A list of Working Papers on the last pages

No. 131, 1984

CONRAD - A Maximum Likelihood Program for Estimation of Simultaneous Equations Models that are Non-Linear in the Parameters

by

Leif Jansson and Erik Mellander

This is a preliminary paper. It is intended for private circulation, and should not be quoted or referred to in publications without permission of the authors. Comments are welcome.

September, 1984

CONTENTS

1 INTRODUCTION

2 ESTIMATION AND TESTING

2.1 Formal description of the model
2.2 The mapping from θ to A
2.2.1 Inequality constraints
2.2.2 General form of the coefficients
2.3 Identification
2.4 Hypothesis testing
Notes

3 PROGRAM CHARACTERISTICS

3.1 General features
3.2 Optimization method
3.3 Computation of Var($\hat{\theta}$)
3.4 Compatability with other computers
Notes

4 INPUT INSTRUCTIONS

4.1 DATAM.DAT
4.2 PARIC.DAT
4.3 GICOF.DAT

5 COMMENTS ON OUTPUT

5.1 RESF1.DAT
5.2 TERMINAL OUTPUT
5.3 RESF2.DAT
5.4 RESF3.DAT
5.5 RESF4.DAT
Notes

Appendix A THE CONCENTRATED LOG-LIKELIHOOD FUNCTION AND ITS FIRST ORDER DERIVATIVES

Appendix B DIMENSION STATEMENTS

Appendix C AN EXAMPLE

REFERENCES

1 INTRODUCTION

CONRAD (Constrained Regression with Analytical Derivatives) is a FIML[1] program for estimation of simultaneous systems of equations.

While requiring the systems to be linear in the variables, the program can handle quite general non-linearities in the parameters. The stochastic specification incorporates both contemporaneous and intertemporal correlations between the error terms, the latter in the form of a first order vector autoregression.

Owing to one of the characteristic features of CONRAD, namely the analytical determination of the gradient of the log-likelihood function, systems of considerable size can be estimated without unduly long execution times. In contrast to most other programs of comparable type, the analytical form of the gradient vector is automatically solved by the program and used to evaluate the first order derivatives. The user thus does not have to concern himself with the maximization of the likelihood function.

CONRAD can be used without prior knowledge of programming. The program offers great flexibility with respect to input and model specifications. All input may be supplied in free format. Variab-

---

[1] Full Information Maximum Likelihood; for the principle of maximum likelihood, see Silvey (1975). Hendry (1976) discusses the relationship between FIML and other simultaneous equations estimates.

les can be chosen freely from a given data matrix and, regarding time series data, the estimation can be based on any coherent subperiod.

By means of an optional line width, printouts can be obtained both on small terminals and on full-scale printers. Not only estimation results can be output; if desired, a thorough documentation of input data and specifications will also be supplied. This facility has been implemented in such a way that it can be used as a means to check the proper arrangement of the input, before the actual estimation. In addition, a number of checks are performed by the program.

Users not interested in technical issues can confine their reading to sections 2, 4 and 5. The last two ones are preferably studied together with the example in Appendix C.

Concerning notation, matrices and vectors have been denoted by boldface type, using capitals for the matrices and small letters for the vectors.

2 ESTIMATION AND TESTING

2.1 Formal description of the model

The simultaneous equations model is assumed to be linear in the variables. It can be written in matrix notation in the following way

$$\mathbf{A}(\theta)\mathbf{x}_t = \mathbf{B}\mathbf{y}_t + \mathbf{C}\mathbf{z}_t = \mathbf{u}_t \qquad t = 1,2,\ldots,T \qquad (1)$$

where the vector $\mathbf{x}_t$ and the matrix $\mathbf{A}$ are partitioned according to

$$\mathbf{x}_t = \begin{pmatrix} \mathbf{y}_t \\ \mathbf{z}_t \end{pmatrix} \quad \text{and} \quad \mathbf{A}(\theta) = (\mathbf{B}:\mathbf{C}).$$

The n×1 and m×1 vectors $\mathbf{y}_t$ and $\mathbf{z}_t$ contain observations at time t on the endogenous and predetermined variables, respectively. Lagged values of the endogenous variables may be among the explanatory variable, i.e. it is possible that $y_{t-i} \in \{\mathbf{z}_t\}$ for $i > 0$.

As indicated by the expression $\mathbf{A}(\boldsymbol{\theta})$, the elements of the coefficient matrix $\mathbf{A}$ are regarded as functions of a set of k unrestricted parameters $\boldsymbol{\theta}' = (\theta_1, \theta_2, \ldots, \theta_k)$. The set of all T realizations of the relationship in (1) can be compiled in the matrix equation

$$\mathbf{A}(\theta)\mathbf{X}' = \mathbf{B}\mathbf{Y}' + \mathbf{C}\mathbf{Z}' = \mathbf{U}'. \qquad (1a)$$

The most general stochastic specification allows for both contemporaneous and intertemporal correlation between the error terms according to:

$$\text{I:} \qquad \mathbf{u}_t = \mathbf{H}\mathbf{u}_{t-1} + \mathbf{e}_t \qquad \text{where} \tag{2}$$

$$\mathbf{e}_t \sim N(\mathbf{0}, \mathbf{\Sigma}) \text{ for all } t,^{1} \tag{3a}$$

such that

$$E(\mathbf{e}_s \mathbf{e}_t') = \delta_{st}\mathbf{\Sigma} \qquad \delta_{st} = \begin{cases} 1 \text{ if } s=t \\ 0 \text{ if } s\neq t \end{cases} \tag{3b}$$

Thus, the errors are assumed to be generated by a first order vector autoregressive process as proposed in Hendry (1971). The $n \times n$ matrix $\mathbf{H}$ is nondiagonal, making every element of $\mathbf{u}_t$ funtionally dependent upon all the elements of $\mathbf{u}_{t-1}$.

Alternatively $\mathbf{H}$ can be set to the zero matrix, implying the more restrictive assumption that the stochastic process is independent of time.

In addition to I, the following assumptions are made:

II: $\mathbf{B}$ is nonsingular; $\det(\mathbf{B}) \neq 0$ where "det" denotes determinant,

III: $\mathbf{X}$ has full column rank, i.e. $r(\mathbf{X}) = n+m$,

IV: $\mathbf{\Sigma}$ is unconstrained (except for being positive definite, i.e. $\mathbf{v}'\mathbf{\Sigma}\mathbf{v} > 0$ for all vectors $\mathbf{v} \neq \mathbf{0}$),

V: $T > n+m$,

VI: $\mathbf{H}$ is unconstrained.

Assumption II is required to ensure that the endogenous vector $\mathbf{y}_t$ be unique for every predetermined vector $\mathbf{z}_t$ and disturbance vector $\mathbf{e}_t$. Condition IV

implies that (1) does not contain any identities, so it is assumed that these have been substituted out. Finally, III and V together make up a sufficient condition for the estimator to be well defined. For a more elaborate discussion about sample size requirements, see Brown (1981).

Although **H** is unconstrained, only those matrices that have all their eigenvalues within the unit circle can be considered meaningful. The eigenvalues λ_j, $j=1,2,\ldots,n$, which are the n roots of the n:th degree equation

$$\det(\mathbf{H}-\lambda\mathbf{I}) = 0 \qquad (4)$$

where

I = the identity matrix of order n

λ = a+bi; a scalar with real and imaginary parts a and b, respectively (the latter oppsibly zero)

thus must have the property that $(a^2+b^2)^{1/2} < 1$. (In the single equation case this simplifies to the requirement that the sole element of **H** be less than one in absolute value.)

If this condition is fulfilled the process generating the errors can be taken to be stationary, which is necessary for the elements of $\mathbf{\Sigma}$ to be finite and independent of t. When some of the eigenvalues lie outside the unit circle the error variances and covariances will instead increase without limit over time. As the program calculates the eigenvalues of the estimated **H** matrix it is possible to check which one of these cases that is prevalent.

2.2 The mapping from Θ to A

The most general mapping from Θ to an element of **A** can be considered to be the outcome of two subsequent operations.

In the first step one may apply transformation functions to the θ_i:s, whose ranges are constrained to (subsets of) the non-negative or non-positive real numbers, i.e.

$$h_1(\theta_i) \geq \gamma_i \geq 0 \quad \text{and} \quad h_2(\theta_i) \leq \upsilon_i \leq 0$$

where γ_i and υ_i are constants chosen by the user. This step is treated in section 2.2.1.

In the second step, the a_{ij}:s can be formed as a ratio of two second-degree polynomials in the transformed (or untransformed) free parameters[2], as shown in section 2.2.2.

2.2.1 Inequality constraints

Often some parameters are restricted with respect to sign and/or size, by the theory underlying the model. As long as the restrictions refer to individual parameters only, they can be tested by means of one-sided t-tests.

In order to make it possible to impose inequality constraints on several parameters simultaneously,[3] the following two hyperbolic transformations have been implemented in CONRAD

$$h_1(\theta_i) = \begin{cases} (\theta_i^6 + \gamma_i^6)^{1/6} & \text{for } \gamma_i \geq 0 \\ (\theta_i^2 + 0.05^2)^{1/2} - 0.05 & \text{for } \gamma_i = 0 \end{cases} \quad (5a)$$

and

$$h_2(\theta_i) = -\begin{cases} (\theta_i^6 + \upsilon_i^6)^{1/6} & \text{for } \upsilon_i \leq 0 \\ (\theta_i^2 + 0.05^2)^{1/2} - 0.05 & \text{for } \upsilon_i = 0 \end{cases} \quad (5b)$$

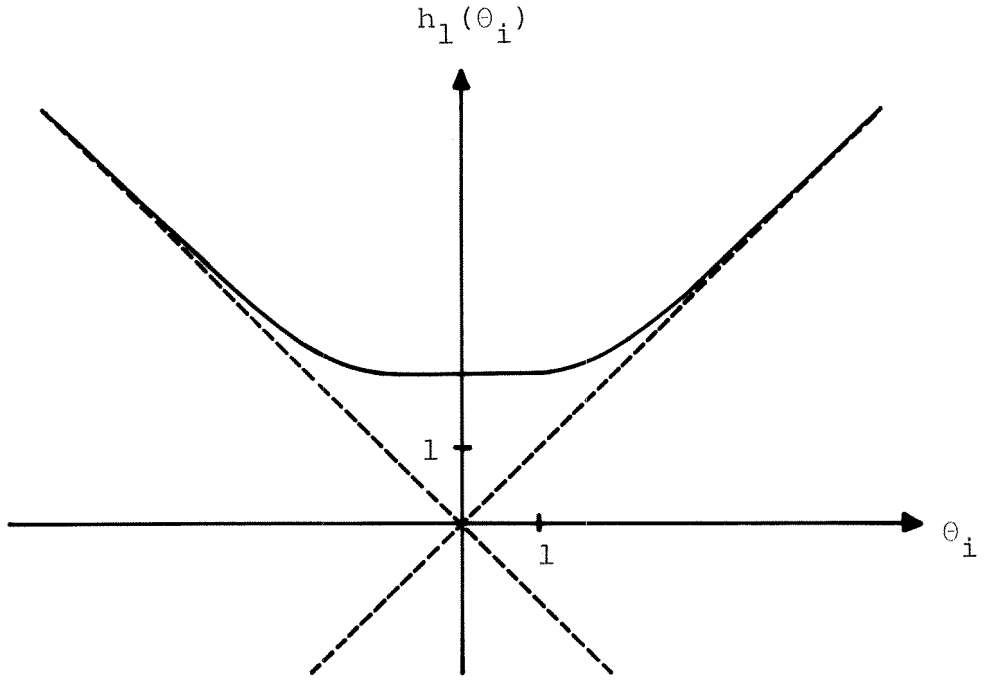
The number 0.05 has been used to ensure differentiability when the limit is equal to zero. As can readily be seen, the limiting values of the transformations when this constant goes to zero are $|\theta_i|$ and $-|\theta_i|$, respectively.

The user can thus impose inequality constraints on the parameters of his/her model by defining them as $h_1(\theta_i)$ or $h_2(\theta_i)$.[4]

An example of (5a) is given in Fig. 1, with the lower limit set equal to 2. For illustrative purposes, 4 has been used for the exponents instead of 6. (By increasing the integer, which must be even, the hyperbola can be made to pass arbitrarily close to its asymptotes, shown by the dotted lines.)

