2} + 1, (t-1) \frac{H}{2} + 2,\ldots, (t+1) \frac{H}{2}\}$,

:

:

$I_T = \{(T-1) \frac{H}{2} + 1, (T-1) \frac{H}{2} + 2,\ldots,(T+1) \frac{H}{2}\}$.

(H is supposed to be an even number.)

The total number of individuals investigated is strongly different, being TH when using disjointed samples, H when using complete CS/TS data, and (T+1)H/2 when using rotation samples. Thus, the first sampling plan involves the least "intensive" investigation of the micro units, the second is the most "intensive", while the third is situated "in between".

In the discussion so far, we have used the subscript h to indicate the number of the individual (household) in the population. From now on, this letter will be reserved to denote the number of the individual observation (household report) in the sample from each period under investigation, i.e. ε_{ht} is reinterpreted as the disturbance of the h'th observation (household report) collected in the t'th period $(h=1,\ldots,H;t=1,\ldots,T)$. Indicating the disturbances relating to the individuals as numbered in the population by the superscript Δ, we thus have:

(i) disjointed samples

$$\varepsilon_{h1} = \varepsilon^{\Delta}_{h1}, \qquad (h=1,\dots,H),$$
$$\varepsilon_{h2} = \varepsilon^{\Delta}_{h+H,2},$$
etc.

(ii) complete CS/TS data

$$\varepsilon_{ht} = \varepsilon^{\Delta}_{ht} \qquad (h=1,\dots,H;\ t=1,\dots,T),$$

(iii) rotation samples

$$\varepsilon_{h1} = \varepsilon^{\Delta}_{h1},$$
$$\varepsilon_{h2} = \varepsilon^{\Delta}_{h+H/2,2}, \qquad (h=1,\dots,H),$$
etc.

In this way, we ensure that the values of the subscript variables h and t form a HxT matrix regardless of the choice of sampling plan (provided, of course, that the sample includes H observations from each period).

3.3 The disturbance variance/covariance matrix: CS/TS data with disjointed samples,

Let ε_t denote the (column) vector of disturbances relating to the individuals observed in the t'th period, i.e.,

$$(3.7) \qquad \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \vdots \\ \varepsilon_{Ht} \end{pmatrix} \qquad (t=1,\dots,T).$$

We find

$$(3.8) \qquad E(\varepsilon_t \varepsilon_t') = \begin{pmatrix} \sigma^2 & \sigma_T^2 & \dots\dots & \sigma_T^2 \\ \sigma_T^2 & \sigma^2 & \dots\dots & \sigma_T^2 \\ \vdots & \vdots & & \vdots \\ \sigma_T^2 & \sigma_T^2 & \dots\dots & \sigma^2 \end{pmatrix} = \sigma^2 A \qquad (t=1,\dots,T),$$

where A is the HxH matrix

$$(3.9) \qquad A = \begin{pmatrix} 1 & \omega & \cdots\cdots & \omega \\ \omega & 1 & \cdots\cdots & \omega \\ \omega & \omega & \cdots\cdots & 1 \end{pmatrix}$$

with ω representing the proportion of the total variance σ^2 which is due to the time specific component. (Compare eq. (3.6).) In compact notation, the matrix A may be written as

$$(3.9a) \qquad A = (1-\omega)\, I_H + \omega(e_H e_H'),$$

where I_H is the identity matrix of order H and e_H is the Hx1 unit vector (i.e., the vector consisting entirely of ones). Moreover, since all observations relate to different individuals, all disturbances with different time subscripts are uncorrelated, i.e.,

$$(3.10) \qquad E(\varepsilon_t \varepsilon_s') = 0_{H,H} \qquad (t=1,\ldots,T; s \neq t),$$

where $0_{H,H}$ is the HxH zero matrix.

Defining the THx1 vector

$$(3.11) \qquad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{pmatrix}$$

consisting of all the disturbances in the sample ordered first by period and then by observation within each period, the variance/covariance structure may be expressed as

$$(3.12) \qquad E(\varepsilon\varepsilon') = \sigma^2 \begin{pmatrix} A & 0 & \cdots & 0 \\ 0 & A & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A \end{pmatrix},$$

or, by using the Kronecker product operator $\otimes$ as

$$(3.12a) \qquad E(\varepsilon\varepsilon') = \sigma^2 \Omega_D = \sigma^2 I_T \otimes A,$$

where I_T is the identity matrix of order T. (The subscript D is an abbreviation of "disjointed".) From (3.9a) and (3.12a) and well-known properties of Kronecker products[1] follows

1) See e.g., Theil [17], pp. 303-306.

(3.13) $E(\varepsilon\varepsilon') = \sigma^2\Omega_D = \sigma^2\{(1-\omega)I_{TH} + \omega(I_T \otimes e_H e_H')\}$,

where $I_{TH} = I_T \otimes I_H$ is the identity matrix of order TH.

We conclude that the variance/covariance matrix of the disturbances when using CS/TS data with disjointed samples is block-diagonal with identical blocks, whose elementes may be expressed fairly simply by the two parameters σ and ω. If no time specific effect is present, i.e. $\omega = 0$, the matrix degenerates to the diagonal matrix $\sigma^2 I_{TH}$.

3.4 The disturbance variance/covariance matrix: Complete CS/TS data

Using this sort of data, where all the T subsamples contain identically the same individuals (i.e., the subscript h identifies the individual), the variance/covariance matrix of the disturbances may be formulated in two different ways, corresponding to two different ways of ordering the disturbances: (i) ordering first by period, second by individual, and (ii) ordering first by individual, second by period. Principle (ii) is the one commonly used in the context of complete CS/TS data (see, for instance, Nerlove [13], [14]), and it may, of course , be applied also when dealing with rotation designs. However, in the latter case, principle (i) appears to be the most convenient, giving somewhat simpler algebra. (This is at least the case with the particular sampling plan considered in section 3.5.)[2] In this section, both principles will be considered, partly for the sake of completeness, and partly in order to facilitate comparisons with sections 3.3 and 3.5.

Needless to say, the distinction between the different ways of ordering the observations is of formal significance only: Changes in the ordering of the elements of the ε vector do not affect its density function. Although the variance/covariance matrix $\Omega = E(\varepsilon\varepsilon')$ is changed, neither the value of the determinant $|\Omega|$ nor the quadratic form $\varepsilon'\Omega^{-1}\varepsilon$ is altered. Consequently, the likelihood function of the observations is the same as before.

3.4.1 Ordering first by period, second by individual

We start by noticing that the choice of sampling plan does not influence the properties of the contemporaneous variances and covariances, i.e. the vector ε_t (defined in (3.7)) has the variance/covariance matrix (3.8)

2) As regards disjointed samples, there is only one natural way of ordering the observations. Differences between principles (i) and (ii) do not exist.

with A given by (3.9). The matrices of non-contemporaneous covariances, $E(\varepsilon_t\varepsilon_s')$ $(s\neq t)$, however are not zero matrices as in section 3.3, but scalar matrices,

$$(3.14)\quad E(\varepsilon_t\varepsilon_s') = \sigma^2 \begin{pmatrix} \rho & 0 & \ldots\ldots & 0 \\ 0 & \rho & \ldots\ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots\ldots & \rho \end{pmatrix} = \sigma^2\rho I_H \qquad (t=1,\ldots,T;\ s\neq t),$$

with the diagonal elements representing disturbances relating to the same household (recalling that ρ is the proportion of the total variance which is due to the individual component).

Combining (3.8) and (3.14), we find that the variance/covariance matrix of the disturbance vector ε (as defined in (3.11)) takes the form

$$(3.15)\quad E(\varepsilon\varepsilon') = \sigma^2\Omega_C = \sigma^2 \begin{pmatrix} A & \rho I & \ldots\ldots & \rho I \\ \rho I & A & \ldots\ldots & \rho I \\ \vdots & \vdots & & \vdots \\ \rho I & \rho I & \ldots\ldots & A \end{pmatrix}.$$

(The subscript C is an abbreviation of "complete".) Using Kronecker product notation, this matrix can be written as

$$\sigma^2\Omega_C = \sigma^2\ \{I_T \otimes (A-\rho I_H) + (e_T e_T') \otimes (\rho I_H)\},$$

or, when inserting for A from (3.9a), as

$$(3.16)\quad E(\varepsilon\varepsilon') = \sigma^2\Omega_C = \sigma^2\{(1-\omega-\rho)I_{TH} + \omega(I_T \otimes e_H e_H') + \rho(e_T e_T' \otimes I_H)\}.$$

3.4.2 Ordering first by individual, second by period

Let $\tilde{\varepsilon}_h$ denote the vector of the T disturbances relating to the h'th individual, i.e.

$$(3.17)\quad \tilde{\varepsilon}_h = \begin{pmatrix} \varepsilon_{h1} \\ \varepsilon_{h2} \\ \vdots \\ \varepsilon_{hT} \end{pmatrix} \qquad (h=1,\ldots,H).$$

We find

$$(3.18)\quad E(\tilde{\varepsilon}_h\tilde{\varepsilon}_h') = \begin{pmatrix} \sigma^2 & \sigma_I^2 & \cdots & \sigma_I^2 \\ \sigma_I^2 & \sigma^2 & \cdots & \sigma_I^2 \\ \vdots & \vdots & & \vdots \\ \sigma_I^2 & \sigma_I^2 & \cdots & \sigma^2 \end{pmatrix} = \sigma^2 B \qquad (h=1,\ldots,H),$$

where B is the TxT matrix

$$(3.19)\quad B = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix} = (1-\rho)I_T + \rho(e_T e_T').$$

The matrices of covariances relating to different individuals are diagonal matrices, the diagonal elements representing covariances between disturbances from the same period:

$$(3.20)\quad E(\tilde{\varepsilon}_h\tilde{\varepsilon}_k') = \sigma^2 \begin{pmatrix} \omega & 0 & \cdots & 0 \\ 0 & \omega & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \omega \end{pmatrix} = \sigma^2\omega I_T \qquad (h=1,\ldots,H;\ k\neq h).$$

The HTx1 vector

$$(3.21)\quad \tilde{\varepsilon} = \begin{pmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \\ \vdots \\ \tilde{\varepsilon}_H \end{pmatrix}$$

consisting of all the disturbances ordered first by individual, second by period, thus has the following variance/covariance matrix:

$$(3.22)\quad E(\tilde{\varepsilon}\tilde{\varepsilon}') = \sigma^2\Omega_{C} = \sigma^2 \begin{pmatrix} B & \omega I & \cdots & \omega I \\ \omega I & B & \cdots & \omega I \\ \vdots & \vdots & & \vdots \\ \omega I & \omega I & \cdots & B \end{pmatrix}.$$

This may be written compactly as

$$\sigma^2\Omega_{C*} = \sigma^2\{I_H \otimes (B-\omega I_T) + e_H e_H' \otimes (\omega I_T)\},$$

or, when inserting for B from (3.19), as

$$(3.23)\quad E(\tilde{\tilde{\varepsilon}}\tilde{\varepsilon}') = \sigma^2\Omega_{C*} = \sigma^2\{(1-\rho-\omega)I_{HT} + \rho(I_H \otimes e_T e_T') + \omega(e_H e_H' \otimes I_T)\}.$$

Disregarding a few differences with respect to the choice of symbols, this expression is identical with the one derived by Nerlove [ ] (eq. (1.7), p. 385).

Summing up: 1. The variance/covariance matrix of ε, as well as that of $\tilde{\varepsilon}$, have identical blocks, of dimensions TxT and HxH respectively, along the main diagonal. The blocks outside the main diagonal are identical scalar matrices.

2. In the absence of individual effects, i.e. $\rho=0$, all off-diagonal blocks of the variance/covariance matrix of ε become zero, and the matrix is identical with the one we get when using disjointed samples ($\Omega_C = \Omega_D$).

3. In the absence of time specific effects, i.e. $\omega=0$, all off-diagonal blocks of the variance/covariance matrix of $\tilde{\varepsilon}$ become zero.

4. In the absence of both individual and time specific effects, the variance/covariance matrice of ε and that of $\tilde{\varepsilon}$ are both scalar matrices ($\Omega_C = \Omega_{C*} = I_{TH}$).

3.5 The disturbance variance/covariance matrix: Rotation samples

We focus on the particular rotation scheme outlined in section 3.2, i.e., the one in which half of the H individuals reporting in period t-1 also report in period t (t=2,...,T). For the sake of convenience, the observations will be ordered first by period, second by individual. (Cf. the beginning of section 3.4. In appendix A , however, we shall briefly discuss the opposite ordering, confining our attention to the individuals reporting twice, i.e. omitting those reporting in period 1 only and those reporting in period T only.)

It is readily observed (a) that the matrices of contemporaneous variances/covariances, $E(\varepsilon_t\varepsilon_t')$, have the form (3.8); and (b) that matrices of covariances between vectors ε_t and ε_s more than one period apart are zero matrices, since all elements relate to different individuals, i.e.

$$(3.24) \quad E(\varepsilon_t\varepsilon_s') = 0_{H,H} \qquad (t=1,\ldots,T;)\ (s=1,\ldots,t-2,t+2,\ldots,T)$$

The only difference between this case and the case with disjointed samples relates to the structure of the covariance matrices of vectors for two adjoining periods. Recalling that the h'th observation in period t and the (h+H/2)'th observation in period t-1 come from the same individual, we have $E(\varepsilon_{ht}\varepsilon_{h+H/2,t-1}) = \sigma_I^2$ (h=1,...,H/2; t=2,...,T). Similarly, $E(\varepsilon_{ht}\varepsilon_{h-H/2,t+1}) = \sigma_I^2$ (h=H/2+1,...,H; t=1,...,T-1). Thus,

$$(3.25)\quad \begin{cases} E(\varepsilon_t\varepsilon_{t-1}') = \sigma_I^2\begin{pmatrix}0 & I\\ 0 & 0\end{pmatrix} = \sigma^2\rho C_H & (t=2,\ldots,T) \\[2ex] E(\varepsilon_t\varepsilon_{t+1}') = \sigma_I^2\begin{pmatrix}0 & 0\\ I & 0\end{pmatrix} = \sigma^2\rho C_H' & (t=1,\ldots,T-1), \end{cases}$$

where C_H is the HxH matrix

$$(3.26) \quad C_H = \begin{pmatrix}0 & I\\ 0 & 0\end{pmatrix}$$

(The four submatrices of C_H are of the orders H/2 x H/2.)

Combining (3.8), (3.24) and (3.25), we get

$$(3.27) \quad E(\varepsilon\varepsilon') = \sigma^2\Omega_R = \sigma^2\begin{pmatrix} A & \rho C_H' & 0 & \cdots & 0 & 0 \\ \rho C_H & A & \rho C_H' & \cdots & 0 & 0 \\ 0 & \rho C_H & A & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots' \\ 0 & 0 & 0 & \cdots & A & \rho C_H \\ 0 & 0 & 0 & \cdots & \rho C_H & A \end{pmatrix}.$$

(The subscript R is an abbreviation of "rotating".)

This matrix can also be written in compact Kronecker product notation by introducing the TxT matrix

$$(3.28) \quad D_T = \begin{pmatrix} 0_{1,T-1} & 0 \\ I_{T-1} & 0_{T-1,1} \end{pmatrix},$$

i.e., the identity matrix of order T-1 bordered by zero vectors in the first row and the last column. We find

$$(3.29)\quad E(\varepsilon\varepsilon') = \sigma^2\Omega_R = \sigma^2\{I_T \otimes A + D_T \otimes (\rho C_H) + D_T' \otimes (\rho C_H')\};$$

$$= \sigma^2\{(1-\omega)I_{TH} + \omega(I_T \otimes e_H e_H') + \rho(D_T \otimes C_H + D_T' \otimes C_H')\},$$

when inserting for A from (3.9a).

By partitioning A as follows

$$(3.30)\quad A = \begin{pmatrix} (1-\omega)I_{H/2} + \omega E & \omega E \\ \omega E & (1-\omega)I_{H/2+\omega E} \end{pmatrix} = \begin{pmatrix} A_* & \omega E \\ \omega E & A_* \end{pmatrix}$$

where $E = e_{H/2}e_{H/2}'$, and using (3.25) the matrix Ω_R can be reformulated as

$$(3.31)\quad \Omega_R = \begin{pmatrix} A_* & \omega E & 0 & 0\cdots 0 & 0 & 0 \\ \omega E & A_* & \rho I & 0\cdots 0 & 0 & 0 \\ 0 & \rho I & A_* & \omega E\cdots 0 & 0 & 0 \\ 0 & 0 & \omega E & A_*\cdots 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0\cdots\cdots & A_* & \rho I & 0 \\ 0 & 0 & 0\cdots\cdots & \rho I & A_* & \omega E \\ 0 & 0 & 0\cdots\cdots & 0 & \omega E & A_* \end{pmatrix}$$

where all submatrices have dimension H/2 x H/2.

The properties of Ω_R can be stated as follows:
1) The main diagonal consists of 2T identical blocks, each having the form A_*. 2) The sub-diagonals just below and above the main diagonal consist of 2T-1 blocks alternating between ωE and ρI, beginning (and ending) with ωE. 3) The remaining submatrices of Ω_R are zero matrices. From this follows: 4) In the absence of individual effets, i.e. $\rho=0$, then Ω_R is block-diagonal with T identical HxH dimensional blocks of the form $A = \begin{pmatrix} A_* & \omega E \\ \omega E & A_* \end{pmatrix}$.

5) In the absence of time specific effects, i.e. $\omega=0$, then Ω_R is block-diagonal with T+1 blocks; the first and last ones are identity matrices of dimension H/2 x H/2, the remaining are HxH matrices of the form $\begin{pmatrix} I & \rho I \\ \rho I & I \end{pmatrix}$ (recalling that $A_* = I_{H/2}$ when $\omega = 0$). 6) In the absence of both individual and time specific effects, Ω_R degenerates to the identity matrix of dimension HT x HT.

4. The simultaneous equations (multi-commodity) model approach

In this chapter, we generalize the one-commodity model in chapter 3 to the case where a complete set of consumer demand equations is specified. We start, as in chapter 3, with discussing the covariance structure of the disturbances corresponding to different sampling schemes (sections 4.1-4.5). The final section (section 4.6) deals with interesting special cases and compares the main results derived in this chapter with those obtained in chapter 3.

4.1. The structure of disturbances: Basic assumptions

The disturbance of the demand function for the i'th commodity relating to the h'th household and the t'th period, ε_{iht}, may be decomposed into a household specific component, a period specific component, and a combined component (a remainder) (cf. eq. (2.5)):

(4.1) $\varepsilon_{iht} = u_{ih} + v_{it} + w_{iht}$ for all h and t, and i=1,...,N.

We assume that all components have zero expectations:

(4.2) $E(u_{ih}) = E(v_{it}) = E(w_{iht}) = 0$ for all h and t, and i=1,...,N,

and that the second order moments satisfy

(4.3a) $E(u_{ih}u_{jk}) = \delta_{hk}\sigma_{ij}^{I}$,

(4.3b) $E(v_{it}v_{js}) = \delta_{ts}\sigma_{ij}^{T}$,

(4.3c) $E(w_{iht}w_{jks}) = \delta_{hk}\delta_{ts}\sigma_{ij}^{C}$,

(4.3d) $E(u_{ih}v_{js}) = E(u_{ih}w_{jks}) = E(v_{it}w_{jks}) = 0$

for all h,k,s,t, and i,j = 1,...,N.

The superscripts I, T, and C symbolize individual (household specific), time (period) specific, and combined, respectively. Assumptions (4.3a-d) imply: (i) homoscedasticity of all components of the disturbances ($var(u_{ih}) = \sigma_{ii}^I$, $var(v_{it}) = \sigma_{ii}^T$, $var(w_{iht}) = \sigma_{ii}^C$ for all h and t), (ii) constant covariances between components relating to different commodities but to the same household and period ($cov(u_{ih},u_{jh}) = \sigma_{ij}^I$, $cov(v_{it},v_{jt}) = \sigma_{ij}^T$, $cov(w_{iht},w_{jht}) = \sigma_{ij}^C$ for all h and t), and (iii) no correlation between components relating to different households and/or periods). These are straightforward generalizations of the corresponding single-equation assumptions in chapter 3, (3.1)-(3.3).

From (4.1)-(4.3) follows

$$E(\varepsilon_{iht}) = 0, \tag{4.4}$$

$$E(\varepsilon_{iht}\varepsilon_{jks}) = \delta_{hk}\sigma_{ij}^I + \delta_{ts}\sigma_{ij}^T + \delta_{hk}\delta_{ts}\sigma_{ij}^C, \tag{4.5}$$

or more explicitly

$$E(\varepsilon_{iht}\varepsilon_{jks}) = \begin{cases} \sigma_{ij}^I + \sigma_{ij}^T + \sigma_{ij}^C & \text{for } k=h \;\&\; s = t, \\ \sigma_{ij}^I & \text{for } k=h \;\&\; s \neq t, \\ \sigma_{ij}^T & \text{for } k\neq h \;\&\; s = t, \\ 0 & \text{for } k\neq h \;\&\; s \neq t. \end{cases} \tag{4.5a}$$

Defining

$$\begin{cases} \sigma_{ij} = \sigma_{ij}^I + \sigma_{ij}^T + \sigma_{ij}^C \\ \rho_{ij} = \sigma_{ij}^I/\sigma_{ij} \\ \omega_{ij} = \sigma_{ij}^T/\sigma_{ij} \end{cases} \qquad (i,\ j=1,\dots,N), \tag{4.6}$$

the variance/covariance structure may alternatively be expressed as

$$(4.5b)\quad E(\varepsilon_{iht}\varepsilon_{jks}) = \begin{cases} \sigma_{ij} & \text{for } k=h \ \&\ s=t, \\ \rho_{ij}\sigma_{ij} & \text{for } k=h \ \&\ s\neq t, \\ \omega_{ij}\sigma_{ij} & \text{for } k\neq h \ \&\ s=t, \\ 0 & \text{for } k\neq h \ \&\ s\neq t, \end{cases}$$

Owing to the adding-up restrictions (2.7), the σ's have to satisfy

$$(4.7)\quad \sum_{i=1}^{N} \sigma_{ij} = \sum_{i=1}^{N} (\sigma_{ij}^{I} + \sigma_{ij}^{T} + \sigma_{ij}^{C}) = 0 \qquad (j=1,\dots,N).$$

We shall also consider the stronger set of restrictions

$$(4.8)\quad \sum_i \sigma_{ij}^{I} = \sum_i \sigma_{ij}^{T} = \sum_i \sigma_{ij}^{C} = 0 \qquad (j=1,\dots,N).$$

Before proceeding further, it is convenient to rewrite the formulae above in matrix notation. Defining the Nx1 disturbance vectors

$$(4.9)\quad \varepsilon_{ht} = \begin{pmatrix} \varepsilon_{1ht} \\ \varepsilon_{2ht} \\ \vdots \\ \varepsilon_{Nht} \end{pmatrix}, \quad u_h = \begin{pmatrix} u_{1h} \\ u_{2h} \\ \vdots \\ u_{Nh} \end{pmatrix}, \quad v_t = \begin{pmatrix} v_{1t} \\ v_{2t} \\ \vdots \\ v_{Nt} \end{pmatrix}, \quad w_{ht} = \begin{pmatrix} w_{1ht} \\ w_{2ht} \\ \vdots \\ w_{Nht} \end{pmatrix},$$

and the NxN matrices of "contemporaneous" variances/covariances

$$(4.10)\quad \begin{cases} \Sigma^{(I)} = (\sigma_{ij}^{I}), \quad \Sigma^{(T)} = (\sigma_{ij}^{T}), \quad \Sigma^{(C)} = (\sigma_{ij}^{C}), \\ \Sigma = (\sigma_{ij}) = \Sigma^{(I)} + \Sigma^{(T)} + \Sigma^{(C)}, \end{cases}$$

equations (4.1)-(4.5), (4.7) and (4.8) may be formulated as

$$(4.1*)\quad \varepsilon_{ht} = u_h + v_t + w_{ht},$$

$$(4.2*)\quad E(u_h) = E(v_t) = E(w_{ht}) = 0_N,$$

$$(4.3a*)\quad E(u_h u_k') = \delta_{hk}\Sigma^{(I)},$$

(4.3b *) $E(v_t v_s') = \delta_{ts}\Sigma^{(T)}$,

(4.3c *) $E(w_{ht} w_{ks}') = \delta_{hk}\delta_{ts}\Sigma^{(C)}$,

(4.3d *) $E(u_h v_s') = E(u_h w_{ks}') = E(v_t w_{ks}') = 0_{N,N}$,

(4.4 *) $E(\varepsilon_{ht}) = 0_N$,

(4.5 *) $E(\varepsilon_{ht}\varepsilon_{ks}') = \delta_{hk}\Sigma^{(I)} + \delta_{ts}\Sigma^{(T)} + \delta_{hk}\delta_{ts}\Sigma^{(C)}$,

(4.5a *) $$E(\varepsilon_{ht}\varepsilon_{ks}') = \begin{cases} \Sigma & \text{for } k=h \text{ \& } s=t, \\ \Sigma^{(I)} & \text{for } k=h \text{ \& } s\neq t, \\ \Sigma^{(T)} & \text{for } k\neq h \text{ \& } s=t, \\ 0_{N,N} & \text{for } k\neq h \text{ \& } s\neq t, \end{cases}$$

(for all h, k, t, and s)

(4.7 *) $\Sigma e_N = 0_N$,

(4.8 *) $\Sigma^{(I)} e_N = \Sigma^{(T)} e_N = \Sigma^{(C)} e_N = 0_N$.

Here $0_{N,N}$ and 0_N are zero matrices of orders NxN and Nx1 respectively, and e_N is the Nx1 unit vector.

4.2 The sampling schemes

Again we shall consider the following three sampling schemes: (i) cross-section/time-series (CS/TS) data with disjointed samples, (ii) complete CS/TS data sets, and (iii) rotation samples with one half of the individuals "in rotation" each period. (Compare the formal definitions of the sets $I_1,\ldots,I_T$ given in section 3.2.) Using from now on the symbol ε_{iht} to denote the disturbance of the i'th demand function in period t relating to the h'th observation (household report) from this period, and letting the superscript Δ indicate the corresponding disturbances when the individuals (households) are numbered as in the population, we have the following correspondence:

(i) disjointed samples

$$\varepsilon_{ih1} = \varepsilon^{\Delta}_{ih1},$$

$$\varepsilon_{ih2} = \varepsilon^{\Delta}_{i,h+H,2}, \qquad (h=1,\ldots,H)$$

etc.

(ii) complete CS/TS data

$$\varepsilon_{iht} = \varepsilon^{\Delta}_{iht} \qquad (h=1,\ldots,H;\ t=1,\ldots,T),$$

(iii) rotation samples

$$\varepsilon_{ih1} = \varepsilon^{\Delta}_{ih1},$$

$$\varepsilon_{ih2} = \varepsilon^{\Delta}_{i,h+H/2,2}, \qquad (h=1,\ldots,H)$$

etc.