Instead of the hyperbolic transformations quadratic ones like, e.g., $g(\theta_i) = \theta_i^2 + \gamma_i$ could have been used. The advantage of the former is that it becomes more "linear" the further away from the vertical axis one moves. So it is only in the neighborhood of the lower (upper) limit (γ_i and υ_i

Figure 1 **The transformation function h_1 with $\gamma = 2$**



respectively) that the restrictions are binding and any severe nonlinearities are imposed. When the optimal value of θ_i is far from γ_i (or υ_i) the search for the optimal set of parameters is thus little affected. This would not be the case if a quadratic transformation had been used.

2.2.2 General form of the coefficients

To simplify the notation we write the $n\times(n+m)$ coefficient matrix **A** in the form of a column vector according to

$$\text{vec } \mathbf{A}(\boldsymbol{\theta}) = \begin{matrix} \mathbf{a}_{.1} \\ \mathbf{a}_{.2} \\ \vdots \\ \mathbf{a}_{.n+m} \end{matrix} \tag{6}$$

where $\mathbf{a}_{.j}$ = the j:th column of $\mathbf{A}(\boldsymbol{\theta})$,

i.e. the columns of $\mathbf{A}(\boldsymbol{\theta})$ are stacked on top of each other. The element a_{ij} of $\mathbf{A}(\boldsymbol{\theta})$ will be in position $g = n(j-1)+i$ of vec $\mathbf{A}(\boldsymbol{\theta})$.

Further, we introduce an identity transformation of θ_i:

$$h_0(\theta_i) = \theta_i \tag{7}$$

An element of vec $\mathbf{A}(\boldsymbol{\theta})$ can then, in the most general case, be expressed in the following way:

$$a_g = \frac{n_g}{d_g}, \tag{8}$$

where

$$\begin{aligned} n_g = r_{go} &+ \sum_{i=1}^{k} r_{gi} h_{\ell_i}(\theta_i) \\ &+ \sum_{i=1}^{k} \sum_{j \geq i}^{k} \rho_{gij} h_{\ell_i}(\theta_i) h_{\ell_j}(\theta_j) \end{aligned} \tag{8a}$$

and

$$d_g = s_{go} + \sum_{i=1}^{k} s_{gi} h_{\ell_i}(\theta_i) + \sum_{i=1}^{k} \sum_{j \geq i}^{k} \sigma_{gij} h_{\ell_i}(\theta_i) h_{\ell_j}(\theta_j) \qquad (8b)$$

$g = 1,2,\ldots,n(n+m)$

$\ell = 0,1,2$

where r, ρ, s and σ denote constants.

As can be seen, the non-linear parts of (8a) and (8b) - the doublesums - contain all the $k(k+1)/2$ possible different product of the free parameters.

2.3 Identification

Whereas for linear models criteria exist for global identification,[5] the corresponding criteria for nonlinear models assure only local identification, i.e. identification in the neighbourhood of the true parameter vector.[6] However, if the nonlinearities are confined to the parameters, as is the case here, the condition for local identification turns out to be quite simple.

Given the assumptions in sec. 2.1, the model is locally identified if the $n(n+m) \times k$ matrix of first partial derivatives

$$\frac{\partial \text{vec } \mathbf{A}(\Theta)}{\partial \Theta} = (\delta_{gi}) \qquad (9)$$

$$= \left(\frac{\partial a_g}{\partial \theta_i}\right) \qquad \begin{matrix} g=1,2,\ldots,n(n+m) \\ i=1,2,\ldots,k \end{matrix}$$

has full column rank. That is, the number of independent columns should be equal to the number of elements in Θ.

The involved derivatives can be found in Appendix A. A simple application of the condition for identification is provided by the example in Appendix C.

2.4 Hypothesis testing

It is assumed to be well known how the significance of single parameters can be tested by means of the t-distribution.

In this section a more general test will be described, namely the likelihood ratio (LR) test. This test which, i.a., can be used in tests relating to several parameters, is shortly described in Maddala (1977, pp. 43-44).

Let H_0 be a hypothesis concerning the parametric structure of the model, which is more restrictive than an alternative hypothesis H_1. Denote the corresponding log-likelihood values by L_0 and L_1, respectively. Then $L_1 > L_0$ and minus twice the logarithm of the likelihood ratio becomes $-2(L_0 - L_1)$, which is asymptotically distributed as a chi-square under H_0. The number of degrees of freedom is equal to the difference in the number of unrestricted parameters. Formally

$$-2(L_0 - L_1) \underset{A}{\sim} \chi^2(p_1 - p_0), \tag{10}$$

where

p_1 = the number of unrestricted parameters in the least restrictive model

p_0 = the number of unrestricted parameters in the more restrictive model,

and $\tilde{A}$ denotes "asymptotically distributed as". The log-likelihood values can easily be obtained according to[7]

$$L = -F + k \tag{11}$$

where

F = the function value calculated by the program

$k = -\frac{1}{2}nT[\ell n(2\Pi)+1]$

n = number of equations

T = number of observations.

As the constant k disappears in the subtraction of L_1 from L_0 the test can be performed directly with the F values.

Inequality constraints can in principle also be tested with the LR test. However, in the presence of such restrictions the asymptotic distribution of the test statistic is not that of a single chi-square variable but rather a weighted average of several chi-square distributions, which makes the practical application of the test very difficult in most cases. See further Gouriéroux, Holly and Monfort (1982) and also Judge, Griffiths, Hill and Lee (1980, Ch. 3).

Another use of the LR test is to check the compatibility of the sample information with different

stochastic specifications. Denoting the more restrictive assumption of no autocorrelation with I' the following relationship can be exploited

$$-2(L_{I'}-L_{I}) \underset{A}{\sim} \chi^2(n^2), \qquad (12)$$

where n is the number of equations. Note that L_I and $L_{I'}$ are assumed to be based on the same number of observations.

Although the LR test is easy to use and has several attractive properties (cf. Maddala op. cit.) it has the disadvantage that it is strictly applicable only to large samples. Thus, when T is low compared to the number of degrees of freedom, the outcomes of the test should be interpreted with care.

Finally, it should be mentioned that there are other general tests beside the LR test, e.g. the Wald and Lagrange Multiplier (LM) tests. A relatively non-technical discussion of the Wald, LR and LM tests can be found in Engle (1984). Although these tests are all asymptotically equivalent, they may yield conflicting results in finite samples. In general, the Wald test is more conservative than the LR test, while the LM test is the one most likely to reject the null hypothesis.[8]

NOTES

[1] The normality assumption is not crucial. For all specifications available in CONRAD the estimate $\hat{\theta}$ is consistent, as long as the expectation of **e** is zero and the variance is in accordance with (3b), cf Hausman (1983).

[2] The second order polynomial mapping is considered in Jansson and Mellander (1983).

[3] Testing procedures for this type of constraints tend in general to be very complicated, however. See further Section 2.4.

[4] Notice the distinction between the user's model and the statistical model. The parameters of the latter are, of course, still the θ_i:s.

[5] Cf. Fischer (1966).

[6] See Rothenberg (1971).

[7] The corresponding formula in Appendix A is (a6). For simplicity, the superindex "*" used there, has been omitted here.

[8] See Berndt and Savin (1977) and Breusch (1979).

3 PROGRAM CHARACTERISTICS

3.1 General features

CONRAD is written in FORTRAN-10, a FORTRAN variant available on Digital Equipment's DEC-10 computers. Aside from some machine-specific characteristics (cf. sec. 3.4), the version used here[1] is close to FORTRAN-66. The program operates on the DEC-10 computer at the Stockholm University Computing Center (QZ).

Including all subroutines the source program has a total length of about 1,600 rows, giving a rough indication of its size.

Since it has been found to considerably increase accuracy, without unduly increasing execution times, all computations are performed in double precison.[2] To keep the price for this higher accuracy as low as possible, effort has been taken to economize on storage space. Consequently, many arrays are used for multiple purposes. Comments have, however, been inserted in the source program where this might create confusion.

To lessen the need for frequent changes of the DIMENSION and COMMON statements, almost all matrices are stored in vector form, i.e. according to (6) in sec. 2.2.2. When changes none the less have to be made, they are faciliated by Appendix B which contains all relevant information on memory allocation requirements.

The use of input/output media is confined to disk files. Further, the program assumes all input to be in alphanumeric form and provides all output in the same format.

3.2 Optimization method

Since restrictions on the elements of **A** are handled by treating them as functions of the unconstrained vector $\boldsymbol{\theta}$, the maximizition of the log-likelihood function is an unconstrained optimization problem. The function to be maximized is the "concentrated" log-likelihood function L* (cf. Appendix A),whose only argument is the vector $\boldsymbol{\theta}$. In practice the program instead follows the equivalent route of minimizing -L*, or, more specifically, F=-L*+k* where k* is a constant.

The minimization is carried out by means of a quasi-Newton[3] routine, VA09AD, from the Harwell Subroutine Library,[4] which is based on an algorithm developed by Fletcher (1970). Although similar to the more well-known Davidon-Fletcher-Powell (DFP) algorithm [Fletcher and Powell (1963)], Fletcher's method is considerbly more efficient since it has almost eliminated the linear search subproblem inherent in the DFP.[5]

Function values and first order derivatives, which are required as input to VA09 are provided by the subroutine MFLD. Concerning second order derivatives, only a positive definite initial estimate of the Hessian matrix is needed. We have chosen the

simpliest possible alternative, i.e. the identity matrix. VA09 then gradually approximates the inverse Hessian. The property of positive definiteness is preserved in the updating process, so as to avoid the tendency of Newton methods to move in the wrong direction when far from the optimum.

The single stopping or convergence criterion employed is the accuracy that the user requires in the estimates. Thus, the algorithm is terminated when

$$\hat{\theta}_i^{r+1} - \hat{\theta}_i^{r} < \text{EPS} \quad \text{for } i=1,2,\ldots,k,$$

where the superindices denote iteration number and EPS is the prespecified accuracy.[6] Although this is one obvious convergence criterion, there are several alternatives, some of which might be more efficient.[7]

According to our experience the algorithm is remarkably reliable. For almost any conceivable starting values for the θ_i:s it converges rapidly to a local minimum. In case of non-convergence the most probable cause is that the optimization problem is not well defined (for some $\boldsymbol{\theta}$:s). An illustration of such a case is given by the example in Appendix C, where the model is not identified for certain sets of parameters.

In order to avoid unnecessary long execution times it is recommended, however, that the problem be scaled such that the θ_i:s do not differ in magnitude by more than a factor of 100. That can always be achieved by a slight change in the mapping from the θ_i:s to the coefficients, e.g. by specifying the relevant parameters according to

$\kappa_i = c\theta_i'$

where c is a constant.

The first order derivatives are usually very close to zero at the solution. In general, the biggest element of the gradient vector should not exceed $1 \cdot 10^{-6}$.

Beside the derivatives there are two additional checks on the solution. One is the integer IEXIT, which gives the reason for the exit from VA09. It can take on the following values.

IEXIT=1 The normal exit, in which the accuracy condition has been fulfilled.

IEXIT=2 The specified accuracy has not been obtained. Probable cause is that EPS has been set too small for computer word length.

IEXIT=3 Maximum number of function evaluations MXFN, has been reached.

The other test concerns the estimate of the Hessian matrix. By construction, this matrix is positive definite and so its eigenvalues are all strictly positive. Equivalently, for the Cholesky factorization **L'DL** of the Hessian, the Cholesky values, i.e. the diagonal elements of **D**, are all positive.[8]

One (or several) of the Cholesky values may be very small, however, indicating that the solution is not wholly satisfactory. That can happen, for instance, if one of the parameters is just barely identified, due to strong collinearities in the data. In such cases the difference between the largest and the smallest of the Cholesky values will be big and so the ratio,

$$\delta = d_{min}/d_{max}$$

will be small.

One measure which can be used as a lower bound for δ is the Euclidean norm of the gradient vector. Denoting the gradient vector by **g**, the Euclidean norm is defined according to

$$\|\mathbf{g}\| = \left(\sum_{i=1}^{k} g_i^2 \right)^{1/2}$$

Thus, for a well-behaved solution the following inequality should hold

$$1 > \frac{\|\mathbf{g}\|}{\delta} = \text{COND},$$

which is the number computed by the program. Ordinarily COND should be much smaller than one, certainly not exceeding 0.1.

Before concluding this section an interesting result reported by Belsley (1980), will be remarked upon.

Considering the DFP algorithm, Belsley finds that execution times can be considerably reduced (up to

50 percent) if the Hessian is initialized by the estimate proposed by Berndt, Hall, Hall and Hausman (1974), instead of by the identity matrix. The fact that the BHHH Hessian is quite easy to program makes the suggested approach particularly attractive.