4.3 The disturbance variance/covariance matrix: Disjointed CS/TS samples

Let ε_t be the HNx1 vector of disturbances of the individuals observed in period t, ordered first by individual, second by commodity, i.e.,

$$(4.11) \quad \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \vdots \\ \vdots \\ \varepsilon_{Ht} \end{pmatrix} \qquad (t=1,\ldots,T),$$

where the subvectors ε_{ht} are defined in (4.9). From (4.5a*) follows

$$(4.12) \quad E(\varepsilon_t\varepsilon_t') = \begin{pmatrix} \Sigma & \Sigma^{(T)} & \cdots & \Sigma^{(T)} \\ \Sigma^{(T)} & \Sigma & \cdots & \Sigma^{(T)} \\ \vdots & \vdots & & \vdots \\ \Sigma^{(T)} & \Sigma^{(T)} & \cdots & \Sigma \end{pmatrix}$$

$$\doteq I_H \otimes (\Sigma-\Sigma^{(T)}) + E_H \otimes \Sigma^{(T)} = J \qquad (t=1,\ldots,T),$$

where $E_H (= e_H e_H')$ is the HxH matrix consisting entirely of ones, and where the last equality defines J. Furthermore

$$(4.13)\quad E(\varepsilon_t\varepsilon_s') = 0_{HN,HN} \qquad (t=1,\ldots,T;\ s\neq t),$$

since all disturbances with different time subscripts come from different individuals.

Let

$$(4.14)\quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \vdots \\ \vdots \\ \varepsilon_T \end{pmatrix}$$

denote the THNx1 vector consisting of all disturbances ordered first by period, second by individual, and third by commodity. Its variance/covariance matrix is

$$(4.15)\quad E(\varepsilon\varepsilon') = \Lambda_D = \begin{pmatrix} J & 0 & \cdots & 0 \\ 0 & J & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & J \end{pmatrix} = I_T \otimes J$$

$$= I_{TH} \otimes (\Sigma - \Sigma^{(T)}) + I_T \otimes E_H \otimes \Sigma^{(T)},$$

after inserting for J from (4.12). (As previously, the subscript D is an abbreviation of "disjointed"). This equation generalizes eq. (3.13). In the absence of time specific effects, Λ_D takes the form $I_{TH} \otimes \Sigma$.

4.4 The disturbance variance/covariance matrix: Complete CS/TS data

In this section, as in section 3.4, we shall consider ordering the observations first by period, second by individual, as well as the opposite ordering. The ordering by commodity is supposed to take place after the ordering by period and individual; i.e. we discuss, as in section 4.3, the ordering of the vectors ε_{ht} (defined in (4.9)). In appendix B, we shall, however, briefly comment on the covariance structure when the ordering by commodity precedes the ordering by period and individual.

4.4.1 Ordering first by period, second by individual

The vector ε_t, as defined in (4.11), obviously has the variance/covariance matrix (4.12) in the present case as well. All the matrices of non-contemperaneous covariances are of the form

$$(4.16)\quad E(\varepsilon_t\varepsilon_s') = \begin{pmatrix} \Sigma^{(I)} & 0 & \cdots & 0 \\ 0 & \Sigma^{(I)} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \Sigma^{(I)} \end{pmatrix} = $$

$$= I_H \otimes \Sigma^{(I)} = L \qquad (t=1,\ldots,T; \quad s\neq t),$$

where the blocks along the main diagonal contain covariances relating to the same households. (The zero matrices have dimension NxN.)

Combining (4.12) and (4.16), we get

$$(4.17)\quad E(\varepsilon\varepsilon') = \Lambda_C = \begin{pmatrix} J & L & \cdots & L \\ L & J & \cdots & L \\ \vdots & \vdots & & \vdots \\ L & L & \cdots & J \end{pmatrix} = $$

$$= I_T \otimes (J-L) + E_T \otimes L$$

with ε defined in (4.14). (The subscript C is an abbreviation of "complete".) Inserting for J and L from (4.12) and (4.16) yields

$$(4.18)\quad E(\varepsilon\varepsilon') = \Lambda_C = I_{TH} \otimes (\Sigma-\Sigma^{(T)} - \Sigma^{(I)}) + I_T \otimes E_H \otimes \Sigma^{(T)} +$$

$$+ E_T \otimes I_H \otimes \Sigma^{(I)},$$

which is a generalization of equation (3.16).

4.4.2 Ordering first by individual, second by period

The TNx1 vector of disturbances relating to the h'th individual

$$(4.19)\quad \tilde{\varepsilon}_h = \begin{pmatrix} \varepsilon_{h1} \\ \varepsilon_{h2} \\ \vdots \\ \vdots \\ \varepsilon_{hT} \end{pmatrix} \qquad (h=1,\ldots,H)$$

with subvectors ε_{ht} defined in (4.9), has the following variance/covariance matrix:

$$(4.20)\quad E(\tilde{\varepsilon}_h\tilde{\varepsilon}_h') = \begin{pmatrix} \Sigma & \Sigma^{(I)} & \cdots & \Sigma^{(I)} \\ \Sigma^{(I)} & \Sigma & \cdots & \Sigma^{(I)} \\ \vdots & \vdots & & \vdots \\ \Sigma^{(I)} & \Sigma^{(I)} & \cdots & \Sigma \end{pmatrix}$$

$$= I_T \otimes (\Sigma - \Sigma^{(I)}) + E_T \otimes \Sigma^{(I)} = K \qquad (h=1,\ldots,H),$$

the last equality defining K. The matrices of covariances relating to different individuals are block diagonal matrices

$$(4.21)\quad E(\tilde{\varepsilon}_h\tilde{\varepsilon}_k') = \begin{pmatrix} \Sigma^{(T)} & 0 & \cdots & 0 \\ 0 & \Sigma^{(T)} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \Sigma^{(T)} \end{pmatrix}$$

$$= I_T \otimes \Sigma^{(T)} = M \qquad (h=1,\ldots,H;\ k\neq h),$$

the last equality defining M. The blocks along the main diagonal contain covariances between observations from the same period. (The zero matrices have dimension NxN.)

The HTNx1 vector

$$(4.22)\quad \tilde{\varepsilon} = \begin{pmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \\ \vdots \\ \tilde{\varepsilon}_H \end{pmatrix} ,$$

containing all disturbances ordered first by individual, second by period, and third by commodity, thus has the variance/covariance matrix

$$(4.23)\quad E(\tilde{\varepsilon}\tilde{\varepsilon}') = \Lambda_{C*} = \begin{pmatrix} K & M & \cdots & M \\ M & K & \cdots & M \\ \vdots & \vdots & & \vdots \\ M & M & \cdots & K \end{pmatrix} =$$

$$= I_H \otimes (K-M) + E_H \otimes M,$$

or

$$(4.24)\quad E(\tilde{\varepsilon}\tilde{\varepsilon}') = \Lambda_{C*} = I_{HT} \otimes (\Sigma-\Sigma^{(I)} - \Sigma^{(T)}) + I_H \otimes E_T \otimes \Sigma^{(I)} + $$

$$+ E_H \otimes I_T \otimes \Sigma^{(T)} ,$$

when inserting for K and M from (4.20) and (4.21). This equation generalizes (3.23).

4.5 The disturbance variance/covariance matrix: Rotation samples

In view of the results derived in sections 3.3-3.5, and 4.3-4.4, the variance/covariance formulae corresponding with rotation samples may be readily established. First, we notice that the matrices of contemporaneous variances/covariances are identical with those in sections 4.3 and 4.4, i.e. $E(\varepsilon_t\varepsilon_t') = J = I_H \otimes (\Sigma-\Sigma^{(T)}) + E_H \otimes \Sigma^{(T)}$. Second, all matrices of covariances between disturbance vectors more than one period apart are zero matrices, since their elements relate to different individuals, i.e.

$$(4.25)\quad E(\varepsilon_t\varepsilon_s') = 0_{HN,HN} \qquad (t=1,\ldots,T;s=1,\ldots,t-2,\ t+2,\ldots,T).$$

Third, $E(\varepsilon_{ht}\varepsilon'_{h+H/2,t-1}) = \Sigma^{(I)}$ $(h=1,\ldots,H/2;\ t=2,\ldots,T)$,

and $E(\varepsilon_{ht}\varepsilon'_{h-H/2,t+1}) = \Sigma^{(I)}$ $(h=H/2+1,\ \ldots,\ H;\ t=1,\ldots,T-1)$,

all other submatrices of $E(\varepsilon_t\varepsilon'_{t-1})$ and $E(\varepsilon_t\varepsilon'_{t+1})$ are zero matrices, i.e.

$$(4.26)\begin{cases} E(\varepsilon_t\varepsilon'_{t-1}) = \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix} \otimes \Sigma^{(I)} = C_H \otimes \Sigma^{(I)} , \\ E(\varepsilon_t\varepsilon'_{t+1}) = \begin{pmatrix} 0 & 0 \\ I & 0 \end{pmatrix} \otimes \Sigma^{(I)} = C'_H \otimes \Sigma^{(I)} \end{cases}$$

(cf. (3.25) and (3.26)).

Thus, the complete variance/covariance matrix takes the form

$$(4.27)\ E(\varepsilon\varepsilon') = \Lambda_R = \begin{pmatrix} J & Q' & 0 & \cdots & 0 & 0 \\ Q & J & Q' & \cdots & 0 & 0 \\ 0 & Q & J & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & J & Q' \\ 0 & 0 & 0 & \cdots & Q & J \end{pmatrix},$$

where $Q = C_H \otimes \Sigma^{(I)}$, or in Kronecker product notation

$$(4.28)\ E(\varepsilon\varepsilon') = \Lambda_R = I_T \otimes J + D_T \otimes Q + D'_T \otimes Q'$$
$$= I_{TH} \otimes (\Sigma - \Sigma^{(T)}) + I_T \otimes E_H \otimes \Sigma^{(T)}$$
$$+ (D_T \otimes C_H + D'_T \otimes C'_H) \otimes \Sigma^{(I)} .$$

4.6 Special cases: comparison with the single-equation model

The THN x THN variance/covariance matrices $\Lambda_D, \Lambda_C, \Lambda_{C*}$, and Λ_R (defined in eqs. (4.15), (4.18), (4.24), and (4.28)) are in general rather complicated and may prove inconvenient to deal with empirically, at least when T, H, and N are not very small. This motivates simplifying the structure by imposing restrictions on $\Sigma^{(I)}$, $\Sigma^{(T)}$, and $\Sigma^{(C)}$, reducing the number of free parameters.

What is the maximal number of elements in the matrices $\Sigma^{(I)}$, $\Sigma^{(T)}$, $\Sigma^{(C)}$, and Σ that may be chosen freely? First, the definitional equations $\sigma_{ij} = \sigma_{ij}^{I} + \sigma_{ij}^{T} + \sigma_{ij}^{C}$ and the symmetry conditions $\sigma_{ji}^{I} = \sigma_{ij}^{I}$, $\sigma_{ji}^{T} = \sigma_{ij}^{T}$, $\sigma_{ji}^{C} = \sigma_{ij}^{C}$ yield $N^2 + 3N(N-1)/2 = (5N^2 - 3N)/2$ restrictions. Second, the adding-up conditions (4.7), or the stronger set (4.8), impose N and 3N restrictions respectively. The maximal number of "degrees of freedom" is thus

$$4N^2 - (\tfrac{5}{2}N^2 - \tfrac{3}{2}N + N) = \tfrac{1}{2}N\,(3N+1)$$

when using (4.7), and

$$4N^2 - (\tfrac{5}{2}N^2 - \tfrac{3}{2}N + 3N) = \tfrac{3}{2}N\,(N-1)$$

when using (4.8).

Suppose that the matrices (ρ_{ij}) and (ω_{ij}), containing the shares in the "total" variance/covariance of the individual and time specific components respectively, have one set of "row specific" and one set of "column specific" components:

$$(4.29)\quad \begin{cases} \rho_{ij} = \lambda_i \lambda_j \\ \omega_{ij} = \mu_i \mu_j \end{cases} \qquad (i,j=1,\ldots,N).$$

(Cf.eq. (4.6).) This implies

$$(4.30)\quad \begin{cases} \sigma_{ij}^{I} = \lambda_i \lambda_j \sigma_{ij} \\ \sigma_{ij}^{T} = \mu_i \mu_j \sigma_{ij}, \end{cases}$$

or

$$(4.30^{*})\quad \begin{cases} \Sigma^{(I)} = \lambda \Sigma \lambda \\ \Sigma^{(T)} = \mu \Sigma \mu, \end{cases}$$

where

$$(4.31)\quad \lambda = \begin{pmatrix} \lambda_1 & 0 \cdots & 0 \\ 0 & \lambda_2 \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 \cdots & \lambda_N \end{pmatrix}, \qquad \mu = \begin{pmatrix} \mu_1 & 0 \cdots & 0 \\ 0 & \mu_2 \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 \cdots & \mu_N \end{pmatrix}.$$

Since $\sigma_{ii}^{C} \geq 0$, the inequalities $\lambda_i^2 + \mu_i^2 \leq 1$ should be satisfied for all i.