According to a (very) small number of tests, Belsley's conclusions do not seem to hold for Fletcher's algorithm, however. In fact, initialization by the BHHH Hessian actually increased execution times slightly in a few cases. The procedure has thus not been implemented in the present program.

3.3 Computation of $\mathrm{Var}(\hat{\hat{\theta}})$

It can be shown that the variance of $\hat{\theta}$ fulfills the following inequality[9]

$$\mathrm{Var}(\hat{\theta}) \geq -\{E[\frac{\partial(\partial L^*(\theta_o)/\partial\theta)'}{\partial\theta}]\}^{-1} = \mathbf{M}^{-1}$$

where θ_o is the true parameter vector and $\mathbf{M}$ the so called information matrix. Since matrices are involved, the interpretation of the inequality sign is that the difference between $\mathrm{Var}(\hat{\theta})$ and $\mathbf{M}^{-1}$ is a positive semidefinite matrix. Equality holds asymptotically, as the sample size approaches infinity.

One way to estimate $\mathrm{Var}(\hat{\theta})$ is thus to determine the analytic expression for $\mathbf{M}^{-1}$ and evaluate it at $\hat{\theta}$. For the model in sec. 2, $\mathbf{M}$ can be found by a simple generalization of Hendry's (1974) result, so as to include nonlinear restrictions on the parameters.[10]

However, since **M** is a rather complicated matrix, which is difficult to program, we have instead based our estimate of Var $(\hat{\boldsymbol{\theta}})$on the Hessian matrix produced by the minimization routine (cf. above). A justification for this is given by the following alternative form of the above inequality,[11]

$$\mathrm{Var}(\hat{\boldsymbol{\theta}}) \geq -\mathrm{plim}\left[\frac{\partial(\partial L^*(\hat{\boldsymbol{\theta}})/\partial\boldsymbol{\theta})'}{\partial\boldsymbol{\theta}}\right]^{-1} = -\mathrm{plim}\left[\mathbf{G}(\hat{\boldsymbol{\theta}})\right]^{-1}$$

where plim denotes probability limit and $\mathbf{G}(\hat{\boldsymbol{\theta}})$ the Hessian, evaluated at $\hat{\boldsymbol{\theta}}$. The inverse of $\mathbf{G}(\hat{\boldsymbol{\theta}})$ is thus a consistent estimate of the lower bound for $\mathrm{Var}(\hat{\boldsymbol{\theta}})$.

As the minimization algorithm makes a second order approximation of the function to be minimized, the Hessian matrix obtained from it will be equal to $\mathbf{G}(\hat{\boldsymbol{\theta}})$ only if L* is quadratic. However, at least in large samples, L* is known to be very nearly so. Further, for the models considered by Belsley (1980) the Hessian matrices produced by two different quasi-Newton routines were very close to their analytic counterparts, in spite of the sample sizes being very moderate, 20-40 observations.

Although the two estimates of $\mathrm{Var}(\hat{\boldsymbol{\theta}})$ described above are asymptotically equivalent they may of course differ in small samples. In general, variance estimates based on the Hessian matrix tend to be greater than those based on the information matrix.[12] Accordingly, one would expect inferences based on the former estimates to be more cautious.

3.4 Compatability with other computers

In some respects the proper functioning of the program relies on characteristics which may be specific to DEC computers. Without any pretension to completenes a few remarks will be given on a couple of such aspects.

The main obstacle to run CONRAD on other computers probably concerns the input and output of data and the handling of the corresponding files. For instance, the subroutine OC which opens and closes input and output files cannot be directly transferred on other computers. Further, the terminal output (TYPE statements) will generally not work either, on other computers.

Concerning computations, extensive use has been made of the automatic zero initialization of all variables and arrays, provided by the DEC FORTRAN compiler. This feature is, however, common to most modern compilers.

Since all parts of the program are not stored in one file, loading and execution are performed with the help of a MIC-file (Macro Interpreted Commands), initialized by the command /CONRAD. This facility, which allows operative system commands to be collected in an executable file, is also machine dependent.

Finally, a minor difficulty might occur because some machines, e.g. IBM, have lower character capacity than the DEC computers. In the call on the printout subroutine PVM one of the arguments is a

character variable, used as heading - e.g. the name of a matrix. This variable has been explicitly declared in double precision, allowing the heading to consist of ten characters. The corresponding character capacity on IBM computers permit only eight characters, however, thus requiring the longest headings to be shortened.

NOTES

[1] Cf. DECSYSTEM10; FORTRAN-10 Programmer's Reference Manual, 6:th printing, 1975.

[2] Since the double precision has been implemented using the IMPLICIT statement, it can easily be eliminated if deemed unnecessary, however.

[3] I.e. of the Newton type but not requiring evaluation of second order derivatives. See e.g. Quandt (1983, pp. 721-722).

[4] An auxiliare routine, MC11AD, from the same library is also employed.

[5] The linear search problem is discussed in Quandt (1983, pp. 735-737). For comparisons between Fletcher's algorithm and the DFP, see Fletcher (1970).

[6] Termination may also occur because the prespecified maximum number of function evaluations, MXFN, has been reached.

[7] See Belsley (1980) and Quandt (1983, pp. 737-738).

[8] The Cholesky factorization is briefly discussed in Theil (1983). An extensive treatment and a proof of the stated equivalence can be found in Lau (1978).

[9] Se Silvey (1975).

[10] See Mellander (1984).

[11] Cf. Pollock (1979, p. 345).

[12] Cf., e.g., Calzolari and Panattoni.

1984-09-10, EM/Conrad-kap4-i, disk TT62

4 INPUT INSTRUCTIONS

Three input files are used. Each of these will be described separately, under headings equal to the names wich should be assigned to them.

Common to all the three files is the use of free format input. Hence, considering a particular row the only requirement is that the strings (values) appear in the prescribed order, separated by at least one blank position.

Regarding the observations on the endogenous and predetermined variables, a distinction is made between the dataset in the input file and the data actually used in the estimation. The latter may be a subset of the former, differing both in number of variables and/or observations.

The choice of variables is completely free. Arbitrary subsystems of the total system comprised in the data material, may thus be estimated without renewing the data input. The choice of observations is somewhat more limited, but, e.g., in the context of time series data it allows the estimation to be based on any coherent subperiod.

Sections 4.2 and 4.3 contain schematic representations of the respective files. These are such that one row in the manual corresponds to one row in the file. The notations have been chosen so as to be easily recognizable with those used in section 2.2.

Neither of the input files contain information needed to actually run the program. Such commands have instead been gathered in a special file. Thus, when the input files have been properly arranged the execution can be started by the simple command /CONRAD. However, before doing so, Appendix C should be used to check that the storage space currently allocated in the program is sufficient for the problem at hand.

4.1 DATAM.DAT

This file contains observations on the endogenous and predetermined variables.

The data must be arranged in matrix form. The matrix should be structured so that the endogenous and predetermined variables form the block matrices, according to

$$\begin{matrix} y_{11} & y_{12} & \cdots & y_{1\eta} & z_{11} & z_{12} & \cdots & z_{1\xi} \\ y_{21} & y_{22} & \cdots & y_{2\eta} & z_{21} & z_{22} & \cdots & z_{2\xi} \\ \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot \\ y_{\tau 1} & y_{\tau 2} & \cdots & y_{\tau\eta} & z_{\tau 1} & z_{\tau 2} & \cdots & z_{\tau\xi} \end{matrix}$$

where

η the number of endogenous variables,
ξ the number of predetermined variables,
τ the number of observations on the variables.

As mentioned in the previous section, the data actually used in the estimatin may be a subset of the data matrix in this file. The indices above thus constitute upper limits for the corresponding indices of the **X** matrix (cf. section 2.1), according to $n \leq \eta$, $m \leq \xi$ and $T \leq \tau$.

In principle, one row in the data matrix should correspond to one row in the input file. However, with many variables that may be impossible, due to lack of space in the file's rows. In that case the matrix can be partioned into submatrices of appropriate sizes. The submatrices should then be "stacked" in the file, one underneath the other. For instance, if the user wants to split up the data in two parts, the first part containing all the endogenous variables and the first predetermined variable, that could be done in the following way:

$$
\begin{array}{llll}
y_{11} & y_{12} & \cdots & y_{1\eta} \quad z_{11} \\
y_{21} & y_{22} & \cdots & y_{2\eta} \quad z_{21} \\
\vdots & \vdots & & \vdots \quad \vdots \\
y_{21} & y_{22} & \cdots & y_{\tau\eta} \quad z_{\tau 1} \\
z_{12} & \cdots & z_{1\xi} & \\
z_{22} & \cdots & z_{2\xi} & \\
\vdots & & \vdots & \\
z_{\tau 2} & \cdots & z_{\tau\xi}. &
\end{array}
$$

The number of blocks and the number of columns in each block is freely determined by the user. Dif-

ferrent blocks need not have equally many columns. They must, however, have the same number of rows (i.e τ).

The blocks may be separated by blank rows. This may be convenient, e.g. making it easier to check the file. The number of blank rows - which may be zero - is set by the user. However, all blocks must be separated by the same number of blank rows. No blank rows should preceed the first block.

4.2 PARIC.DAT

Initial values for the unrestricted parameters and information about inequality constraints should be provided in this file.

File structure

k			
$\bar{\theta}^1$	$i(\bar{\theta}^1)$	$\ell(\bar{\theta}^1)$	$\gamma/\upsilon(\bar{\theta}^1)$
$\bar{\theta}^2$	$i(\bar{\theta}^2)$	$\ell(\bar{\theta}^2)$	$\gamma/\upsilon(\bar{\theta}^2)$
.	.	.	.
.	.	.	.
.	.	.	.
$\bar{\theta}^k$	$i(\bar{\theta}^k)$	$\ell(\bar{\theta}^k)$	$\gamma/\upsilon(\bar{\theta}^k)$

where

k the number of unrestricted parameters to be estimated

$\bar{\theta}^{\alpha}$ α=1, ..., k. The initial value of the parameter having the α:th ordinal number. (Ordinal numbers will always be indicated by superindicies, so as to avoid confusion with the subindices used in section 2.)

$i(\bar{\theta}^{\alpha})$ $1 \leq i \leq k$. The (sub)index of the parameter whose initial value is $\bar{\theta}^{\alpha}$.

$\ell(\bar{\theta}^{\alpha})$ =0 if no hyperbolic transformation should be applied to the parameter (default),
=1 if the parameter should be transformed according to (5a),
=2 if the parameter should be transformed according to (5b).

$\gamma/\upsilon(\bar{\theta}^{\alpha})$ Required only if $\ell=\ell(\bar{\theta}^{\alpha}) \geq 1$. If ℓ=1 the parameter's lower limit, i.e. γ, should be given and if ℓ=2 the upper limit, υ. Default value is zero.

4.3 GICOF.DAT

In this file some general information should first be given, regarding, i.a., the stochastic specification and the format of output. Then the file DATAM.DAT should be described. Finally, the desired data set and the nonzero coefficients have to be specified.

A subset of the data in DATAM.DAT is determined in the following way. The relevant observations are indicated by the row numbers of the data matrix, that correspond to the "first" and "last" observations. To select the appropriate variables, maximum numbers are first set for both the endogenous

and the predetermined variables. In the respective block matrices (cf. 4.1) these numbers are equal to the column numbers corresponding to the rightmost of the considered variables. Whether the maximum numbers of variables will be equal to the numbers of variables actually taken into account in the estimation, is decided by the specification of the coefficients in the **A** matrix.

All coefficients not explicitly specified will be set equal to zero. This also applies to the diagonal elements of the **B** matrix (which usually are automatically set to -1 in most programs). Among the variables chosen in the first step (cf. above), certain ones can thus be ignored in the estimation by not specifying the coefficients associated with them.

No difference is made between Θ_i, $h_1(\Theta_i)$ and $h_2(\Theta_i)$ when the coefficients are specified. All three of these expressions are identified with the index i. The task here is thus only to assign values to the constants in the polynomials in (8a) and (8b) in section 2.2.2.

As all nonzero coefficients should be defined according to the same principle only one coefficient is shown in the schematic file structure below. For easy reference to (8), constants in the polynomials have been written with small letters. To simplify the notation the common index g has been suppressed, however.

As in the previous section, superindicies denote ordinal numbers.

File structure

LP EPS MXFN WIDTH HEAD

NY NZ NO

NBLC ICL(1) ... ICL(NBLC) NBR

IFO ILO NMX MMX

(Title)

NA IOR^1 IOR^2 ...

IA JA ITYPE

r_o ℓn $n\ell n$ s_o ℓd $n\ell d$

r^1 $i[r^1]$

. .

. .

$r^{\ell n}$ $i[r^{\ell n}]$

ρ^1 $i[\rho^1]$ $j[\rho^1]$

. . .

. . .

$\rho^{n\ell n}$ $i[\rho^{n\ell n}]$ $j[\rho^{n\ell n}]$

s^1 $i[s^1]$

. .

. .

$s^{\ell d}$ $i[s^{\ell d}]$

σ^1 $i[\sigma^1]$ $j[\sigma^1]$

. . .

. . .

$\sigma^{n\ell d}$ $i[\sigma^{n\ell d}]$ $j[\sigma^{n\ell d}]$

} a_g

LP Determines the stochastic specification;
=1 for autocorrelated errors,
=0 otherwise.

EPS A real positive number, specifying the accuracy of the estimates (cf. 3.2).

MXFN Should be set to 0 in the first run, resulting in printout of all input data in RESF-1.DAT. Otherwise set equal to the maximum number of evaluations of the log-likelihood function.

WIDTH The maximum number of character positions per row, in the main output file.

HEAD Required only if the user wants to provide the output files with a title, then HEAD=1.

NY The number of endogenous variables in the data matrix in DATAM.DAT.

NZ The number of predetermined variables in the data matrix in DATAM.DAT.

NO The number of observations on the variables in DATAM.DAT.

NBLC The number of blocks into which the data matrix in DATAM.DAT is partitioned.

ICL(i) Required only if NBLC>1. The number of columns in the i:th block, i=1,2, ..., NBLC.

NBR The number of blank rows between the blocks. Assumed to be zero if omitted.

IFO, ILO — The row numbers of the data matrix in DATAM.DAT, corresponding to the first and last observations, respectively, that are to be considered in the estimation.

NMX, MMX — The maximum numbers of endogenous and predetermined variables, respectively, that are to be considered in the estimation (cf. the beginning of this section).

Title — Only if HEAD=1. The maximum length of the title is 76 characters, including blanks.

NA — The number of <u>all</u> nonzero coefficients in the **A** matrix.

IOR^i — The numbers of the equations containing intercepts. These are required only for the calculation of goodness-of-fit measures. A particular equation is identified by the number of the column in the data matrix that contains its endogenous variable.

IA, JA — The coefficients' row and column indices, respectively, in the **A** matrix.

ITYPE
=0 if the coefficient is restricted, a priori, to be equal to a constant,
=1 if the coefficient is equal to (a special case of) n_g in (8a),
=2 otherwise.

r_o — The constant in the polynomial n_g which constitutes the numerator of a_g. If ITYPE=0, r_o is set equal to the a priori constant and the definition of the coefficient is completed.

ℓn, $n\ell n$ Specified only if ITYPE>0. The number of terms in the linear and nonlinear parts of n_g, respectively. Thus, ℓn is equal to the number of (nonzero) r:s and $n\ell n$ to number of ρ:s.

s_o, ℓd, $n\ell d$ Specified only if ITYPE=2. The s_o is the constant in d_g. ℓd and $n\ell d$ are defined in analogy with ℓn and $n\ell n$, i.e. ℓd is the number of s:s and $n\ell d$ the number of σ:s.

r^κ, $i[r^\kappa]$ Specified only if $\ell n>0$. $\kappa=1,\ldots,\ \ell n$. r^κ is the κ:th constant and $i[r^\kappa]$ the index of the parameter associated with this constant. As indicated by the use of superindices, no particular order is required.

ρ^μ, $i[\rho^\mu]$ $j[\rho^\mu]$ Specified only if $n\ell n>0$. $\mu=1,\ldots,\ n\ell n$. The notations are analogous with those used for the r:s. The i- and j-indices may be equal.

If ITYPE=2, the nonzero s- and σ-constants with related indices are then specified, in the same way as the r:s and ρ:s.

EM/Conrad-kap5-h 1984-09-19 disk TT6,

5 COMMENTS ON OUTPUT

The results are presented in four output files. In addition, some information - e.g. concerning the development of the likelihood function in the the search for the optimum - is provided as terminal output, during the execution.

The content of this section is organized in the same way as that of the preceeding one, i.e. it is divided into subsections corresponding to the different output files. Further, there is also a short subsection about the terminal output.

It is recommended that this section and the example in Appendix C be read in parallel.

5.1 RESF1.DAT

This file will be created only if the parameter MXFN is set to zero. In that case the input to CONRAD, organized in easily readable form, will be printed here, whereupon the execution will be terminated. This should enable the user to check that the input has been properly arranged, before turning to the actual estimation. Further, the file provides a thorough documentation of the data and specifications used in the regression.

Concerning the data, only the NMX and MMX first columns of the **Y** and **Z** matrices are printed, respectively, beginning with the IFO:th observation

and ending with the ILO:th. (The format is F11.6.) The number of observations (i.e. ILO-IFO+1) is then explicitly given, together with the values of IFO and ILO.

If the vector autoregressive stochastic specification has been chosen, that will be indicated by a statement saying that the number of observations on the model is one less than the number of observations on the variables.

Then the initial values for the parameters are given, together with information about the inequality constraints. Actually, the initial values subject to the chosen hyperbolic transformations are given, rather than the initial values themselves. The reason for this is just that it makes it easier to check the equality constraints (cf. below).

The required accuracy in the estimates, i.e. the value of the parameter EPS (cf. sec. 3.2), is also shown.

Since the program sorts the rows of PARIC.DAT according to parameter index, the initial values and information about the transformations will appear in increasing order, irrespective of the order in which they have been input. The transformations are identified by the numbers 0 (no transformation), 1 and 2 as set out in sec 2.2.

Finally, the values of the nonzero elements of the coefficient matrix **A** are given, together with the corresponding equality constraints. The values of the a_{ij}:s should be the ones implied by the specified constraints and the initial values of

the (hyperbolically transformed) parameters. No change is made with respect to order, i.e. the coefficients appear in the same order as they have been input in GICOF.DAT.

For each of the coefficients the type of constraint is indicated by the value of ITYPE (cf. sec 4.3). The constants in the mapping from the (transformed) parameters are also given, together with the indices of the corresponding parameters. The constants are denoted in the same way as in (8a) and (8b) in sec. 2.2.2, except that the ρ:s and σ:s are denoted "greek r" and "greek s", respectively.

5.2 TERMINAL OUTPUT

One type of terminal output concerns checks on input and problem specification. The program performs a number of such checks, before starting on the actual estimation. If an error is detected it is communicated to the user via the terminal, whereafter the execution is stopped.

During the optimization the function value and its two basic components will be printed according to

$$F \quad \ell n(\det(\mathbf{B})) \quad \ell n(\det(\mathbf{\Sigma}))$$

each time the function is evaluated ("ℓn" and "det" denote natural logarithm and determinant, respectively). The relationship between the function F and the two determinants is given by

$$F = \frac{1}{T}\left[\frac{1}{2}\ell n(\det(\mathbf{\Sigma})) - \ell n(\det(\mathbf{B}))\right]$$

whereas the relationship between F and the log-likelihood function is given in sec. 2.4.

After the final iteration, "END OF VA09" will be printed. The estimated parameter vector and the corresponding gradient vector are then given, their elements ordered according to parameter index. A final printing of F and the logarithms of the two determinants conclude the terminal output.

5.3 RESF2.DAT

This is the main output file. The first line contains some technical information, conveyed through the variables IEXIT and COND. If the run has been succesful IEXIT should equal unity and COND be less than 0.1 (cf. sec. 3.2). The final F-value and the logarithms of the corresponding determinants are then given (cf. the previous section).

To enable computation of the generalized R^2 measure proposed by Berndt[1] the logarithm of the determinant of a matrix denoted **(Y-YBAR)'(Y-YBAR)** is also given. The generalized R^2, which measures the goodness of fit for the whole system of equations, can be obtained according to

$$\tilde{R}^2 = 1 - \exp\{\ell n(\det(\mathbf{\Sigma})) - 2\,\ell n(\det(\mathbf{B})) - \ell n[\det((\mathbf{Y-YBAR})'(\mathbf{Y-YBAR}))]\}$$

Ordinarily, the i:th column of **YBAR** will contain the mean value of the i:th endogenous variable, the matrix **Y-YBAR** thus containing the deviations of the endogenous variables from their respective means. However, if the i:th equation does not contain an intercept, the i:th column of **YBAR** will contain zeros only.

In case some of the NMX first equations are not considered in the estimation they are of course not considered in the calculation of det[(**Y-YBAR**)'(**Y-YBAR**)] either.

To indicate the relevant alternative of these three possibilities with respect to the different equations, a row of the matrix **YBAR** has been printed below the logarithm of det[(**Y-YBAR**)'(**Y-YBAR**)]. Equations which have not been considered in the computation of the determinant are indicated by the value of -999.

The estimated parameter vector is then given, followed by the corresponding vectors of first order derivatives (the gradient) and standard errors, respectively.

To make it possible to calculate the variances of coefficients involving several parameters, the complete covariance matrix of the parameter vector is also printed. (The standard errors have, of course, been obtained as the square root of the diagonal elements of this matrix.) The heading, HESIAN(-1), where (-1) indicates inverse, refers to the way this matrix has been estimated; see further sec. 3.3.

Most of the following output should be easily understood, since the notation corresponds directly to the one used in sec. 2.

The goodness of fit measure used for the individual equations needs some explanation, however. It is equal to the squared cosine of the angle between either the vector of observations on the endoge-

nous variable (**y**) and the corresponding predicted vector (**yp**), or between the same vectors, measured as deviations from their respective means.[2] The latter variant, which is used if the equation contains an intercept, is equal to the squared correlation coefficient between **y** and **yp**. [Cf. Wonnacott and Wonnacott (1970, p. 303).]

In contrast to the conventional R^2, which is not bounded from below when applied to one of the equations in a simultaneous system, this measure always lies in the closed interval [0,1]. It is also applicable in the case of nonlinearities, which R^2 is not.[3]

The d statistic of Durbin and Watson (1950, 1951) is supplied in connection with the goodness of fit measure (under the heading "D-W"). It should be pointed out, however, that the distributional properties of this statistic are not well known in the context of simultaneously estimated systems of equations.[4] This holds in particular if nonlinearities are present. It is thus intended as an indicative measure rather than as a test statistic. For testing purposes the LR test described at the end of sec. 2.4 is instead recommended.[5]

5.4 RESF3.DAT

The content of this file is supplied by the optimization algorithm. Presumably, it will be of interest only if the execution is unsatisfactory in some technical sense - as indicated, e.g., by the values of IEXIT and COND, given at the beginning of RESF2.DAT.

For every fifth iteration, and also on exit, the following is printed

Iteration number, Number of calls of MFLD
The value of F
$\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k$ (5 to a line)
$g_1, g_2, \ldots, g_k$ ("-)

where g_i is the first order derivate of F with respect to $\hat{\theta}_i$.

On exit, the value of IEXIT is also printed, after the number of calls on MFLD.

It is the number of calls on MFLD that is delimited by the setting of the parameter MXFN (cf. sec. 4.3).The number of iterations will be at most equal to the number of calls on MFLD but usually somewhat smaller.[6]

5.5 RESF4.DAT

This file is organized in the same way as the input file PARIC.DAT, the only difference being that the final estimates of the parameters have been substituted for the initial values.

After having been renamed it can thus serve as an input file in the next run. This is convenient if, e.g., the optimization algorithm has been stopped before having reached the optimum, due to a too low value having been assigned to MXFN. The search can then continue from the point where it was previously terminated.

Another conceivable use of the file is to provide starting values when the stochastic specification is changed from the one assuming time-independent residuals to the one allowing them to follow a first order vector autoregression.

NOTES

[1] Cf. Berndt and Khaled (1979, p. 1228). Other generalized measures have been proposed by e.g. Carter and Nagar (1977). The advantage of Berndt's measure is that it is the only one which is invariant to whether the structural or the reduced form of the system is considered.

[2] Cf. Haessel (1978).

[3] There is a close relationship between the two measures, however, in that they take on identical values when applied to a linear equation with intercept, estimated by OLS.

[4] Cf., however, Durbin (1957).

[5] If, still, used as a test statistic Farebrother's (1980) tables should be used if the equation does not contain an intercept. Further, if the equation contains lagged endogenous variables the "h" statistic should be substituted for the d statistic. The former can, however, be computed with the help of the latter, cf. Durbin (1970 p. 419).