With this reparametrization, the strong set of adding-up restrictions, (4.8), will not be satisfied in general: $\Sigma_i \sigma_{ij}^{I} = \Sigma_i \sigma_{ij}^{T} = 0$ would imply $\Sigma_i \lambda_i \sigma_{ij} = \Sigma_i \mu_i \sigma_{ij} = 0$ (j=1,...,N), which cannot hold unless λ_i and μ_i have the same value for all i (assuming that Σ has rank N-1). The number of degrees of freedom is: N(N-1)/2 (= the maximal number of σ_{ij}'s that can be varied freely) + 2N (= the number of λ_i's and μ_i's), i.e.

$$\frac{1}{2}N\ (N+3).$$

Example: N=3

The number of degrees of freedom in the σ_{ij}'s is N(N-1)/2=3. Choose σ_{11}, σ_{22} and σ_{33} freely. The 6 covariances then must satisfy

$$\sigma_{21} + \sigma_{31} = -\sigma_{11},$$

$$\sigma_{21} + \sigma_{32} = -\sigma_{22},$$

$$\sigma_{31} + \sigma_{32} = -\sigma_{33},$$

$$\sigma_{21} = \sigma_{12},$$

$$\sigma_{31} = \sigma_{13},$$

$$\sigma_{32} = \sigma_{23},$$

which implies

$$\sigma_{12} = \sigma_{21} = (-\sigma_{11} - \sigma_{22} + \sigma_{33})/2,$$

$$\sigma_{13} = \sigma_{31} = (-\sigma_{11} + \sigma_{22} - \sigma_{33})/2,$$

$$\sigma_{23} = \sigma_{32} = (\sigma_{11} - \sigma_{22} - \sigma_{33})/2.$$

In addition, the 6 λ's and μ's can be varied freely. Thus, the total number of degrees of freedom is 3+6=9.

We now impose the following stronger set of constraints:

$$(4.32)\quad \begin{cases} \lambda_i = \sqrt{\rho} \\ \mu_i = \sqrt{\omega} \end{cases} \qquad (i=1,\dots,N),$$

where, of course, ρ and ω are non-negative. In view of (4.30), this implies

$$(4.33)\quad \begin{cases} \sigma_{ij}^{\;I} = \rho\sigma_{ij}, \\ \sigma_{ij}^{\;T} = \omega\sigma_{ij}, \end{cases}$$

or

$$(4.33^*)\quad \begin{cases} \Sigma^{(I)} = \rho\Sigma \\ \Sigma^{(T)} = \omega\Sigma \end{cases}$$

Inserting (4.33*) into eqs. (4.15), (4.18), (4.24), and (4.28), using (3.13) (3.16), (3.23), and (3.29), respectively, we get

$$(4.34)\quad \Lambda_D = \{(1-\omega)I_{TH} + \omega I_T \otimes E_H\} \otimes \Sigma \doteq \Omega_D \otimes \Sigma,$$

$$(4.35)\quad \Lambda_C = \{(1-\omega-\rho)I_{TH} + \omega I_T \otimes E_H + \rho E_T \otimes I_H\} \otimes \Sigma = \Omega_C \otimes \Sigma,$$

$$(4.36)\quad \Lambda_{C^*} = \{(1-\rho-\omega)I_{HT} + \rho I_H \otimes E_T + \omega E_H \otimes I_T\} \otimes \Sigma = \Omega_{C^*} \otimes \Sigma,$$

$$(4.37)\quad \Lambda_R = \{(1-\omega)I_{TH} + \omega I_T \otimes E_H + \rho(D_T \otimes C_H + D_T' \otimes C_H')\} \otimes \Sigma \doteq \Omega_R \otimes \Sigma.$$

In this way, we obtain that the variance/covariance matrix of the complete vector of residuals in the multi-equation model can be written as the Kronecker product of two matrices: one of dimension THxTH and proportional with the variance/covariance matrix in the single-equation model, the other of dimension NxN and equal to the Σ matrix of "contemporaneous" variances/covariances in the multi-equation model. In view of the simple rules that exist for inverting and calculating determinant values of matrices expressed in terms of Kronecker products (cf.eqs. (C.2) and (C.3) in appendix C), this simplification represents a considerable gain when it comes to estimation. The specification (4.32), admittedly restrictive since it implies that the relative importance of the individual and time specific components of the variances and covariances is the same for all commodities, should be interpreted on this background. The structure of the Σ matrices in this case has

$$\frac{1}{2}N\ (N-1) + 2$$

degrees of freedom, and we notice that restrictions (4.8) are satisfied automatically when (4.7) is imposed.

5. Estimation

Combining the specification of the demand structure (chapter 2) with the stochastic specification of the disturbances (chapters 3 and 4) we now proceed to the problem of estimation. We shall first sketch the problem in general terms (section 5.1) and then discuss practically interesting special cases in some detail (sections 5.2 and 5.3).

5.1. Preliminaries: The Full Information Maximum Likelihood (FIML) principle

Provided that the disturbances are normally distributed, the following general scheme contains all the models and situations considered in the previous chapters as particular cases:

$$(5.1)\quad \begin{cases} x = f(z;\beta) + w, \\ \text{where } w \text{ is distributed as } N(0,\Omega). \end{cases}$$

Here, x denotes the vector of budget shares (ordered in a prescribed way), w is the corresponding vector of disturbances, f is a vector function, z is the vector containing all the values of the exogenous variables, β is the vector containing all the coefficients of the budget share functions, and Ω is the variance/covariance matrix of w.

Letting, in general, n denote the dimension of x and w, the log-likelihood function (i.e., the (natural) logarithm of the density function) of x is

$$(5.2)\quad L = L(x,z,\beta,\Omega) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\Omega| - \frac{1}{2}\{x-f(z;\beta)\}'\Omega^{-1}\{x-f(z;\beta)\},$$

provided Ω is non-singular. The symmetry constraint, $\Omega' = \Omega$, leaves $n(n+1)/2$ free elements in Ω. Meaningful estimation requires some additional restrictions on this matrix.

The Full Information Maximum Likelihood (FIML) estimators of (the unknown coefficients of) β and Ω are those which maximise L simultaneously, given the values of x and z. If, in particular, Ω is known up to a factor of proportionality, the FIML estimator of β is found by minimising the quadratic form

$$(5.3) \quad Q = w' \Omega^{-1} w = \{x-f(z;\beta)\}' \Omega^{-1} \{x-f(z;\beta)\} .$$

We shall now turn our attention more specifically to three particular cases:

(i) Disjointed CS/TS data.

(ii) Complete CS/TS data in the absence of time specific disturbance effects.

(iii) Rotation samples in the absence of time specific disturbance effects.

The single equation case is discussed in section 5.2, section 5.3 deals with simultaneous estimation of the complete model in the particular case where the time specific and individual component respectively of the variances/covariances represent the same proportion of the corresponding total: $\sigma_{ij}^T = \omega\sigma_{ij}$, resp. $\sigma_{ij}^I = \rho\sigma_{ij}$ for all i and j; cf. section 4.6.

5.2 Estimation in the single-equation model

5.2.1 Disjointed CS/TS data.[1)]

Let a denote the vector of budget shares of the commodity considered, ordered first by period, second by individual, i.e.

---

1) Evidently, disjointed CS/TS data are formally identical with both complete CS/TS data and data from rotation samples in the absence of individual disturbance components. Thus, the results obtained in this section are valid for the two latter categories of data as well.

$$(5.4) \qquad a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ \vdots \\ a_T \end{pmatrix},$$

where

$$(5.5) \qquad a_t = \begin{pmatrix} a_{1t} \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ a_{Ht} \end{pmatrix} \qquad\qquad (t = 1,\dots, T),$$

omitting, for simplicity, the commodity subscript (cf. sections 2.1, 3.1 and 3.3). We then have the following situation:

$n = TH,$

$x = a,$

$w = \varepsilon$ (as defined in eq. (3.11)),

$\Omega = \sigma^2\Omega_D$ (as defined in eq. (3.13)).

In order to establish the likelihood function, it is necessary to derive expressions for the determinant $|\Omega|$ and the quadratic form Q. From eqs. (3.12a) and (C.3) we obtain

$$|\Omega| = \sigma^{2TH}\ |\Omega_D| = \sigma^{2TH}\ |I_T \otimes A| = \sigma^{2TH}\ |A|^T.$$

By induction we find that the determinant value of A (as defined in eq. (3.9)) is equal to [2]

$$(5.6) \qquad |A| = (1-\omega)^{H-1} \{1 + (H-1)\omega\}.$$

This gives

$$(5.7) \qquad |\Omega| = \sigma^{2TH}\ (1-\omega)^{T(H-1)} \{1 + (H-1)\omega\}^T.$$

Moreover, using eq. (C.6), we find

$$Q = \varepsilon'\ (\sigma^2\Omega_D)^{-1}\ \varepsilon = \frac{1}{\sigma^2}\ \varepsilon'\ (I_T \otimes A)^{-1}\varepsilon = \frac{1}{\sigma^2} \sum_{t=1}^{T} \varepsilon_t'\ A^{-1}\ \varepsilon_t.$$

---

2) Cf. also Balestra [2], p. 125.

The inverse of A is equal to[3)]

$$(5.8)\qquad A^{-1} = \frac{1}{1-\omega}\left(I_H - \frac{\omega}{1+(H-1)\omega}E_H\right).$$

Thus,

$$(5.9)\qquad Q = \frac{1}{\sigma^2(1-\omega)} \sum_{t=1}^{T} \left(\varepsilon_t' I_H \varepsilon_t - \frac{\omega}{1+(H-1)\omega}\varepsilon_t' E_H \varepsilon_t\right)$$

$$= \frac{1}{\sigma^2(1-\omega)} \sum_{t=1}^{T} \left\{ \sum_{h=1}^{H} \varepsilon_{ht}^2 - \frac{\omega}{1+(H-1)\omega}\left(\sum_{h=1}^{H}\varepsilon_{ht}\right)^2 \right\}.$$

Inserting (5.7) and (5.9) into (5.2), we obtain

$$(5.10)\qquad L = -\frac{TH}{2}\log(2\pi) - \frac{1}{2}\{TH \log\sigma^2 + T(H-1)\log(1-\omega) + T\log(1+(H-1)\omega)\} - \frac{1}{2\sigma^2(1-\omega)}\left\{Q_A - \frac{\omega}{1+(H-1)\omega}Q_B\right\},$$

where

$$(5.11)\qquad Q_A = \sum_{t=1}^{T}\sum_{h=1}^{H}\varepsilon_{ht}^2,$$

$$(5.12)\qquad Q_B = \sum_{t=1}^{T}\left\{\sum_{h=1}^{H}\varepsilon_{ht}\right\}^2,$$

interpreting the ε's as shorthand expressions for the corresponding differences when inserting eq. (2.4).

Maximising L partially with respect to σ^2 with ω and β fixed, we find the conditional estimator

$$(5.13)\qquad \hat{\sigma}^2 = \frac{1}{TH(1-\omega)}\left\{Q_A - \frac{\omega}{1+(H-1)\omega}Q_B\right\}.$$

The concentrated log-likelihood function thus has the form (cf. Rothenberg and Leenders [16])

$$(5.14)\qquad L^* = \text{constant} - \frac{1}{2}\{TH\log\{(1+(H-1)\omega)Q_A - \omega Q_B\} - T(H-1)\log(1+(H-1)\omega) - T\log(1-\omega)\}$$

---

3) Cf. Balestra [2], Appendix B, or Nerlove [14], eq. (4.3).

Maximising L^* partially with respect to ω, with β fixed, we get, after some algebra, the conditional estimator

$$(5.15)\quad \hat{\omega} = \frac{1}{H-1}\left(\frac{Q_B}{Q_A}-1\right).$$

Inserting $\omega = \hat{\omega}$ into (5.13) we obtain

$$(5.16)\quad \hat{\sigma}^2 = \frac{Q_A}{TH}.$$

Thus, the FIML estimators of ω and σ^2 satisfy

$$(5.17)\quad \hat{\omega} = \frac{1}{H-1}\;\frac{\sum_{t=1}^{T}\{(\sum_{h=1}^{H}\hat{\varepsilon}_{ht})^2 - \sum_{h=1}^{H}\hat{\varepsilon}_{ht}^2\}}{\sum_{t=1}^{T}\sum_{h=1}^{H}\hat{\varepsilon}_{ht}^2},$$

$$(5.18)\quad \hat{\sigma}^2 = \frac{1}{TH}\sum_{t=1}^{T}\sum_{h=1}^{H}\hat{\varepsilon}_{ht}^2,$$

the 'hats' denoting the residuals calculated from the estimated equation.

Simultaneous FIML estimation of β, ω and σ^2 may prove practically troublesome. The following approximate three-stage procedure, however, seems useful:

(i) Estimate β by (non-linear) OLS, i.e. by minimising Q_A.[4)]

(ii) Estimate ω and σ^2 from the residuals by using (5.17) and (5.18).

(iii) Reestimate β by minimising

$$Q = \frac{1}{\sigma^2(1-\omega)}\left\{Q_A - \frac{\omega}{1+(H-1)\omega}Q_B\right\},$$

with ω set equal to the estimate obtained at stage (ii).[5)]

A practical way to proceed when carrying out stage (iii) is the following: Noticing that the matrix A^{-1} can be factorised as follows[6)]

$$(5.19)\quad A^{-1} = \Phi'\Phi, \text{ where } \Phi = \frac{1}{\sqrt{1-\omega}}\left\{I_H - \frac{1-R_A}{H}E_H\right\} = \Phi'$$

4) Any other consistent method might be used.

5) In order to obtain a better approximation, we might return to stage (ii) and repeat the process.

6) Cf. Balestra [2], section 5.2.3.

with

$$(5.20)\quad R_A = \sqrt{\frac{1-\omega}{1+(H-1)\omega}},$$

then Q can be written as

$$(5.21)\quad Q = \frac{1}{\sigma^2} \sum_{t=1}^{T} (\Phi\varepsilon_t)'(\Phi\varepsilon_t) = \frac{1}{\sigma^2(1-\omega)} \sum_{t=1}^{T} \sum_{h=1}^{H} \{\varepsilon_{ht} - \frac{1-R_A}{H} (\sum_{k=1}^{H} \varepsilon_{kt})\}^2$$

i.e. minimisation of Q is equivalent to minimisation of the sum of squares of the transformed disturbances

$$\varepsilon_{ht} - \frac{1-R_A}{H} (\sum_{k=1}^{H} \varepsilon_{kt}).$$

5.2.2. Complete CS/TS data with no time specific effects.

With reference to the general scheme (5.1), this special case is the following:

$$n = TH,$$
$$x = a,$$
$$w = \varepsilon,$$
$$\Omega = \sigma^2 \Omega_{C/\omega=0} = \sigma^2 \{(1-\rho) I_{TH} + \rho E_T \otimes I_H\} = \sigma^2 B \otimes I_H,$$

where Ω_C and B are defined in eqs. (3.16) and (3.19) respectively. Using the analogy between this case and the previous one, we directly obtain (cf. (5.7) and (5.9))

$$(5.22)\quad |\Omega| = \sigma^{2TH} (1-\rho)^{H(T-1)} \left\{1+(T-1)\rho\right\}^H$$

$$(5.23)\quad Q = \frac{1}{\sigma^2(1-\omega)} \sum_{h=1}^{H} \left\{\sum_{t=1}^{T} \varepsilon_{ht}^2 - \frac{\rho}{1+(T-1)\rho} (\sum_{s=1}^{T} \varepsilon_{hs})^2\right\},$$

which when inserted into (5.2) yields

$$(5.24)\quad L = -\frac{TH}{2} \log (2\pi) - \frac{1}{2}\left\{TH \log \sigma^2 + H(T-1) \log (1-\rho) + H \log (1+(T-1)\rho)\right\}$$
$$- \frac{1}{2\sigma^2 (1-\rho)} \left\{ Q_A - \frac{\rho}{1+(T-1)\rho} Q_C\right\},$$

where Q_A is as defined in (5.11), and

(5.25) $$Q_C = \sum_{h=1}^{H} \left\{ \sum_{t=1}^{T} \varepsilon_{ht} \right\}^2.$$

As before, the ε's are interpreted as shorthand expressions for the corresponding differences when inserting (2.4).

Partial maximisation of L with respect to ρ and σ^2 , with β fixed, yields (cf. the derivation of eqs. (5.17) and (5.18))

(5.26) $$\hat{\rho} = \frac{1}{T-1} \frac{\sum_{h=1}^{H} \left\{ \left(\sum_{t=1}^{T} \hat{\varepsilon}_{ht} \right)^2 - \sum_{t=1}^{T} \hat{\varepsilon}_{ht}^2 \right\}}{\sum_{h=1}^{H} \sum_{t=1}^{T} \hat{\varepsilon}_{ht}^2},$$

(5.27) $$\hat{\sigma}^2 = \frac{1}{TH} \sum_{h=1}^{H} \sum_{t=1}^{T} \hat{\varepsilon}_{ht}^2.$$

This suggests the following procedure as an approximation to FIML estimation:

(i) Estimate β by (non-linear) OLS, i.e. by minimising Q_A.

(ii) Estimate ρ from the residuals by using eq. (5.26).

(iii) Reestimate β by minimising

$$Q = \frac{1}{\sigma^2(1-\rho)} \left\{ Q_A - \frac{\rho}{1+(T-1)\rho} Q_C \right\},$$

with ρ set equal to the estimate obtained at stage (ii).

Minimisation of Q is equivalent to minimisation of the following sum of squares:

(5.28) $$\sum_{h=1}^{H} \sum_{t=1}^{T} \left\{ \varepsilon_{ht} - \frac{1-R_B}{T} \left(\sum_{s=1}^{T} \varepsilon_{hs} \right) \right\}^2,$$

where

(5.29) $$R_B = \sqrt{\frac{1-\rho}{1+(T-1)\rho}}.$$

5.2.3 Rotation samples with no time specific effects

In this case, the specification is the following:

$n = TH,$

$x = a,$

$w = \varepsilon,$

$\Omega = \sigma^2 \; {}_{R|\omega=0},$

where Ω_R is defined in eq. (3.29). We have

$$\Omega = \sigma^2 \begin{pmatrix} I_G & 0 & 0 & \cdots & 0 & 0 \\ 0 & F & 0 & \cdots & 0 & 0 \\ 0 & 0 & F & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & F & 0 \\ 0 & 0 & 0 & \cdots & 0 & I_G \end{pmatrix}$$

where $G = H/2$, $F = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \otimes I_G$. (The number of F's is T-1.)

Then

$$(5.30) \quad |\Omega| = \sigma^{2TH} |I_G| \; |F|^{T-1} \; |I_G| = \sigma^{2TH} \, |F|^{T-1}$$

$$= \sigma^{2TH} \left| \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \otimes I_G \right|^{T-1} = \sigma^{2TH} \begin{vmatrix} 1 & \rho \\ \rho & 1 \end{vmatrix}^{(T-1)G}$$

$$= \sigma^{2TH} \, (1-\rho^2)^{(T-1)G},$$

and

$$\Omega^{-1} = \frac{1}{\sigma^2} \begin{pmatrix} I_G & 0 & \cdots & 0 & 0 \\ 0 & F^{-1} & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & F^{-1} & 0 \\ 0 & 0 & \cdots & 0 & I_G \end{pmatrix},$$

where

$$F^{-1} = \frac{1}{1-\sigma^2} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix} \otimes I_G .$$

In order to simplify the expression for Q, we write

ε as

$$\varepsilon = \begin{pmatrix} \varepsilon_{1(1)} \\ \varepsilon_{1(2)} \\ \varepsilon_{2(1)} \\ \varepsilon_{2(2)} \\ \vdots \\ \vdots \\ \vdots \\ \varepsilon_{T(1)} \\ \varepsilon_{T(2)} \end{pmatrix}$$

where

$$\varepsilon_{t(1)} = \begin{pmatrix} \varepsilon_{1t} \\ \vdots \\ \vdots \\ \vdots \\ \varepsilon_{Gt} \end{pmatrix}, \quad \varepsilon_{t(2)} = \begin{pmatrix} \varepsilon_{G+1,t} \\ \vdots \\ \vdots \\ \vdots \\ \varepsilon_{Ht} \end{pmatrix} \qquad (t=1, \ldots, T).$$

We then have

$$(5.31) \quad Q = \varepsilon' (\sigma^2 \Omega_{R|\omega = 0})^{-1} \varepsilon = \frac{1}{\sigma^2} \Big[\varepsilon'_{1(1)} \varepsilon_{1(1)} + \sum_{t=2}^{T} (\varepsilon'_{t-1(2)}, \varepsilon'_{t(1)}) F^{-1} \begin{pmatrix} \varepsilon_{t-1(2)} \\ \varepsilon_{t(1)} \end{pmatrix} + \varepsilon'_{T(2)} \varepsilon_{T(2)} \Big]$$

$$= \frac{1}{\sigma^2} \Big\{ \sum_{h=1}^{G} \varepsilon_{h1}^2 + \frac{1}{1-\rho^2} \sum_{t=2}^{T} \Big(\sum_{h=1}^{G} \varepsilon_{h+G,t-1}^2 - 2\rho \sum_{h=1}^{G} \varepsilon_{h+G,t-1} \varepsilon_{ht} + \sum_{h=1}^{G} \varepsilon_{ht}^2 \Big) + \sum_{h=1}^{G} \varepsilon_{h+G,T}^2 \Big\},$$

or alternatively,

$$(5.32) \quad Q = \frac{1}{\sigma^2(1-\rho^2)} \Big\{ Q_A - 2\rho Q_D - \rho^2 Q_E \Big\},$$

where Q_A is defined as in eq. (5.11) and

(5.33) $$Q_D = \sum_{t=2}^{T} \sum_{h=1}^{G} \varepsilon_{h+G,t-1} \varepsilon_{ht},$$

(5.34) $$Q_E = \sum_{h=1}^{G} \varepsilon_{h1}^2 + \sum_{h=1}^{G} \varepsilon_{h+G,T}^2.$$

The log-likelihood function thus takes the form

(5.35) $$L = -\frac{TH}{2}(2\pi) - \frac{1}{2}\left\{TH \log \sigma^2 + (T-1)\, G \log (1-\rho^2)\right\} - \frac{1}{2\sigma^2(1-\rho^2)}\left\{Q_A - 2\rho Q_D - \rho^2 Q_E\right\}.$$

Partial maximisation of L with respect to σ^2, with ρ and β fixed, yields the conditional estimator

(5.36) $$\hat{\sigma}^2 = \frac{1}{TH(1-\rho^2)}\left\{Q_A - 2\rho Q_D' - \rho^2 Q_E\right\},$$

which when inserted into (5.35) gives the concentrated log-likelihood function

(5.37) $$L^* = \text{constant} - \frac{1}{2}\left\{TH \log (Q_A - 2\rho Q_D - \rho^2 Q_E) - \frac{(T+1)H}{2} \log (1-\rho^2)\right\}.$$

We find, after some algebra, that partial maximisation of L^* with respect to ρ, with β fixed, implies solution of the following equation:

(5.38) $$(T-1)Q_E \rho^3 - 2Q_D \rho^2 + \left\{Q_A (T+1) - 2TQ_E\right\}\rho - 2TQ_D = 0.$$

By utilizing the fact that Q_E is approximately equal to Q_A/T, provided T is not too small, an approximate solution to (5.38) can be found. Setting $Q_E = Q_A/T$, (5.38) can be written as

(5.38 a) $$\left(\frac{1}{T}\rho^2+1\right)\left((T-1)Q_A\rho - 2TQ_D\right) = 0,$$

which has

$$(5.39)\quad \rho^{*} = \frac{2T}{T-1}\,\frac{Q_D}{Q_A}$$

as its only real solution. Inserting $\rho=\rho^{*}$ into eq. (5.36), letting $Q_E = Q_A/T$, the variance estimator reduces to

$$(5.40)\quad \hat{\sigma}^2 = \frac{Q_A}{TH}.$$

Thus, when using rotation samples, the FIML estimators of ρ and σ^2 satisfy approximately

$$(5.41)\quad \rho^{*} = \frac{2T}{T-1}\;\frac{\sum_{t=2}^{T}\sum_{h=1}^{G}\hat{\varepsilon}_{h+G,t-1}\,\hat{\varepsilon}_{ht}}{\sum_{t=1}^{T}\sum_{h=1}^{H}\hat{\varepsilon}_{ht}^{2}}\qquad (G = H/2),$$

$$(5.42)\quad \hat{\sigma}^2 = \frac{1}{TH}\sum_{t=1}^{T}\sum_{h=1}^{H}\hat{\varepsilon}_{ht}^{2}.$$

If we omit all observations from households observed only once, i.e. observations for which t=1 & h=1,...,G, and t=T & h=G+1,...,H, it can be shown that the FIML estimators satisfy the following equations exactly:

$$(5.43)\quad \tilde{\rho}^2 = \frac{2Q_D}{Q_A-Q_E} = \frac{2\sum_{t=2}^{T}\sum_{h=1}^{G}\hat{\varepsilon}_{h+G,t-1}\,\hat{\varepsilon}_{ht}}{\sum_{t=2}^{T}\sum_{h=1}^{G}\hat{\varepsilon}_{ht}^{2} + \sum_{t=1}^{T-1}\sum_{h=G+1}^{H}\hat{\varepsilon}_{ht}^{2}},$$

$$(5.44)\quad \tilde{\sigma}^2 = \frac{1}{(T-1)H}\,(Q_A-Q_E) = \sum_{t=2}^{T}\sum_{h=1}^{G}\hat{\varepsilon}_{ht}^{2} + \sum_{t=1}^{T-1}\sum_{h=G+1}^{H}\hat{\varepsilon}_{ht}^{2}.$$

We propose the following approximate estimation procedure:

(i) Estimate β by means of (non-linear) OLS.

(ii) Estimate ρ from the residuals by using eq. (5.41) or (5.43).

(iii) Reestimate β by minimising

$$Q = \frac{1}{\sigma^2(1-\rho^2)}\left\{Q_A - 2\rho\,Q_D - \rho^2\,Q_E\right\},$$

or

$$Q' = \frac{1}{\sigma^2(1-\rho^2)}\left\{Q_A - Q_E - 2\rho Q_D\right\} = Q + \frac{Q_E}{\sigma^2},$$

with ρ set equal to the estimate obtained at stage (ii).

The minimisation of Q' can be carried out practically as follows: We have

$$Q' = \frac{1}{\sigma^2(1-\rho^2)} \sum_{t=2}^{T} \sum_{h=1}^{G} (\varepsilon_{h+G,t-1}, \varepsilon_{ht}) \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{h+G,t-1} \\ \varepsilon_{ht} \end{pmatrix}.$$

Noticing that

$$(5.45) \quad M = \frac{1}{2\sqrt{1-\rho}} \begin{pmatrix} 1+\alpha & -(1-\alpha) \\ -(1-\alpha) & 1+\alpha \end{pmatrix},$$

where

$$(5.46) \quad \alpha = \sqrt{\frac{1-\rho}{1+\rho}},$$

has the property that

$$M'M = \frac{1}{1-\rho^2} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix},$$

then Q' can be written as

$$(5.47) \quad Q' = \frac{1}{\sigma^2(1-\rho)} \sum_{t=2}^{T} \sum_{h=1}^{G} \left\{ \xi_{1ht}^2 + \xi_{2ht}^2 \right\},$$

where $\begin{pmatrix} \xi_{1ht} \\ \xi_{2ht} \end{pmatrix} = \sqrt{1-\rho}\, M \begin{pmatrix} \varepsilon_{h+G,t-1} \\ \varepsilon_{ht} \end{pmatrix}$, or

$$\xi_{1ht} = \frac{1}{2} \left\{ (1+\alpha)\, \varepsilon_{h+G,t-1} - (1-\alpha)\, \varepsilon_{ht} \right\}$$

$$= \varepsilon_{h+G,t-1} - \frac{1-\alpha}{2} (\varepsilon_{h+G,t-1} + \varepsilon_{ht})$$

$$\xi_{2ht} = \frac{1}{2} \left\{ (1+\alpha)\, \varepsilon_{ht} - (1-\alpha)\, \varepsilon_{h+G,t-1} \right\}$$

$$= \varepsilon_{ht} - \frac{1-\alpha}{2} (\varepsilon_{h+G,t-1} + \varepsilon_{ht}).$$

Thus minimisation of Q' is equivalent to minimisation of the sum of squares of the transformed disturbances ξ_{1ht} and ξ_{2ht}.