[6] The reason is that the algorithm sometimes has to seek the minimum along a line, in order to obtain a sufficient reduction in F, in which case the iteration will involve several function evaluations. This is the "linear search subproblem" referred to in sec. 3.2.

Appendix A THE CONCENTRATED LOG-LIKELIHOOD FUNCTION AND ITS FIRST ORDER DERIVATIVES

The (unconcentrated) log-likelihood function corresponding to the equations (1), (2) and (3) is given by[1]

$$L(\Theta,\mathbf{H},\boldsymbol{\Sigma}) = k + T\ell n|\mathbf{B}| - \frac{1}{2}T\ell n|\boldsymbol{\Sigma}| - \frac{1}{2}tr[(\mathbf{U}'\mathbf{U} - 2\mathbf{H}\mathbf{U}_1'\mathbf{U} + \mathbf{H}\mathbf{U}_1'\mathbf{U}_1\mathbf{H}')\boldsymbol{\Sigma}^{-1}], \tag{a1}$$

where index 1 denotes one period lag and

$$k = -\frac{1}{2}nT\ell n(2\Pi).$$

It should be noticed that T is the number of observations on the <u>model</u>, which is one less the number of observations on the variables.

Application of matrix differentiation techniques,[2] yields the following first order conditions for $\mathbf{H}$ and $\boldsymbol{\Sigma}$, respectively.

$$-\mathbf{U}_1'\mathbf{U}_1\mathbf{H}'\boldsymbol{\Sigma}^{-1} + \mathbf{U}_1'\mathbf{U}\boldsymbol{\Sigma}^{-1} = \mathbf{0} \tag{a2}$$

$$-\frac{1}{2}T\boldsymbol{\Sigma}^{-1} - \frac{1}{2}(\mathbf{U}'\mathbf{U} - \mathbf{U}'\mathbf{U}_1\mathbf{H}' - \mathbf{H}\mathbf{U}_1'\mathbf{U} + \mathbf{H}\mathbf{U}_1'\mathbf{U}_1\mathbf{H}') = \mathbf{0}. \tag{a3}$$

Since $\mathbf{H}$ and $\boldsymbol{\Sigma}$ are both unrestricted, they can be solved for from (a2) and (a3).[3] Substitution of the solutions into (a1) gives the concentrated log-likelihood function.

The solution for $\mathbf{H}$, implied by (a2), is

$$\tilde{\mathbf{H}} = \mathbf{H}(\Theta)$$

$$= \mathbf{U}'\mathbf{U}_1(\mathbf{U}_1'\mathbf{U}_1)^{-1}. \tag{a4}$$

By substitution of (a4) in (a3) $\boldsymbol{\Sigma}$ can also be expressed as a function of Θ, according to

$$\tilde{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}(\Theta)$$

$$= T^{-1}(\mathbf{U}'\mathbf{U} - \tilde{\mathbf{H}}\mathbf{U}_1'\mathbf{U}). \tag{a5}$$

The concentrated log-likelihood function thus becomes

$$L^*(\Theta) = k^* + T\ell n|\mathbf{B}| - \frac{1}{2}T\ell n|\tilde{\boldsymbol{\Sigma}}|, \tag{a6}$$

where

$$k^* = -\frac{1}{2}nT[\ell n(2\Pi) + 1].$$

Differention of L^* yields[3,4]

$$\left(\frac{\partial L^*}{\partial \Theta}\right)' = \left(\frac{\partial \mathrm{vec}\mathbf{A}}{\partial \Theta}\right)' \mathrm{vec}[T(\mathbf{B}^{-1}:\mathbf{0}) - \tilde{\boldsymbol{\Sigma}}^{-1}(\mathbf{U}'-\tilde{\mathbf{H}}\mathbf{U}_1')\mathbf{X} + \tilde{\mathbf{H}}'\tilde{\boldsymbol{\Sigma}}^{-1}(\mathbf{U}'-\tilde{\mathbf{H}}\mathbf{U}_1')\mathbf{X}_1], \tag{a7}$$

where the $n(n+m)\times k$ matrix $\partial \mathrm{vec}\mathbf{A}/\partial\Theta$ can be written

$$\frac{\partial \mathrm{vec}\mathbf{A}}{\partial \Theta} = \left(\frac{\partial \mathrm{vec}\mathbf{A}}{\partial \Theta_1}, \frac{\partial \mathrm{vec}\mathbf{A}}{\partial \Theta_2}, \ldots, \frac{\partial \mathrm{vec}\mathbf{A}}{\partial \Theta_k}\right). \tag{a8}$$

Denoting a particular element of vec**A** by a_g (cf. sec. 2.2.2), the derivative of this coefficient with respect to the i:th parameter can be obtained according to

$$\frac{\partial a_g}{\partial \theta_i} = \frac{\partial a_g}{\partial h_\ell(\theta_i)} \frac{\partial h_\ell(\theta_i)}{\partial \theta_i} . \tag{a9}$$

Expansion of the first factor on the RHS yields

$$\frac{\partial a_g}{\partial h_\ell(\theta_i)} = \left[d_g \frac{\partial n_g}{\partial h_\ell(\theta_i)} - n_g \frac{\partial d_g}{\partial h_\ell(\theta_i)} \right] / d_g^2 . \tag{a10}$$

where

$$\frac{\partial n_g}{\partial h_\ell(\theta_i)} = r_{gi} + 2\rho_{gij} h_\ell(\theta_i) + \sum_{j \neq i} \rho_{gij} h_\ell(\theta_j) \tag{a11}$$

and an analogous expression for $\partial d_g / \partial h_\ell(\theta_i)$, with r and ρ replaced by s and σ, respectively.

For $\ell=0$ the last factor on the RHS of (a9) equals unity (cf. sec. 2.2.2). If $\ell=1$ then

$$\frac{\partial h_1(\theta_i)}{\partial \theta_i} = \begin{cases} [\theta_i / h_1(\theta_i)]^5 & \text{for } \gamma_i > 0 \\ \theta_i / [h_1(\theta_i) + 0.05] & \text{for } \gamma_i = 0 \end{cases} \tag{a12}$$

(cf. sec. 2.2.1). The corresponding formula for $\ell=2$ is obtained by mere substitution of 2 and υ_i for 1 and γ_i, respectively, and multiplication by -1.

NOTES

[1] Cf. Hendry (1971).

[2] Cf., e.g., Neudecker (1969) and Pollcock (1979). Derivations based on Pollock's conventions can be found in Mellander (1984).

[3] The solutions for $\mathbf{H}$ and $\boldsymbol{\Sigma}$, given below, can also be found in Hendry, op.cit.

[4] A derivation is given in Mellander (1984). Following Pollock (op.cit.) $\partial L^*/\partial \boldsymbol{\theta}$ is defined to be a row vector; thus the transposition.

[5] The corresponding expression for the case when $\mathbf{A}$ is unrestricted, is given in Hendry (op.cit.). Due to a misprinting, the prime on the first $\mathbf{H}$ in the last term has been omitted there, however.

EM/Conrad-AppB-c 1984-09-11, disk TT62

Appendix B DIMENSION STATEMENTS

Minimum dimension requirements are given in the following copy of the file CONRAD.DOC.("*" indicates multiplication.)

```
     CONRAD.DOC

MAIN PROGRAM:
=============
DIMENSION ICL(15),TITLE(19),IOR(NMX),ITETA(KP)
DIMENSION TETA(KP),GRAD(KP),EPSV(KP),HES(KP*KP)
DIMENSION YBR(NMX),YP(T*NMX),YPBR(NMX),YLS(NMX*NMX)
COMMON/DATA/ X(T*(NY+MMX)),X1(T*ICMX),NMX,MMX,T,ICMX,TINV
COMMON/COEFF/ NA,NOA(NA,2),ITYP(NA),NI(NA,4),ROSO(NA,2),
1              RS(&,2),IXL(&,2),GRGS($,2),IXNL($,4),
2              A(NMX*ICMX),BT(NMX*NMX),HA(NMX*ICMX)
COMMON/PAR/ IHT(KP),CHT(KP),KDUM
COMMON/RES/ U(T*NMX),U1(T*NMX),U1HT(T*NMX),
1           UTU1(NMX*NMX),U1SQ(NMX*NMX),SIGMA(NMX*NMX)

SUBROUTINE MFLD:
================
DIMENSION ZZ(KP),NG(NA),DG(NA),FI(NMX*ICMX)
COMMON/DATA/
COMMON/COEFF/
COMMON/PAR/
COMMON/RES/

SUBROUTINE POL:
===============
COMMON/COEFF/
COMMON/PAR/

SUBROUTINE DPOL:
================
COMMON/COEFF/

SYMBOLS NOT USED (IN SEC. 4) IN THE MANUAL:
KP    = THE NUMBER OF PARAMETERS TO BE ESTIMATED (DENOTED BY SMALL
        K IN THE MANUAL).
T     = ILO-IFO+1
ICMX  = NMX+MMX
&     = MAX(&1,&2) WHERE &1 AND &2 ARE THE NUMBER OF "R"- AND
        "S"-CONSTANTS DECLARED, RESPECTIVELY.
$     = MAX($1,$2) WHERE $1 AND $2 ARE THE NUMBER OF "GREEK R"-
        AND "GREEK S"-CONSTANTS DECLARED, RESPECTIVELY.
```

The COMMON blocks of the subroutines are identical with the corresponding blocks in the main program. Thus, when dimensions are changed the newly declared blocks can simply be copied from the main program into the subroutines.

EM/Conrad-AppC-g 1984-09-19, disk TT62

Appendix C AN EXAMPLE[1]

The example is based on the disequilibrium version of the supply and demand model for exports, suggested by Goldstein and Khan (1978). The model will be estimated on Swedish annual data 1959-80, under the two alternative stochastic specifications that the program can allow for.

Goldstein and Khan consider the following demand and supply equations for exports.

$$\log x_t^d = \alpha_0 + \alpha_1 \log(px/pxw)_t + \alpha_2 \log yw_t \qquad (c1)$$

where

x^d = quantity of exports demanded
px = price of exports
pxw = weighted average of the export prices of the country's trading partners
yw = weighted average of the real incomes of the country's trading partners

and

$$\log x_t^s = \beta_0 + \beta_1 \log(px/p)_t + \beta_3 y_t^* \qquad (c2)$$

where

x^s = quantity of exports supplied
px = price of exports
p = domestic price index
y* = logarithm of an index of domestic capacity.

---

[1] The dataset used in this example was kindly provided by Eva Christina Horwitz, OECD, Paris.

To these equations the following partial adjustment mechanisms are added

$$\log x_t - \log x_{t-1} = \gamma(\log x_t^d - \log x_{t-1}) \qquad (c3)$$

and

$$\log px_t - \log px_{t-1} = \lambda(\log x_t - \log x_t^s). \qquad (c4)$$

To get the notation in accordance with that used in section 2, let

$$\theta_1 = \gamma \ , \ \theta_2 = \alpha_0 \ , \ \theta_3 = \alpha_1 \ , \ \theta_4 = \alpha_2$$

$$\theta_5 = \lambda \ , \ \theta_6 = \beta_0 \ , \ \theta_7 = \theta_1 \ , \ \theta_8 = \beta_2 .$$

Substituion of (c1) in (c3) and (c2) in (c4) then yields the following system

$$\log x_t = a_{12}\log px_t + a_{13} + a_{14}\log pxw_t + a_{15}\log yw_t + a_{18}\log x_{t-1} \qquad (c5)$$

$$\log px_t = a_{21}\log x_t + a_{23} + a_{26}\log p_t + a_{27}y_t^* + a_{29}\log px_{t-1} \qquad (c6)$$

where

$$a_{12} = \theta_1\theta_3, \quad a_{13} = \theta_1\theta_2, \quad a_{14} = -a_{12},$$
$$a_{15} = \theta_1\theta_4, \quad a_{18} = 1-\theta_1,$$

$$a_{21} = \frac{\theta_5}{1+\theta_5\theta_7}, \qquad a_{23} = \frac{-\theta_5\theta_6}{1+\theta_5\theta_7},$$

$$a_{26} = \frac{\theta_5\theta_7}{1+\theta_5\theta_7}, \qquad a_{27} = \frac{-\theta_5\theta_8}{1+\theta_5\theta_7},$$

$$a_{29} = \frac{1}{1+\theta_5\theta_7} .$$

The **Y** and **Z** matrices are thus

$$\mathbf{Y} = (\mathbf{log\ x},\ \mathbf{log\ px})$$

and

$$\mathbf{Z} = (\mathbf{1},\ \mathbf{log\ pxw},\ \mathbf{log\ yw},\ \mathbf{log\ p},\ \mathbf{y^*},\ \mathbf{log\ x_{-1}},\ \mathbf{log\ px_{-1}})$$

where **1** is the unit vector.