5.3 Estimation in the multi-equation model.

5.3.1 Disjointed CS/TS data with $\omega_{ij} = \omega$

Let a denote the vector of budget shares ordered first by period, second by individual, third by commodity, i.e.

$$(5.48)\quad a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_T \end{pmatrix},$$

where

$$(5.49)\quad a_t = \begin{pmatrix} a_{1t} \\ \vdots \\ a_{Ht} \end{pmatrix}, \text{ with } a_{ht} = \begin{pmatrix} a_{1ht} \\ \vdots \\ a_{Nht} \end{pmatrix} \qquad \begin{matrix} (h=1,\ldots,H) \\ (t=1,\ldots,T). \end{matrix}$$

With reference to the general scheme (5.1), we have the following situation:

$h = THN$,

$x = a$,

$w = \varepsilon$ (as defined in eq. (4.14)),

$\Omega = \Omega_D \otimes \Sigma$ (as defined in eqs. (3.13) and (4.10)),

assuming $\omega_{ij} = \omega$ for all i and j (cf. eq. (4.34)).

For FIML estimation to be possible, Ω must be nonsingular, i.e. the rank of Σ must be equal to N. This will,however, not be satisfied if the parametric specification of the budget share functions implies satisfaction of (2.6) identically, as is the case with e.g.specifications A-D in section 2.3. Then (2.7) must hold, and the rank of Σ can at most be N-1. In such cases, we delete one commodity from the model, redefine x, w and Ω correspondingly, and replace N by N-1. The likelihood function and, consequently, the FIML estimators are independent of which commodity is deleted.[7] We shall not comment further

7) Cf. Pollak and Wales [15], Appendix A, and Deaton [7], section 4.2.

on this way of reformulating the model in the subsequent sections.

We start, as in section 5.2.1, by deriving expressions for $|\Omega|$ and Q. We have (cf. (5.7))

$$(5.50)\quad |\Omega| = |\Omega_D \otimes \Sigma| = |\Omega_D|^N |\Sigma|^{TH} = (1-\omega)^{T(H-1)N}(1+(H-1)\omega)^{TN}|\Sigma|^{TH}.$$

Moreover,

$$(5.51)\quad Q = \varepsilon'(\Omega_D \otimes \Sigma)^{-1}\varepsilon = \varepsilon'(I_T \otimes A \otimes \Sigma)^{-1} = \varepsilon'(I_T \otimes A^{-1} \otimes \Sigma^{-1})\varepsilon$$

$$= \sum_{t=1}^{T} \varepsilon_t'(A^{-1} \otimes \Sigma^{-1})\varepsilon_t.$$

Using (5.8), this can be written as

$$(5.52)\quad Q = \frac{1}{1-\omega}\sum_{t=1}^{T}\varepsilon_t'\{(I_H - \frac{\omega}{1+(H-1)\omega}E_H) \otimes \Sigma^{-1}\}\varepsilon_t$$

$$= \frac{1}{1-\omega}\left[\sum_{t=1}^{T}\sum_{h=1}^{H}\varepsilon_{ht}'\Sigma^{-1}\varepsilon_{ht} - \frac{\omega}{1+(H-1)\omega}\sum_{t=1}^{T}\sum_{h=1}^{H}\sum_{k=1}^{H}\varepsilon_{ht}'\Sigma^{-1}\varepsilon_{kt}\right].$$

Inserting (5.50) and (5.52) into (5.2), the log-likelihood function takes the form

$$(5.53)\quad L = -\frac{THN}{2}\log(2\pi) - \tfrac{1}{2}\{T(H-1)N\log(1-\omega) + TN\log(1+(H-1)\omega) +$$

$$+ TH\log|\Sigma|\} - \frac{1}{2(1-\omega)}\{Q_A - \frac{\omega}{1+(H-1)\omega}Q_B\},$$

where

$$(5.54)\quad Q_A = \sum_{t=1}^{T}\sum_{h=1}^{H}\varepsilon_{ht}'\Sigma^{-1}\varepsilon_{ht},$$

$$(5.55)\quad Q_B = \sum_{t=1}^{T}\sum_{h=1}^{H}\sum_{k=1}^{H}\varepsilon_{ht}'\Sigma^{-1}\varepsilon_{kt}.$$

The ε's are, as before, shorthand expressions for the corresponding differences when inserting (2.4).

Simultaneous maximisation of L with respect to ω and the elements of β and Σ is even more awkward than in the single-commodity case. The following five-stage procedure may be a practically applicable substitute:

(i) Estimate β by (non-linear) OLS to each demand equation separately (disregarding restrictions between coefficients in different equations). These estimates are consistent (although, of course, not efficient).

(ii) Estimate Σ from the residuals from stage (i):

$$\hat{\Sigma} = (\hat{\sigma}_{ij}) ,$$

where

$$\hat{\sigma}_{ij} = \frac{1}{TH} \sum_{t=1}^{T} \sum_{h=1}^{H} \hat{\varepsilon}_{iht}\hat{\varepsilon}_{jht}.$$

(iii) Calculate estimates of ω from the residuals from stage (i) as in the single equation case. The estimate obtained from the i'th demand equation is (cf. eq. (5.17))

$$\hat{\omega}_i = \frac{1}{H-1} \frac{\sum_{t=1}^{T}\{(\sum_{h=1}^{H}\hat{\varepsilon}_{iht})^2 - \sum_{h=1}^{H} \hat{\varepsilon}_{iht}^2\}}{\sum_{t=1}^{T} \sum_{h=1}^{H} \hat{\varepsilon}_{iht}^2} \qquad (i=1,\ldots,N).$$

The estimates are consistent, by will, of course, generally differ.

(iv) Form a "compromise estimate" of ω:

$$\hat{\omega} = \sum_{i=1}^{N} w_i\hat{\omega}_i,$$

where the w's are weights adding up to unity. We may, for instance, let w_i be equal to 1/N or $(1/\hat{\sigma}_{ii}) / \sum_{j=1}^{N}(1/\hat{\sigma}_{jj})$.

(v) Reestimate β by minimising Q (as given by (5.51)) with Σ and ω set equal to the estimates derived at stages (ii) - (iv)[8]. This stage is similar to a (non-linear) Zellner "Seemingly Unrelated Regressions" method (cf. Zellner [19]).

Stages (iii) and (iv) may be replaced by:

(iii') Estimate ω by solving $\partial L/\partial\omega = 0$, with β and Σ set equal to the estimates obtained at stages (i) and (ii). (This involves solution of an equation of the third degree.)

Stage (v) may be carried out practically in the following way: Since Σ^{-1} is symmetric and positive definite, there exists a N x N matrix Z such that

$$(5.56) \quad Z'Z = \Sigma^{-1}.$$

The Z matrix can be found numerically by computer routines. Then, Q as given by (5.51), can be written as

$$(5.57) \quad Q = \sum_{t=1}^{T} \varepsilon_t'\{(\Phi'\Phi) \otimes (Z'Z)\}\varepsilon_t = \sum_{t=1}^{T} \varepsilon_t'\{(\Phi' \otimes Z')(\Phi \otimes Z)\}\varepsilon_t,$$

paying regard to the factorisation of A^{-1} given by (5.19). Thus, minimisation of Q is equivalent to minimisation of the sum of squares

$$\sum_{t=1}^{T} \nu_t'\nu_t , \text{ where } \nu_t = (\Phi \otimes Z)\varepsilon_t.$$

---

8) By repeating stages (ii)-(v) better approximation might be obtained.

5.3.2 Complete CS/TS data with no time specific effects and with $\rho_{ij} = \rho$

Our specification in this case is the following:

$n = THN,$

$x = a,$

$w = \varepsilon,$

$\Omega = \Omega_{C|\omega=0} \otimes \Sigma = \{(1-\rho)I_T + \rho E_T\} \otimes I_H \otimes \Sigma = B \otimes I_H \otimes \Sigma,$

(cf. eqs. (4.34) and (3.19)).

In complete analogy with the derivation of (5.51)-(5.53), we find

$$(5.58) \quad |\Omega| = (1-\rho)^{H(T-1)N}(1+(T-1)\rho)^{HN}|\Sigma|^{TH},$$

$$(5.59) \quad Q = \frac{1}{1-\rho}\left[\sum_{h=1}^{H}\sum_{t=1}^{T}\varepsilon_{ht}'\Sigma^{-1}\varepsilon_{ht}, - \frac{\rho}{1+(T-1)\rho}\sum_{h=1}^{H}\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{ht}'\Sigma^{-1}\varepsilon_{hs}\right],$$

$$(5.60) \quad L = -\frac{THN}{2}\log(2\pi) - \tfrac{1}{2}\{H(T-1)N\log(1-\rho) + HN\log(1+(T-1)\rho) + + TH\log|\Sigma|\} - \frac{1}{2(1-\rho)}\{Q_A - \frac{\rho}{1+(T-1)\rho}Q_C\},$$

where Q_A is as defined in (5.54), and

$$(5.61) \quad Q_C = \sum_{h=1}^{H}\sum_{t=1}^{T}\sum_{s=1}^{T}\varepsilon_{ht}'\Sigma^{-1}\varepsilon_{hs}.$$

The estimation procedure is

(i)-(ii): Identical with stages (i) and (ii) in section 5.3.1.

(iii) Calculate equation specific estimates of ρ from the OLS residuals from stage (i). The estimate corresponding to the i'th equation is (cf. (5.26))

$$\hat{\rho}_i = \frac{1}{T-1}\frac{\sum_{h=1}^{H}\{(\sum_{t=1}^{T}\hat{\varepsilon}_{iht})^2 - \sum_{t=1}^{T}\hat{\varepsilon}_{iht}^2\}}{\sum_{h=1}^{H}\sum_{t=1}^{T}\hat{\varepsilon}_{iht}^2} \qquad (i=1,\ldots,N).$$

(iv) Form a "compromise estimate" of ρ:

$$\hat{\rho} = \sum_{i=1}^{N} w_i \hat{\rho}_i,$$

where the w's are weights adding up to unity.

(v) Reestimate β by minimising Q (as given by (5.59)) with Σ and ρ set equal to the estimates obtained at stages (ii) - (iv).

5.3.3 Rotation samples with no time specific effects and with $\rho_{ij} = \rho$.

We have

$n = THN,$

$x = a,$

$w = \varepsilon,$

$\Omega = \Omega_{R|\omega=0} \otimes \Sigma,$

where Ω_R is defined in eq. (3.29). Then (compare the derivation of (5.30), (5.32) and (5.35))

$$(5.62) \quad |\Omega| = |\Omega_{R|\omega=0}|^N |\Sigma|^{TH} = (1-\rho^2)^{(T-1)GN} |\Sigma|^{TH},$$

$$(5.63) \quad Q = \frac{1}{1-\rho^2}\Bigg[\sum_{t=1}^{T}\sum_{h=1}^{H} \varepsilon_{ht}' \Sigma^{-1} \varepsilon_{ht} - 2\rho \sum_{t=2}^{T}\sum_{h=1}^{G} \varepsilon_{h+G,t-1}' \Sigma^{-1} \varepsilon_{ht} - \rho^2\Big(\sum_{h=1}^{G} \varepsilon_{h1}' \Sigma^{-1} \varepsilon_{h1} + \sum_{h=1}^{G} \varepsilon_{h+G,T}' \Sigma^{-1} \varepsilon_{h+G,T}\Big)\Bigg],$$

$$(5.64) \quad L = -\frac{THN}{2}\log(2\pi) - \tfrac{1}{2}\{(T-1)GN \log(1-\rho^2) + TH \log|\Sigma|\} - \frac{1}{2(1-\rho^2)}\{Q_A - 2\rho Q_D - \rho^2 Q_E\},$$

where $G = H/2$, Q_A is defined as in (5.54), and

$$(5.65) \quad Q_D = \sum_{t=2}^{T} \sum_{h=1}^{G} \varepsilon'_{h+G,t-1} \Sigma^{-1} \varepsilon_{ht},$$

$$(5.66) \quad Q_E = \sum_{h=1}^{G} \varepsilon'_{h1} \Sigma^{-1} \varepsilon_{h1} + \sum_{h=1}^{G} \varepsilon'_{h+G,T} \Sigma^{-1} \varepsilon_{h+G,T}.$$

Also in this case, simultaneous FIML estimation is numerically rather awkward, and we propose the following approximative procedure:

(i)-(ii) Identical with stages (i) and (ii) in section 5.3.1.

(iii) Calculate estimates of ρ from the OLS residuals from stage (i) as in the single equation case (cf. eqs. (5.38), (5.41), and (5.43)). Let $\hat{\rho}_i$ denote the estimate corresponding to the i'th equation.