Regarding identification, it can readily be seen that problems are bound to occur whenever θ_1, and/or θ_5 are close to zero, since several of the columns of $\partial \mathrm{vec}\mathbf{A}/\partial\boldsymbol{\theta}$ will then approach zero vectors. In fact, the latter parameter can cause trouble even without being particularly small, provided the product $\theta_5\theta_7$ is large.

Input

```
 DATAM.DAT
0.59333  4.34251 1.  4.42004  4.03954  4.30136  4.08933  0.50682  4.34251
0.72271  4.36437 1.  4.44030  4.15779  4.32413  4.16200  0.59333  4.34251
0.81978  4.37450 1.  4.45085  4.19192  4.35927  4.22975  0.72271  4.36437
0.92028  4.37450 1.  4.44969  4.25731  4.36945  4.29456  0.81978  4.37450
1.00063  4.39445 1.  4.45783  4.32744  4.37450  4.35543  0.92028  4.37450
1.10856  4.40672 1.  4.47050  4.44524  4.41764  4.41159  1.00063  4.39445
1.18173  4.43082 1.  4.48526  4.54287  4.45899  4.46591  1.10856  4.40672
1.26976  4.45435 1.  4.50756  4.63337  4.47392  4.51743  1.18173  4.43082
1.31909  4.48864 1.  4.51305  4.69583  4.46706  4.56539  1.26976  4.45435
1.39377  4.48864 1.  4.50976  4.83961  4.45202  4.61115  1.31909  4.48864
1.53687  4.52179 1.  4.54436  4.96375  4.51415  4.65586  1.39377  4.48864
1.65250  4.60517 1.  4.60517  5.04349  4.60517  4.69866  1.53687  4.52179
1.70656  4.65396 1.  4.65110  5.11883  4.65205  4.73883  1.65250  4.60517
1.73519  4.78749 1.  4.73180  5.20351  4.77576  4.77828  1.70656  4.65396
1.88555  4.96284 1.  4.89485  5.33383  4.97949  4.81624  1.73519  4.78749
1.95445  5.15329 1.  5.09006  5.43149  5.18122  4.85203  1.88555  4.96284
1.88555  5.34711 1.  5.20483  5.38564  5.34568  4.88734  1.95445  5.15329
1.90658  5.38907 1.  5.21330  5.48935  5.36317  4.92071  1.88555  5.34711
1.89912  5.45104 1.  5.29079  5.54869  5.42495  4.95371  1.90658  5.38907
1.95303  5.53733 1.  5.44415  5.58953  5.51423  4.98498  1.89912  5.45104
2.01357  5.69709 1.  5.57405  5.63679  5.67332  5.01595  1.95303  5.53733
1.96991  5.84064 1.  5.67195  5.69235  5.82393  5.04600  2.01357  5.69709
```

The 22 observations cover the period 1959-80. However, to make the estimations under the two alternative stochastic specifications comparable, the first observation will be disregarded when the residuals are assumed to be only contemporaneously correlated. (This is also required for the LR test of **H=O** to to be applicable, cf. sec. 2.4.)

```
     PARIC.DAT
8
0.30 1
-4.31 2
-3.30 3
1.22 4
0.70 5
3.77 7
-0.94 6
0.48 8
```

In this case, it is hard to think of interesting inequality constraints, relating to several parameters simultaneously. Consequently no use has been made of the hyperbolic transformations.

```
     GICOF.DAT
0 1.E-10 0 78 1
2 7 22
1
2 22 2 7
EXPORT SYSTEM ACC. TO GOLDSTEIN & KHAN (1978). ESTIMATION PERIOD 1960-80.
12 1 2
1 1 0
-1.0
1 3 1
0.0 0 1
1.0 1 2
1 4 1
0.0 0 1
-1.0 1 3
1 5 1
0.0 0 1
1.0 1 4
1 8 1
1.0 1
-1.0 1
2 1 2
0.0 1 0 1.0 0 1
1.0 5
1.0 5 7
2 2 0
-1.0
2 3 2
0.0 0 1 1.0 0 1
-1.0 5 6
1.0 5 7
2 6 2
0.0 0 1 1.0 0 1
1.0 5 7
1.0 5 7
2 7 2
0.0 0 1 1.0 0 1
-1.0 5 8
1.0 5 7
2 9 2
1.0 0 0 1.0 0 1
1.0 5 7
1 2 1
0.0 0 1
1.0 1 3
```

Notice that LP=0 since the first estimation is to be performed with **H=0.** Further, MXFN has been set to zero, in order to obtain a printout of the input in RESF1.DAT. As indicated by IFO (first number on the fourth line) the first row of the matrix in DATAM.DAT should be ignored.

The execution is initialized by the command /CONRAD (cf. section 3.4).

Results

RESF1.DAT

```
|=============================================================================|
| EXPORT SYSTEM ACC. TO GOLDSTEIN & KHAN (1978). ESTIMATION PERIOD 1960-80.   |
|=============================================================================|
```

Y-MATRIX

0.722710	4.364370
0.819780	4.374500
0.920280	4.374500
1.000630	4.394450
1.108560	4.406720
1.181730	4.430820
1.269760	4.454350
1.319090	4.488640
1.393770	4.488640
1.536870	4.521790
1.652500	4.605170
1.706560	4.653960
1.735190	4.787490
1.885550	4.962840
1.954450	5.153290
1.885550	5.347110
1.906580	5.389070
1.899120	5.451040
1.953030	5.537330
2.013570	5.697090
1.969910	5.840640

Z-MATRIX

1.000000	4.440300	4.157790	4.324130	4.162000	0.593330	4.342510
1.000000	4.450850	4.191920	4.359270	4.229750	0.722710	4.364370
1.000000	4.449690	4.257310	4.369450	4.294560	0.819780	4.374500
1.000000	4.457830	4.327440	4.374500	4.355430	0.920280	4.374500
1.000000	4.470500	4.445240	4.417640	4.411590	1.000630	4.394450
1.000000	4.485260	4.542870	4.458990	4.465910	1.108560	4.406720
1.000000	4.507560	4.633370	4.473920	4.517430	1.181730	4.430820
1.000000	4.513050	4.695830	4.467060	4.565390	1.269760	4.454350
1.000000	4.509760	4.839610	4.452020	4.611150	1.319090	4.488640
1.000000	4.544360	4.963750	4.514150	4.655860	1.393770	4.488640
1.000000	4.605170	5.043490	4.605170	4.698660	1.536870	4.521790
1.000000	4.651100	5.118830	4.652050	4.738830	1.652500	4.605170
1.000000	4.731800	5.203510	4.775760	4.778280	1.706560	4.653960
1.000000	4.894850	5.333830	4.979490	4.816240	1.735190	4.787490
1.000000	5.090060	5.431490	5.181220	4.852030	1.885550	4.962840
1.000000	5.204830	5.385640	5.345680	4.887340	1.954450	5.153290
1.000000	5.213300	5.489350	5.363170	4.920710	1.885550	5.347110
1.000000	5.290790	5.548690	5.424950	4.953710	1.906580	5.389070
1.000000	5.444150	5.589530	5.514230	4.984980	1.899120	5.451040
1.000000	5.574050	5.636790	5.673320	5.015950	1.953030	5.537330
1.000000	5.671950	5.692350	5.823930	5.046000	2.013570	5.697090

21 OSERVATIONS: NUMBER 2 TO 22

PARAMETERS, INEQUALITY CONSTRAINTS

INIT. VALUES	INDEX	TRANSF.	BOUND
0.300000	1	0	
-4.310000	2	0	
-3.300000	3	0	
1.220000	4	0	
0.700000	5	0	
-0.940000	6	0	
3.770000	7	0	
0.480000	8	0	

REQUIRED ACCURACY IN THE ESTIMATES: .10E-09

```
EQUALITY CONSTRAINTS

A( 1, 1) =   -1.00000    TYPE OF CONSTRAINT: 0

     R0 =  -1.00

A( 1, 3) =   -1.29300    TYPE OF CONSTRAINT: 1

     CONSTANTS GREEK R     1.00
     CORR. PAR.-INDICES    1  2

A( 1, 4) =    0.99000    TYPE OF CONSTRAINT: 1

     CONSTANTS GREEK R    -1.00
     CORR. PAR.-INDICES    1  3
A( 1, 5) =    0.36600    TYPE OF CONSTRAINT: 1

     CONSTANTS GREEK R     1.00
     CORR. PAR.-INDICES    1  4

A( 1, 8) =    0.70000    TYPE OF CONSTRAINT: 1

     R0 =   1.00

     CONSTANTS R          -1.00
     CORR. PAR.-INDEX       1

A( 2, 1) =    0.19236    TYPE OF CONSTRAINT: 2

     CONSTANTS R           1.00
     CORR. PAR.-INDEX       5

     S0 =   1.00

     CONSTANTS GREEK S     1.00
     CORR. PAR.-INDICES    5  7

A( 2, 2) =   -1.00000    TYPE OF CONSTRAINT: 0

     R0 =  -1.00

A( 2, 3) =    0.18082    TYPE OF CONSTRAINT: 2

     CONSTANTS GREEK R    -1.00
     CORR. PAR.-INDICES    5  6

     S0 =   1.00

     CONSTANTS GREEK S     1.00
     CORR. PAR.-INDICES    5  7

A( 2, 6) =    0.72520    TYPE OF CONSTRAINT: 2

     CONSTANTS GREEK R     1.00
     CORR. PAR.-INDICES    5  7

     S0 =   1.00

     CONSTANTS GREEK S     1.00
     CORR. PAR.-INDICES    5  7

A( 2, 7) =   -0.09233    TYPE OF CONSTRAINT: 2

     CONSTANTS GREEK R    -1.00
     CORR. PAR.-INDICES    5  8

     S0 =   1.00

     CONSTANTS GREEK S     1.00
     CORR. PAR.-INDICES    5  7

A( 2, 9) =    0.27480    TYPE OF CONSTRAINT: 2

     R0 =   1.00

     S0 =   1.00

     CONSTANTS GREEK S     1.00
     CORR. PAR.-INDICES    5  7

A( 1, 2) =   -0.99000    TYPE OF CONSTRAINT: 1

     CONSTANTS GREEK R     1.00
     CORR. PAR.-INDICES    1  3
```

RESF2.DAT

```
|=============================================================================|
| EXPORT SYSTEM ACC. TO GOLDSTEIN & KHAN (1978). ESTIMATION PERIOD 1960-80.   |
|=============================================================================|

IEXIT= 1     COND= 0.00051

F= -0.1639077E+03   LNDET(B)=  0.7642503E-01   LNDET(SIGMA)= -0.1545741E+02

LNDET((Y-YBAR)'(Y-YBAR))=  0.1638678E+01

YBAR

    1.5160     4.8440

MINIMUM AT

   0.430094  -3.482521  -1.844085   1.030875   0.409488  -3.988291   7.544305

   1.129218

GRADIENT

 -0.261E-08  0.174E-08  0.160E-08  0.166E-07  0.398E-09  0.445E-09  0.136E-10

  0.131E-08

STD. DEV.

   0.133503   0.621433   1.059768   0.136895   0.502542   2.327499  10.388072

   0.559935


HESIAN(-1)

  0.178232E-01   0.594090E-01   0.122313E+00  -0.138205E-01   0.445125E-01
  0.594090E-01   0.386179E+00   0.624069E+00  -0.848954E-01   0.212298E+00
  0.122313E+00   0.624069E+00   0.112311E+01  -0.139720E+00   0.390561E+00
 -0.138205E-01  -0.848954E-01  -0.139720E+00   0.187402E-01  -0.477737E-01
  0.445125E-01   0.212298E+00   0.390561E+00  -0.477737E-01   0.252548E+00
 -0.195525E+00  -0.885888E+00  -0.172579E+01   0.200573E+00  -0.111886E+01
 -0.842102E+00  -0.401861E+01  -0.738414E+01   0.904124E+00  -0.501254E+01
  0.474743E-01   0.216497E+00   0.418723E+00  -0.489658E-01   0.271793E+00

 -0.195525E+00  -0.842102E+00   0.474743E-01
 -0.885888E+00  -0.401861E+01   0.216497E+00
 -0.172579E+01  -0.738414E+01   0.418723E+00
  0.200573E+00   0.904124E+00  -0.489658E-01
 -0.111886E+01  -0.501254E+01   0.271793E+00
  0.541725E+01   0.221032E+02  -0.130229E+01
  0.221032E+02   0.107912E+03  -0.538755E+01
 -0.130229E+01  -0.538755E+01   0.313527E+00
```