(iv) Form the "compromise estimate"

$$\hat{\rho} = \sum_{i=1}^{N} w_i \hat{\rho}_i \qquad (\text{where } \Sigma_i w_i = 1).$$

(v) Reestimate β by minimising Q (as given by (5.63)) or $Q' = \{Q_A - Q_E - 2\rho Q_D\}/(1-\rho^2)$ with Σ and ρ set equal to the estimates obtained at stages (ii)-(iv).

Stages (iii) and (iv) may be replaced by:

(iii') Estimate ρ by solving $\partial L/\partial\rho = 0$, with β and Σ set equal to the estimates obtained at stages (i) and (ii).

A practical way of minimising Q' is the following (cf. the derivation of (5.47)): By means of the factorisation (5.56), defining

$$\nu_{ht} = \begin{pmatrix} \nu_{1ht} \\ \vdots \\ \nu_{Nht} \end{pmatrix} = Z\varepsilon_{ht},$$

Q' can be written as

$$(5.67) \quad Q' = \frac{1}{1-\rho^2} \sum_{t=2}^{T} \sum_{h=1}^{G} \sum_{i=1}^{N} (\nu_{h+G,t-1,i}, \nu_{hti}) \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix} \begin{pmatrix} \nu_{h+G,t-1,i} \\ \nu_{hti} \end{pmatrix}$$

$$= \frac{1}{1-\rho} \sum_{t=2}^{T} \sum_{h=1}^{G} \sum_{i=1}^{N} \Big[\{\nu_{h+G,t-1,i} - \frac{1-\alpha}{2}(\nu_{h+G,t-1,i} + \nu_{hti})\}^2 +$$

$$+ \{\nu_{hti} - \frac{1-\alpha}{2}(\nu_{h+G,t-1,i} + \nu_{hti})\}^2 \Big]$$

where α is defined as in (5.46). Thus minimisation of Q' is equivalent to minimisation of a sum of squares of transformed disturbances, as in the single equation case.

Appendix A

THE DISTURBANCE VARIANCE/COVARIANCE MATRIX CORRESPONDING WITH ROTATION SAMPLES, WHEN ORDERING THE OBSERVATIONS FIRST BY INDIVIDUAL, SECOND BY PERIOD. ADDENDUM TO SECTION 3.5

Disregarding the individuals reporting only once, i.e. the first H/2 individuals reporting in period 1 and the last H/2 individuals reporting in period T (H is supposed to be an even number)[1], the sample considered in section 3.5 can be described as follows: One set of G = H/2 individuals report in periodes 1 and 2, a second set of G individuals report in periods 2 and 3,...,a (T-1)th set of G individuals report in periods T-1 and T.

Consider the 2 x 1 vector

(A.1) $$\varepsilon^{*}_{ht} = \begin{pmatrix} \varepsilon_{h+G,t-1} \\ \varepsilon_{ht} \end{pmatrix} \qquad \begin{pmatrix} h=1,\ldots G=H/2 \\ t=2,\ldots,T \end{pmatrix},$$

where the second element is the disturbance in period t of the individual reporting as no. h in this period, and the first element is the disturbance of the same individual in period t-1. Arrange these individual vectors along one 2G x 1 (H x 1) vector,

(A.2) $$\varepsilon^{*}_{t} = \begin{pmatrix} \varepsilon^{*}_{1t} \\ \varepsilon^{*}_{2t} \\ \vdots \\ \varepsilon^{*}_{Gt} \end{pmatrix} \qquad (t=2,\ldots,T),$$

containing the disturbances of all the individuals reporting in periods t-1 and t, ordered by individual.

We then have

(A.3) $$E(\varepsilon^{*}_{ht}\varepsilon^{*\prime}_{ht}) = \begin{pmatrix} \sigma^2 & \sigma^2_I \\ \sigma^2_I & \sigma^2 \end{pmatrix} = \sigma^2 B_2 \qquad (h=1,\ldots,G),$$

where (cf. eq. (3.19))

---

1) I.e. individuals 1,2,...,H/2, and TH/2+1, TH/2+2,..., (T+1)H/2. (Cf. section 3.2.)

(A.4) $$B_2 = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} = (1-\rho)I_2 + \rho(e_2 e_2'),$$

and (cf.eq.(3.20))

(A.5) $$E(\varepsilon^*_{ht}\varepsilon^{*\prime}_{kt}) = \begin{pmatrix} \sigma^2_T & 0 \\ 0 & \sigma^2_T \end{pmatrix} = \sigma^2 \omega I_2 \qquad (h=1,\dots,G;\ k \neq h).$$

Moreover,

(A.6) $$\begin{cases} E(\varepsilon^*_{ht}\varepsilon^{*\prime}_{k,t-1}) = \begin{pmatrix} 0 & \sigma^2_T \\ 0 & 0 \end{pmatrix} = \sigma^2 \omega C_2 & (h,k=1,\dots,G;\ t=3,\dots,T) \\ E(\varepsilon^*_{ht}\varepsilon^{*\prime}_{k,t+1}) = \begin{pmatrix} 0 & 0 \\ \sigma^2_T & 0 \end{pmatrix} = \sigma^2 \omega C_2' & (h,k=1,\dots,G;\ t=2,\dots,T-1) \end{cases}$$

where (cf.eq. (3.26))

(A.7) $$C_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

and

(A.8) $$E(\varepsilon^*_{ht}\varepsilon^{*\prime}_{ks}) = 0_{2,2} \qquad (h,k=1,\dots,G;\ t=2,\dots,T;\ s=2,\dots,t-2,\ t+2,\dots,T).$$

From (A.2), (A.3), and (A.5) then follows

(A.9) $$E(\varepsilon^*_t\varepsilon^{*\prime}_t) = \sigma^2 \begin{pmatrix} B_2 & \omega I_2 & \cdots & \omega I_2 \\ \omega I_2 & B_2 & \cdots & \omega I_2 \\ \vdots & \vdots & & \vdots \\ \omega I_2 & \omega I_2 & \cdots & B_2 \end{pmatrix} = \sigma^2 P \qquad (t=2,\dots,T),$$

where

(A.10) $$P = I_G \otimes (B_2 - \omega I_2) + (e_G e_G') \otimes (\omega I_2) = (1-\rho-\omega)I_{2G} + \rho I_G \otimes (e_2 e_2') + \omega(e_G e_G') \otimes I_2$$

after inserting from (A.4). Similarly, from (A.2), and (A.6) we have

$$(A.11)\begin{cases} E(\varepsilon_t^{*}\varepsilon_{t-1}^{*\prime}) = \sigma^2(e_G e_G^{\prime}) \otimes \begin{pmatrix} 0 & \omega \\ 0 & 0 \end{pmatrix} = \sigma^2 \omega Q & (t=3,\ldots,T), \\ E(\varepsilon_t^{*}\varepsilon_{t+1}^{*\prime}) = \sigma^2(e_G e_G^{\prime}) \otimes \begin{pmatrix} 0 & 0 \\ \omega & 0 \end{pmatrix} = \sigma^2 \omega Q^{\prime} & (t=2,\ldots,T-1), \end{cases}$$

where

$$(A.12)\quad Q = (e_G e_G^{\prime}) \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = (e_G e_G^{\prime}) \otimes C_2.$$

Moreover, owing to (A.8),

$$(A.13)\quad E(\varepsilon_t^{*}\varepsilon_s^{*\prime}) = 0_{2G,2G} \qquad (t=2,\ldots,T;\ s=2,\ldots,t-2,t+2,\ldots,T)$$

The matrices P and Q have dimension 2G x 2G (H x H).

Finally, form the 2G(T-1)x1 (i.e.H(T-1)x1) vector

$$(A.14)\quad \varepsilon^{*} = \begin{pmatrix} \varepsilon_2^{*} \\ \varepsilon_3^{*} \\ \vdots \\ \varepsilon_T^{*} \end{pmatrix},$$

containing all the disturbances of the rotation samples ordered first by individual then by period. In view of (A.9), (A.11), and (A.13), its variance/covariance matrix takes the form

$$(A.15)\quad E(\varepsilon^{*}\varepsilon^{*\prime}) = \sigma^2\Omega_{R*} = \sigma^2 \begin{pmatrix} P & \omega Q^{\prime} & 0 & \cdots & 0 & 0 \\ \omega Q & P & \omega Q^{\prime} & \cdots & 0 & 0 \\ 0 & \omega Q & P & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & P & \omega Q^{\prime} \\ 0 & 0 & 0 & \cdots & \omega Q & P \end{pmatrix},$$

or

$$(A.16)\quad E(\varepsilon^{*}\varepsilon^{*\prime}) = \sigma^2\Omega_{R*} = \sigma^2\ \{I_{T-1} \otimes P + \omega(D_{T-1} \otimes Q + D_{T-1}^{\prime} \otimes Q^{\prime})\},$$

where D_{T-1} is equal to the matrix D_T after deleting its first column and its last row (cf.eq. (3.28)). By inserting from (A.10) and (A.12), the matrix

(A.16) can be written explicity in terms of σ^2, ρ and ω as

$$\text{(A.17)}\quad E(\varepsilon^{*}\varepsilon^{*\prime}) = \sigma^2\Omega_{R*} = \sigma^2 \{(1-\rho-\omega)I_{2G(T-1)} + \rho I_{G(T-1)} \otimes (e_2 e_2^{\prime}) + \omega I_{T-1} \otimes (e_G e_G^{\prime}) \otimes I_2 + \omega(D_{T-1} \otimes (e_G e_G^{\prime}) \otimes C_2 + D_{T-1}^{\prime} \otimes (e_G e_G^{\prime}) \otimes C_2^{\prime})\}.$$

The properties of Ω_{R*} can be stated as follows:

1) The main diagonal consists of T-1 identical blocks P, each of dimension 2G x 2G. 2) The first sub-diagonal below the main diagonal has T-2 identical blocks ωQ, each of dimension 2G x 2G. The first sub-diagonal above the main diagonal is identical with this, except that all blocks are transposed. 3) The remaining submatrices of Ω_{R*} are zero matrices. Notice in particular that when time specific effects are absent, i.e. when $\omega = 0$, then Ω_{R*} gets a particularly simple form, viz.,

$$\Omega_{R*} = I_{T-1} \otimes P = I_{G(T-1)} \otimes \{(1-\rho)I_2 + \rho(e_2 e_2^{\prime})\} = I_{G(T-1)} \otimes \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

Appendix B

THE DISTURBANCE VARIANCE/COVARIANCE MATRIX CORRESPONDING WITH COMPLETE CS/TS DATA WHEN THE ORDERING BY COMMODITY PRECEDES THE ORDERING BY PERIOD AND INDIVIDUAL. ADDENDUM TO SECTION 4.4

In developing the variance/covariance formulae in ch. 4, we supposed that the ordering by commodity took place after the ordering by period and individual. In this appendix, we shall examine the covariance matrix of the disturbances when reversing this ordering, confining our attention to complete CS/TS data only.

1. Ordering first by commodity, second by period, third by individual: the CPI ordering

Define the Hx1 vector

$$\varepsilon_{i*t} = \begin{pmatrix} \varepsilon_{i1t} \\ \varepsilon_{i2t} \\ \vdots \\ \varepsilon_{iHt} \end{pmatrix}, \tag{B.1}$$

containing the disturbances of the i'th demand function for each of the H individuals in period t. From (4.5b) we have

$$E(\varepsilon_{i*t}\varepsilon_{j*t}') = \sigma_{ij}\begin{pmatrix} 1 & \omega_{ij} & \cdots & \omega_{ij} \\ \omega_{ij} & 1 & \cdots & \omega_{ij} \\ \vdots & \vdots & & \vdots \\ \omega_{ij} & \omega_{ij} & \cdots & 1 \end{pmatrix} \qquad \begin{matrix}(t=1,\ldots,T;\\ i,j=1,\ \ldots,N)\end{matrix} \tag{B.2}$$

$$= \sigma_{ij}\left[(1-\omega_{ij})I_H + \omega_{ij}E_H\right] = \sigma_{ij}A_{ij},$$

where $E_H(=e_He_H')$ is the HxH matrix consisting entirely of ones, and A_{ij} is the matrix in the square bracket. Furthermore,

$$E(\varepsilon_{i*t}\varepsilon_{j*s}') = \sigma_{ij}\rho_{ij}I_H \qquad \begin{matrix}(t=1,\ldots,T;\ s\neq t;\\ i,j = 1,\ldots,N).\end{matrix} \tag{B.3}$$

From (B.2) and (B.3), we easily find that the THx1 vectors

$$\text{(B.4)} \qquad \varepsilon_{i**} = \begin{pmatrix} \varepsilon_{i*1} \\ \varepsilon_{i*2} \\ \vdots \\ \varepsilon_{i*T} \end{pmatrix} \qquad (i=1,\ldots,N)$$

have variance/covariance matrices of the form

$$\text{(B.5)} \qquad E(\varepsilon_{i**}\varepsilon_{j**}') = \sigma_{ij}\begin{pmatrix} A_{ij} & \rho_{ij}I_H & \cdots & \rho_{ij}I_H \\ \rho_{ij}I_H & A_{ij} & \cdots & \rho_{ij}I_H \\ \vdots & \vdots & & \vdots \\ \rho_{ij}I_H & \rho_{ij}I_H & \cdots & A_{ij} \end{pmatrix}$$

$$= \sigma_{ij}\,\{I_T \otimes (A_{ij} - \rho_{ij}I_H) + E_T \otimes (\rho_{ij}I_H)\}$$

$$= \sigma_{ij}\,\{(1-\omega_{ij} - \rho_{ij})I_{TH} + \omega_{ij}I_T \otimes E_H + \rho_{ij}E_T \otimes I_H\}$$

$$= \sigma_{ij}\Omega_{Cij} \qquad (i,j=1,\ldots,N),$$

the last equality defining Ω_{Cij}. Finally, form the NTHx1 vector

$$\text{(B.6)} \qquad \varepsilon_{**} = \begin{pmatrix} \varepsilon_{1**} \\ \vdots \\ \varepsilon_{N**} \end{pmatrix},$$

consisting of all the disturbances ordered first by commodity, second by period, third by individual. Its variance/covariance matrix is

$$\text{(B.7)} \qquad E(\varepsilon_{**}\varepsilon_{**}') = \begin{pmatrix} \sigma_{11}\Omega_{C11} & \cdots & \sigma_{1N}\Omega_{C1N} \\ \vdots & & \vdots \\ \sigma_{N1}\Omega_{CN1} & \cdots & \sigma_{NN}\Omega_{CNN} \end{pmatrix}$$