A

-1.000000	-0.793131	-1.497813	0.793131	0.443373	0.000000	0.000000
0.100136	-1.000000	0.399373	0.000000	0.000000	0.755460	-0.113076

0.569906	0.000000
0.000000	0.244540

SIGMA

0.000898	-0.000260
-0.000260	0.000291

STRUCTURAL FORM EQ., NR.: 1

OBS.	Y	YP	RES.
2	0.72271	0.74401	-0.02130
3	0.81978	0.83321	-0.01343
4	0.92028	0.91660	0.00368
5	1.00063	0.99560	0.00503
6	1.10856	1.09394	0.01462
7	1.18173	1.19133	-0.00960
8	1.26976	1.27218	-0.00242
9	1.31909	1.32720	-0.00811
10	1.39377	1.41645	-0.02268
11	1.53687	1.51520	0.02167
12	1.65250	1.61421	0.03829
13	1.70656	1.71124	-0.00468
14	1.73519	1.73769	-0.00250
15	1.88555	1.80204	0.08351
16	1.95445	1.93480	0.01965
17	1.88555	1.89104	-0.00549
18	1.90658	1.87120	0.03538
19	1.89912	1.92180	-0.02268
20	1.95303	1.98885	-0.03582
21	2.01357	2.01685	-0.00328
22	1.96991	2.03978	-0.06987

(COS(P))**2	D-W
0.9948	1.4975

P=ANGLE BETWEEN (Y-YBAR) AND (YP-YPBAR)

STRUCTURAL FORM EQ., NR.: 2

OBS.	Y	YP	RES.
2	4.36437	4.32975	0.03462
3	4.37450	4.36370	0.01080
4	4.37450	4.37660	-0.00210
5	4.39445	4.38158	0.01287
6	4.40672	4.42351	-0.01679
7	4.43082	4.45893	-0.02811
8	4.45435	4.47909	-0.02474
9	4.48864	4.47918	0.00946
10	4.48864	4.47851	0.01013
11	4.52179	4.53472	-0.01293
12	4.60517	4.61832	-0.01315
13	4.65396	4.67500	-0.02104
14	4.78749	4.77880	0.00869
15	4.96284	4.97612	-0.01328
16	5.15329	5.17425	-0.02096
17	5.34711	5.33418	0.01293
18	5.38907	5.39312	-0.00405
19	5.45104	5.44558	0.00546
20	5.53733	5.53004	0.00729
21	5.69709	5.67389	0.02320
22	5.84064	5.81896	0.02168

(COS(P))**2	D-W
0.9989	1.1380

P=ANGLE BETWEEN (Y-YBAR) AND (YP-YPBAR)

-B(-1)C

-1.681056	0.734774	0.410751	-0.555092	0.083085	0.527973	-0.179682
0.231038	0.073578	0.041131	0.699875	-0.104756	0.052869	0.226548

OMEGA

0.001282	-0.000327
-0.000327	0.000213

REDUCED FORM EQ., NR.: 1

OBS.	Y	YP	RES.
2	0.72271	0.76788	-0.04517
3	0.81978	0.84015	-0.02037
4	0.92028	0.91533	0.00495
5	1.00063	1.00543	-0.00480
6	1.10856	1.08268	0.02588
7	1.18173	1.16997	0.01176
8	1.26976	1.25382	0.01594
9	1.31909	1.33355	-0.01446
10	1.39377	1.42223	-0.02846
11	1.53687	1.50730	0.02957
12	1.65250	1.60736	0.04514
13	1.70656	1.69544	0.01112
14	1.73519	1.74390	-0.00871
15	1.88555	1.79842	0.08713
16	1.95445	1.92084	0.03361
17	1.88555	1.90014	-0.01459
18	1.90658	1.87082	0.03576
19	1.89912	1.92415	-0.02503
20	1.95303	1.99157	-0.03854
21	2.01357	2.03365	-0.02008
22	1.96991	2.05056	-0.08065

(COS(P))**2	D-W
0.9926	1.2471

P=ANGLE BETWEEN (Y-YBAR) AND (YP-YPBAR)

REDUCED FORM EQ., NR.: 2

OBS.	Y	YP	RES.
2	4.36437	4.33427	0.03010
3	4.37450	4.36574	0.00876
4	4.37450	4.37610	-0.00160
5	4.39445	4.38206	0.01239
6	4.40672	4.42091	-0.01419
7	4.43082	4.45775	-0.02693
8	4.45435	4.47749	-0.02314
9	4.48864	4.48063	0.00801
10	4.48864	4.48136	0.00728
11	4.52179	4.53175	-0.00996
12	4.60517	4.61380	-0.00863
13	4.65396	4.67389	-0.01993
14	4.78749	4.77967	0.00782
15	4.96284	4.96740	-0.00456
16	5.15329	5.17089	-0.01760
17	5.34711	5.33564	0.01147
18	5.38907	5.38954	-0.00047
19	5.45104	5.44808	0.00296
20	5.53733	5.53390	0.00343
21	5.69709	5.67590	0.02119
22	5.84064	5.82704	0.01360

(COS(P))**2	D-W
0.9992	1.2325

P=ANGLE BETWEEN (Y-YBAR) AND (YP-YPBAR)

Concerning RESF1.DAT it can be seen that the first observation has been left out (cf. DATAM.DAT above), in accordance with the specification in GICOF.DAT. Notice also the reordering of the parameters relative PARIC.DAT.

To obtain RESF2.DAT the value of MXFN was changed to 100. The run required 3.04 seconds of CPU time.

Notice that the printout has been vertically partitioned with respect to several vectors and matrices, e.g. Θ and $\mathbf{A}$, because WIDTH was set to 78.

The notation "(-1)" has been used to denote inverse.

OMEGA is the covariance matrix of the reduced form residuals.

It might be remarked that the extremely high value of Berndt's generalized R^2 measure; 0.99999997, is not due to a misprinting of the formula in sec. 5.3. Rather, it has been caused by the combined effecs of a short time series and a general tendency of this measure to lie very close to unity. [R^2:s of almost the same magnitude are reported in Berndt and Khaled (1979, p.1235)].

The number of iterations and function evaluations needed to find the optimum are given in RESF3.DAT.

RESF3.DAT

```
|=============================================================================|
| EXPORT SYSTEM ACC. TO GOLDSTEIN & KHAN (1978). ESTIMATION PERIOD 1960-80.   |
|=============================================================================|
ENTRY TO VA09A

   0    1
-0.1411646E+03
 0.3000000E+00 -0.4310000E+01 -0.3300000E+01  0.1220000E+01  0.7000000E+00
-0.9400000E+00  0.3770000E+01  0.4800000E+00
 0.1098669E+01  0.2650563E+02  0.3334622E+01  0.1438580E+03  0.1737695E+02
-0.2710711E+02 -0.3121616E+01 -0.1448753E+03
   5   11
-0.1508769E+03
 0.3657473E+00 -0.4292283E+01 -0.3322863E+01  0.1202043E+01  0.2764811E+00
-0.1069310E+01  0.3855606E+01  0.5134900E+00
-0.8361313E+00 -0.6253816E+02 -0.6888012E+01 -0.3378299E+03 -0.1308631E+02
 0.4918533E+02 -0.1583253E+01  0.2160029E+03
  10   17
-0.1635536E+03
 0.3530550E+00 -0.3712914E+01 -0.2386145E+01  0.1086749E+01  0.3884849E+00
-0.3828899E+01  0.6689173E+01  0.1092958E+01
-0.1709415E+01  0.9547018E+01  0.3818401E+00  0.4838898E+02  0.3747190E+01
-0.8254307E+01 -0.2065499E+00 -0.3933032E+02
  15   22
-0.1639069E+03
 0.4312800E+00 -0.3476031E+01 -0.1833431E+01  0.1029451E+01  0.4090560E+00
-0.3993622E+01  0.7468016E+01  0.1130456E+01
 0.3667082E+00  0.5329087E+00 -0.2942269E-01  0.2218442E+01 -0.3617625E+00
-0.1130278E+00 -0.1217911E-01 -0.3962095E+00
  20   27
-0.1639077E+03
 0.4305119E+00 -0.3480557E+01 -0.1840486E+01  0.1030433E+01  0.4118111E+00
-0.3998439E+01  0.7496381E+01  0.1131690E+01
 0.3860476E-02  0.2827210E-02 -0.5938718E-03  0.9683600E-02  0.2183746E-02
 0.2658529E-02 -0.2633014E-03  0.1250141E-01
  25   32
-0.1639077E+03
 0.4303535E+00 -0.3481386E+01 -0.1841954E+01  0.1030620E+01  0.4107859E+00
-0.3994164E+01  0.7517233E+01  0.1130656E+01
 0.1028749E-01  0.1295723E-01 -0.6197918E-03  0.5885880E-01 -0.1938937E-01
 0.3971422E-01  0.1305091E-03  0.1883437E+00
  30   37
-0.1639077E+03
 0.4300943E+00 -0.3482522E+01 -0.1844087E+01  0.1030875E+01  0.4094866E+00
-0.3988286E+01  0.7544337E+01  0.1129217E+01
-0.8039693E-04 -0.2802129E-03 -0.9999877E-05 -0.1408005E-02 -0.2853085E-05
-0.4331575E-05  0.4162234E-06 -0.1381698E-04
  35   43    1
-0.1639077E+03
 0.4300944E+00 -0.3482521E+01 -0.1844085E+01  0.1030875E+01  0.4094880E+00
-0.3988291E+01  0.7544305E+01  0.1129218E+01
-0.2609705E-08  0.1737107E-08  0.1597867E-08  0.1656687E-07  0.3976537E-09
 0.4448386E-09  0.1363655E-10  0.1305095E-08
```

RESF4.DAT

```
 8
   0.430094     1       0
  -3.482521     2       0
  -1.844085     3       0
   1.030875     4       0
   0.409488     5       0
  -3.988291     6       0
   7.544305     7       0
   1.129218     8       0
```

As can be seen, the number of function evaluations amounted only to 43, quite far from the set limit of 100.

The second estimation differed from the first one by allowing for autocorrelated errors, i.e. LP was set equal to 1. Further, the first estimation was included in the sample - IFO = 1 - to get the runs comparable. Since, otherwise, all specifications were unchanged and the same initial values were used, only the first part of RESF1.DAT is reproduced below. Notice the last line, which indicates that the autoregressive specification has been chosen.

This time, the main execution took 6.42 seconds, i.e. slightly more than a doubling of the time used in the first run.

As a comparison, Hendry (1971, p. 263) reports a tripling of the execution time, for linear models. The difference is presumably mainly due to the fact that Hendry used an algorithm in which the first order derivatives of the log-likelihood function are numerically computed, whereas in CONRAD they are calculated analytically.

With more parameters to estimate, CONRAD would compare even more favourably, since the gain from the analytical calculation of the derivatives would then be greater (in relative terms).