Generally, this matrix cannot be written in Kronecker product notation. In the particular case where $\rho_{ij} = \rho$, and $\omega_{ij} = \omega$ for all i and j, however, (cf. section 4.6) then $\Omega_{Cij} = \Omega_C$ (as defined in eq. (3.16)), and (B.7) degenerates to

$$\text{(B.8)} \quad E(\varepsilon_{**}\varepsilon_{**}') = \Sigma \otimes \Omega_C = \Sigma \otimes \{(1-\omega-\rho) I_{TH} + \omega I_T \otimes E_H + \rho E_T \otimes I_H\}.$$

2. Ordering first by commodity, second by individual, third by period: the CIP ordering

The reversing of the ordering of individuals and periods considered in section 1 is straightforward. Define the Tx1 vector

$$\text{(B.9)} \quad \varepsilon_{ih*} = \begin{pmatrix} \varepsilon_{ih1} \\ \varepsilon_{ih2} \\ \vdots \\ \varepsilon_{ihT} \end{pmatrix},$$

its elements being the disturbances of the i'th demand function for the h'th individual in each of the T periods. From (4.5b) follows

$$\text{(B.10)} \quad E(\varepsilon_{ih*}\varepsilon_{jh*}') = \sigma_{ij}\left[(1-\rho_{ij}) I_T + \rho_{ij} E_T\right] = \sigma_{ij} B_{ij} \qquad (h=1,\dots,H; i,j=1,\dots,N),$$

where B_{ij} is the matrix in the square bracket. Moreover,

$$\text{(B.11)} \quad E(\varepsilon_{ih*}\varepsilon_{jk*}') = \sigma_{ij}\omega_{ij} I_T \qquad (h=1,\dots,H;\ k \neq h;\ i,j=1,\dots,N).$$

From (B.10), and (B.11) we find that the HTx1 vectors

$$\text{(B.12)} \quad \tilde{\varepsilon}_{i**} = \begin{pmatrix} \varepsilon_{i1*} \\ \varepsilon_{i2*} \\ \vdots \\ \varepsilon_{iH*} \end{pmatrix} \qquad (i=1,\dots,N)$$

have variance/covariance matrices given by

$$\text{(B.13)}\quad E(\tilde{\varepsilon}_{i**}\tilde{\varepsilon}_{j**}') = \sigma_{ij}\{I_H \otimes (B_{ij} - \omega_{ij} I_T) + E_H \otimes (\omega_{ij} I_T)\}$$

$$= \sigma_{ij}\{(1-\rho_{ij} - \omega_{ij}) I_{HT} + \rho_{ij} I_H \otimes E_T + \omega_{ij} E_H \otimes I_T\}$$

$$\stackrel{1}{=} \sigma_{ij}\Omega_{C*ij} \qquad (i,j=1,\ldots,N),$$

the last equality defining Ω_{C*ij}. Thus, the NHTx1 vector

$$\text{(B.14)}\quad \tilde{\varepsilon}_{**} = \begin{pmatrix} \tilde{\varepsilon}_{1**} \\ \vdots \\ \tilde{\varepsilon}_{N**} \end{pmatrix}$$

containing all the disturbances ordered first by commodity, second by individual, third by period, has the following variance/covariance matrix:

$$\text{(B.15)}\quad E(\tilde{\varepsilon}_{**}\tilde{\varepsilon}_{**}') = \begin{pmatrix} \sigma_{11}\Omega_{C*11} & \cdots & \sigma_{1N}\Omega_{C*1N} \\ \vdots & & \vdots \\ \sigma_{N1}\Omega_{C*N1} & \cdots & \sigma_{NN}\Omega_{C*NN} \end{pmatrix}.$$

An ordering of disturbances identical with the CIP ordering considered above is used in a recent article by Avery [1]. Eq. (B.15) corresponds with eq. (2.7) in Avery's article (our matrix Ω_{C*ij} corresponding with Avery's matrix Σ_{ij}).

Generally, the right hand side of (B.15) cannot be written in Kronecker product notation. In the particular case where $\rho_{ij} = \rho$, and $\omega_{ij} = \omega$ for all i and j, (cf. section 4.6), we have $\Omega_{C*ij} = \Omega_{C*}$ (as defined in eq. (3.23)), and (B.15) degenerates to

$$\text{(B.16)}\quad E(\tilde{\varepsilon}_{**}\tilde{\varepsilon}_{**}') = \Sigma \otimes \Omega_{C*} = \Sigma \otimes \{(1-\rho-\omega) I_{HT} + \rho I_H \otimes E_T + \omega E_H \otimes I_T\}.$$

Appendix C

SOME USEFUL PROPERTIES OF KRONECKER PRODUCTS

The aim of this appendix is to refer, partly without proofs, some properties of matrices expressed as Kronecker products, which have been used in developing the estimation methods in chapter 5.

Let $U = (u_{ij})$ and $V = (v_{ij})$ be non-singular matrices of dimension M x M and G x G respectively. By definition we have

$$(C.1) \qquad C = U \otimes V = \begin{pmatrix} u_{11}V & \cdots & u_{1M}V \\ \vdots & & \vdots \\ u_{M1}V & \cdots & u_{MM}V \end{pmatrix}$$

The dimension of C is MG x MG.

(i) Matrix inversion

$$(C.2) \qquad \underline{C^{-1} = (U \otimes V)^{-1} = U^{-1} \otimes V^{-1}}$$

(ii) Determinant values

$$(C.3) \qquad \underline{|C| = |U \otimes V| = |U|^G \, |V|^M}$$

(iii) Ranks

$$(C.4) \qquad \underline{\text{rank } (C) = \text{rank } (U \otimes V) = \text{rank } (U) \cdot \text{rank } (V)}$$

(= MG, since U and V are non-singular).

(Proofs of (C.2) - (C.4) are found in e.g. Theil [17], pp. 304-306.)

(iv) Quadratic forms

Let

$$x_r = \begin{pmatrix} x_{1r} \\ \vdots \\ x_{Mr} \end{pmatrix} \qquad (r = 1, \ldots, G)$$

be G vectors of dimension M x 1. Form the GM x 1 vector

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_G \end{pmatrix}.$$

Similarly, define the M vectors of dimension G x 1

$$\tilde{x}_i = \begin{pmatrix} x_{i1} \\ \vdots \\ x_{iG} \end{pmatrix} \qquad (i = 1, \ldots, M),$$

and the composite MG x 1 vector

$$\tilde{x} = \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_M \end{pmatrix}.$$

Then,

$$\text{(C.5)} \qquad Q = x'(V \otimes U)x = \tilde{x}'(U \otimes V)\tilde{x}$$

$$= \sum_{r=1}^{G} \sum_{s=1}^{G} \sum_{j=1}^{M} \sum_{i=1}^{M} x_{ir} u_{ij} v_{rs} x_{js}$$

Proof of (C.5):

From the definitions of U, V and x, we have

(*) $$Q = (x_1{}' \; x_2{}' \cdots x_G{}') \begin{pmatrix} v_{11}U & \cdots & v_{1G}U \\ \vdots & & \vdots \\ v_{G1}U & \cdots & v_{GG}U \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_G \end{pmatrix}$$

$$= (\sum_r v_{r1}x_r{}'U, \; \sum_r v_{r2}x_r{}'U \cdots \sum_r v_{rG}x_r{}'U) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_G \end{pmatrix}$$

$$= \sum_{r=1}^{G} \sum_{s=1}^{G} v_{rs}x_r{}'Ux_s .$$

Expanding the bilinear form $x_r{}'Ux_s$ we get

(**) $$x_r{}'Ux_s = \sum_{i=1}^{M} \sum_{j=1}^{M} x_{ir}u_{ij}x_{js} \qquad (r,s=1,\ldots,G)$$

In a similar way, we find

(***) $$\tilde{x}'(U \otimes V)\tilde{x} = \sum_{j=1}^{M} \sum_{i=1}^{M} u_{ij}\tilde{x}_i{}'V\tilde{x}_j$$

and

(****) $$\tilde{x}_i{}'V\tilde{x}_j = \sum_{r=1}^{G} \sum_{s=1}^{G} x_{ir}v_{rs}x_{js} \qquad (i,j=1,\ldots,M)$$

Eq. (C.5) follows by inserting (**) into (*) and (****) into (***). Q.E.D.

Eq. (C.5) may be readily generalized to bilinear forms.

Examples:

1) $V = I_G$; i.e. $v_{ss}=1$, $v_{rs}=0$ for $r \neq s$.

(C.6) $$x'(I_G \otimes U)x = \sum_{s=1}^{G} \sum_{j=1}^{M} \sum_{i=1}^{M} x_{is} u_{ij} x_{js}.$$

2) $V = I_G$, $U = E_M$; i.e. $v_{ss}=1$, $v_{rs}= 0$ for $r \neq s$; $u_{ij}=1$ for all i,j.

(C.7) $$x'(I_G \otimes E_M)x = \sum_{s=1}^{G} \sum_{j=1}^{M} \sum_{i=1}^{M} x_{is} x_{js} = \sum_{s=1}^{G} \{\sum_{i=1}^{M} x_{is}\}^2.$$

3) $U = I_M$; i.e. $u_{ii}=1$, $u_{ij}=0$ for $j \neq i$.

(C.8) $$x'(V \otimes I_M)x = \sum_{i=1}^{M} \sum_{r=1}^{G} \sum_{s=1}^{G} x_{ir} v_{rs} x_{is}.$$

4) $U = I_M$, $V = E_G$; i.e. $u_{ii}=1$, $u_{ij}=0$ for $j \neq i$; $v_{rs}=1$ for all r,s.

(C.9) $$x'(E_G \otimes I_M)x = \sum_{i=1}^{M} \sum_{r=1}^{G} \sum_{s=1}^{G} x_{ir} x_{is} = \sum_{i=1}^{M} \{\sum_{r=1}^{G} x_{ir}\}^2 .$$

REFERENCES

[1] Avery, R.B.: Error Components and Seemingly Unrelated Regressions. Econometrica, vol. 45 (1977), pp. 199-209.

[2] Balestra, P.: The Demand for Natural Gas in the United States. (Amsterdam: North-Holland Publishing Company, 1967.)

[3] Balestra, P., and Nerlove, M.: Pooling Cross Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas. Econometrica, vol. 34 (1966), pp. 585-612.

[4] Biørn, E.: Two Notes on the Stochastic Specification of a Complete Set of Consumer Demand Functions. Working Papers from the Central Bureau of Statistics of Norway, IO 77/17 (1977).

[5] Carlevaro, F.: A Generalization of the Linear Expenditure System. In L. Solari, and J.-N. Du Pasquier (eds.): Private and Enlarged Consumption. (Amsterdam: North-Holland Publishing Company, 1976.)

[6] Chetty, V.K.: Pooling of Time Series and Cross Section Data. Econometrica, vol. 36 (1968), pp. 279-290.

[7] Deaton, A.: Models and Projections of Demand in Post-War Britain. (London: Chapman and Hall, 1975.)

[8] Johansen, L.: On the Relationships between Some Systems of Demand Functions. Liiketaloudellinen Aikakauskirja vol. 1 (1969), pp. 30-41.

[9] Kuh, E.: The Validity of Cross-Sectionally Estimated Behavior Equations in Time Series Applications. Econometrica, vol. 27 (1959), pp. 197-214.

[10] Lybeck, J.A.: The Allocation of Consumer Expenditures in a Swedish Macroeconomic Model. European Economic Review, vol. 8 (1976), pp. 371-396.

[11] Maddala, G.S.: The Use of Variance Components Models in Pooling Cross Section and Time Series Data. Econometrica, vol. 39 (1971), pp. 341-358.

[12] Nasse, P.: Un Système Complet de Fonctions de Demande: Les Équations de Fourgeand et Nataf. Econometrica, vol. 41 (1973), pp. 1137-1158.

[13] Nerlove, M.: Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross Sections. Econometrica, vol. 39 (1971), pp. 359-382.

[14] Nerlove, M.: A Note on Error Components Models. Econometrica, vol. 39 (1971), pp. 383-396.

[15] Pollak, R.A., and Wales, T.J.: Estimation of the Linear Expenditure System. Econometrica, vol 37 (1969), pp. 611-623.

[16] Rothenberg T.J., and Leenders, C.T.: Efficient Estimation of Simultaneous Equation Systems. Econometrica, vol. 32 (1964), pp. 57-76.

[17] Theil, H.: Principles of Econometries. (New York: John Wiley & Sons, Inc., 1971.)

[18] Wallace, T.D., and Hussain, A.: The Use of Error Components in Combining Cross Section with Time Series Data. Econometrica, vol. 37 (1969), pp. 55-72.

[19] Zellner, A.: An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. Journal of the American Statistical Association, vol. 57 (1962), pp. 348-368.