RESF1.DAT

```
|=============================================================================|
| EXPORT SYSTEM ACC. TO GOLDSTEIN & KHAN (1978). ESTIMATION PERIOD 1960-80.   |
|=============================================================================|
```

Y-MATRIX

0.593330	4.342510
0.722710	4.364370
0.819780	4.374500
0.920280	4.374500
1.000630	4.394450
1.108560	4.406720
1.181730	4.430820
1.269760	4.454350
1.319090	4.488640
1.393770	4.488640
1.536870	4.521790
1.652500	4.605170
1.706560	4.653960
1.735190	4.787490
1.885550	4.962840
1.954450	5.153290
1.885550	5.347110
1.906580	5.389070
1.899120	5.451040
1.953030	5.537330
2.013570	5.697090
1.969910	5.840640

Z-MATRIX

1.000000	4.420040	4.039540	4.301360	4.089330	0.506820	4.342510
1.000000	4.440300	4.157790	4.324130	4.162000	0.593330	4.342510
1.000000	4.450850	4.191920	4.359270	4.229750	0.722710	4.364370
1.000000	4.449690	4.257310	4.369450	4.294560	0.819780	4.374500
1.000000	4.457830	4.327440	4.374500	4.355430	0.920280	4.374500
1.000000	4.470500	4.445240	4.417640	4.411590	1.000630	4.394450
1.000000	4.485260	4.542870	4.458990	4.465910	1.108560	4.406720
1.000000	4.507560	4.633370	4.473920	4.517430	1.181730	4.430820
1.000000	4.513050	4.695830	4.467060	4.565390	1.269760	4.454350
1.000000	4.509760	4.839610	4.452020	4.611150	1.319090	4.488640
1.000000	4.544360	4.963750	4.514150	4.655860	1.393770	4.488640
1.000000	4.605170	5.043490	4.605170	4.698660	1.536870	4.521790
1.000000	4.651100	5.118830	4.652050	4.738830	1.652500	4.605170
1.000000	4.731800	5.203510	4.775760	4.778280	1.706560	4.653960
1.000000	4.894850	5.333830	4.979490	4.816240	1.735190	4.787490
1.000000	5.090060	5.431490	5.181220	4.852030	1.885550	4.962840
1.000000	5.204830	5.385640	5.345680	4.887340	1.954450	5.153290
1.000000	5.213300	5.489350	5.363170	4.920710	1.885550	5.347110
1.000000	5.290790	5.548690	5.424950	4.953710	1.906580	5.389070
1.000000	5.444150	5.589530	5.514230	4.984980	1.899120	5.451040
1.000000	5.574050	5.636790	5.673320	5.015950	1.953030	5.537330
1.000000	5.671950	5.692350	5.823930	5.046000	2.013570	5.697090

22 OBSERVATIONS ON THE VARIABLES: NUMBER 1 TO 22,
CORRESPONDING TO 21 OBSERVATIONS ON THE AUTOREGRESSIVE MODEL

RESF2.DAT

```
|=============================================================================|
| EXPORT SYSTEM ACC. TO GOLDSTEIN & KHAN (1978). ESTIMATION PERIOD 1960-80.   |
|=============================================================================|
```

IEXIT= 1 COND= 0.00591

F= -0.1711345E+03 LNDET(B)= 0.1601129E+00 LNDET(SIGMA)= -0.1597830E+02

LNDET((Y-YBAR)'(Y-YBAR))= 0.1638678E+01

YBAR

1.5160 4.8440

MINIMUM AT

0.425328 -3.006924 -1.408521 0.933795 1.356911 -4.591157 2.713114

1.293701

GRADIENT

-0.310E-07 -0.135E-06 -0.141E-07 -0.706E-06 -0.190E-08 0.617E-07 0.162E-08

0.296E-06

STD. DEV.

0.103382 0.438081 0.470797 0.092731 0.588959 0.819553 1.154950

0.174010

HESIAN(-1)

0.106878E-01	0.216557E-01	0.409893E-01	-0.529282E-02	-0.267388E-01
0.216557E-01	0.191915E+00	0.152373E+00	-0.404077E-01	0.412561E-01
0.409893E-01	0.152373E+00	0.221649E+00	-0.344897E-01	-0.607983E-01
-0.529282E-02	-0.404077E-01	-0.344897E-01	0.859898E-02	-0.595164E-02
-0.267388E-01	0.412561E-01	-0.607983E-01	-0.595164E-02	0.346873E+00
0.349962E-01	0.279195E+00	0.160178E+00	-0.585239E-01	-0.524212E-01
0.132461E-01	-0.130982E+00	-0.117543E-01	0.251947E-01	-0.410211E+00
-0.765867E-02	-0.582016E-01	-0.343495E-01	0.122693E-01	0.150463E-01

0.349962E-01	0.132461E-01	-0.765867E-02
0.279195E+00	-0.130982E+00	-0.582016E-01
0.160178E+00	-0.117543E-01	-0.343495E-01
-0.585239E-01	0.251947E-01	0.122693E-01
-0.524212E-01	-0.410211E+00	0.150463E-01
0.671668E+00	0.293478E-01	-0.142335E+00
0.293478E-01	0.133391E+01	-0.124724E-01
-0.142335E+00	-0.124724E-01	0.302794E-01

A

-1.000000	-0.599084	-1.278929	0.599084	0.397169	0.000000	0.000000
0.289848	-1.000000	1.330739	0.000000	0.000000	0.786391	-0.374977

0.574672	0.000000
0.000000	0.213609

H

0.084911	-0.265410
-0.461199	0.220157

HA

-0.161839	0.214541	-0.461786	0.050869	0.033724	-0.208716	0.099523
0.525011	0.056139	0.882812	-0.276297	-0.183174	0.173130	-0.082554

0.048796	-0.056694
-0.265038	0.047028

SIGMA

```
   0.000918  -0.000492
  -0.000492   0.000389
```

STRUCTURAL FORM EQ., NR.: 1

OBS.	Y	YP	RES.
2	0.72271	0.73621	-0.01350
3	0.81978	0.82890	-0.00912
4	0.92028	0.91733	0.00295
5	1.00063	1.00149	-0.00086
6	1.10856	1.09073	0.01783
7	1.18173	1.19678	-0.01505
8	1.26976	1.27531	-0.00555
9	1.31909	1.33425	-0.01516
10	1.39377	1.40681	-0.01304
11	1.53687	1.49915	0.03772
12	1.65250	1.61451	0.03799
13	1.70656	1.71384	-0.00728
14	1.73519	1.74517	-0.00998
15	1.88555	1.79706	0.08849
16	1.95445	1.94405	0.01040
17	1.88555	1.91629	-0.03074
18	1.90658	1.87904	0.02754
19	1.89912	1.92915	-0.03003
20	1.95303	1.97178	-0.01875
21	2.01357	2.00407	0.00950
22	1.96991	2.03329	-0.06338

(COS(P))**2	D-W
0.9947	1.9128

P=ANGLE BETWEEN (Y-YBAR) AND (YP-YPBAR)

STRUCTURAL FORM EQ., NR.: 2

OBS.	Y	YP	RES.
2	4.36437	4.35370	0.01067
3	4.37450	4.37183	0.00267
4	4.37450	4.37723	-0.00273
5	4.39445	4.36940	0.02505
6	4.40672	4.42039	-0.01367
7	4.43082	4.44161	-0.01079
8	4.45435	4.47225	-0.01790
9	4.48864	4.46332	0.02532
10	4.48864	4.47581	0.01283
11	4.52179	4.55344	-0.03165
12	4.60517	4.61761	-0.01244
13	4.65396	4.66305	-0.00909
14	4.78749	4.78156	0.00593
15	4.96284	5.00879	-0.04595
16	5.15329	5.16075	-0.00746
17	5.34711	5.32214	0.02497
18	5.38907	5.40323	-0.01416
19	5.45104	5.42653	0.02451
20	5.53733	5.54222	-0.00489
21	5.69709	5.69064	0.00645
22	5.84064	5.80832	0.03232

(COS(P))**2	D-W
0.9985	2.0761

P=ANGLE BETWEEN (Y-YBAR) AND (YP-YPBAR)

-B(-1)C

```
  -1.768981   0.510448   0.338407  -0.401412   0.191406   0.489648  -0.109036
   0.818003   0.147952   0.098087   0.670043  -0.319498   0.141924   0.182005
```

B(-1)HB

```
   0.405885  -0.154143
  -0.407366  -0.100817
```

B(-1)HC

0.844093	-0.184377	-0.122235	0.266210	-0.126937	-0.176864	0.072311
-0.638153	0.222855	0.147744	-0.095969	0.045761	0.213774	-0.026068

OMEGA

0.001195	-0.000271
-0.000271	0.000131

REDUCED FORM EQ., NR.: 1

OBS.	Y	YP	RES.
2	0.72271	0.73966	-0.01695
3	0.81978	0.82892	-0.00914
4	0.92028	0.91637	0.00391
5	1.00063	1.01415	-0.01352
6	1.10856	1.08639	0.02217
7	1.18173	1.18904	-0.00731
8	1.26976	1.26536	0.00440
9	1.31909	1.34493	-0.02584
10	1.39377	1.41143	-0.01766
11	1.53687	1.48857	0.04830
12	1.65250	1.61378	0.03872
13	1.70656	1.70812	-0.00156
14	1.73519	1.74672	-0.01153
15	1.88555	1.78669	0.09886
16	1.95445	1.94178	0.01267
17	1.88555	1.92449	-0.03894
18	1.90658	1.87589	0.03069
19	1.89912	1.93721	-0.03809
20	1.95303	1.96651	-0.01348
21	2.01357	2.00877	0.00480
22	1.96991	2.04041	-0.07050

(COS(P))**2	D-W
0.9931	1.9295

P=ANGLE BETWEEN (Y-YBAR) AND (YP-YPBAR)

REDUCED FORM EQ., NR.: 2

OBS.	Y	YP	RES.
2	4.36437	4.35861	0.00576
3	4.37450	4.37447	0.00003
4	4.37450	4.37610	-0.00160
5	4.39445	4.37332	0.02113
6	4.40672	4.41396	-0.00724
7	4.43082	4.44373	-0.01291
8	4.45435	4.47097	-0.01662
9	4.48864	4.47081	0.01783
10	4.48864	4.48093	0.00771
11	4.52179	4.53945	-0.01766
12	4.60517	4.60638	-0.00121
13	4.65396	4.66351	-0.00955
14	4.78749	4.78490	0.00259
15	4.96284	4.98014	-0.01730
16	5.15329	5.15707	-0.00378
17	5.34711	5.33343	0.01368
18	5.38907	5.39433	-0.00526
19	5.45104	5.43757	0.01347
20	5.53733	5.54613	-0.00880
21	5.69709	5.68925	0.00784
22	5.84064	5.82875	0.01189

(COS(P))**2	D-W
0.9995	2.2513

P=ANGLE BETWEEN (Y-YBAR) AND (YP-YPBAR)

EIGENVALUES OF THE H MATRIX

REAL PART	IMAGINARY
-0.203808E+00	0.000000E+00
0.508876E+00	0.000000E+00

The values of COND and the first order derivatives indicate that this solution is somewhat less satisfactory than the previous one. This deterioration will always occur when the more general stochastic specification is substituted for the more restrictive one.

As shown at the far end of the file, both the eigenvalues of **H** lie within the unit circle (cf. sec. 2.1).

The LR test statistic for testing the hypothesis that **H=0** is found to be 14.41 [cf. (12) in sec. 2.4]. Since the critical value at the 1 percent significance level is 13.28, the hypothesis is decisively rejected.

According to RESF3.DAT (not reproduced) the number of function evaluations increased by 50 percent, relative to the first estimation. As an indication of the increase in general, this is somewhat on the low side. The increase is seldom higher than 60-70 percent, however.

REFERENCES

Belsley, D.A., 1980, "On the Efficient Computation of the Nonlinear Full-Information Maximum-Likelihood Estimator", Journal of Econometrics, 14, pp. 203-225.

Berndt, E.R., Hall, B.H., Hall, R.E. and Hausman, J.H., 1974, "Estimation and Inference in Nonlinear Structural Models", Annals of Economic and Social Measurement, 3, pp. 653-666.

Berndt, E.R. and Khaled, M.S., 1979, "Parameter Productivity Measurement and Choice Among Flexible Functional Forms", Journal of Political Economy, 81, pp. 1220-1245.

Berndt, E.R. and Savin, N.E., 1977, "Conflict Among Criteria for Testing Hypotheses in the Multivariate Linear Regression Model", Econometrics, 45, pp. 1263-1278.

Breusch, T.S., 1979, "Conflict Among Criteria for Testing Hypothesis: Extensions and Comments", Econometrica, 47, pp. 203-207.

Brown, B.W., 1981, "Sample Size Requirements in Full Information Maximum Likelihood Estimation", International Economic Review, 22, pp. 443-460.

Calzolari, G. and Panattoni, L., 1983, "Hessian and Approximated Hessian Matrices in Maximum Likelihood Estimation: A Monte Carlo Study", European Meeting of the Econometric Society, Pisa, Italy.

Carter, R.A.L. and Nagar, A.L., 1977, "Coefficients of Correlation for Simultaneous Equation Systems", Journal of Econometrics, 6, pp. 39-50.

Durbin, J., 1957, "Testing for Serial Correlation in Systems of Simultaneous Regression Equations", Biometrika, 44, pp. 370-377.

Durbin, J., 1970, "Testing for Serial Correlation in Least Squares Regression When Some of the Regressors Are Lagged Dependent Variables", Econometrica, 38, pp. 410-421.

Durbin, J. and Watson, G.S., 1951 and 1951, "Testing for Serial Correlation in Least Squares Regression", Biometrika, 37, pp. 409-428 and 38, pp. 159-178.

Engle, R.F., 1984, "Wald, Likelihood Ratio and Lagrange Multiplier Tests in Econometrics", in: Z. Griliches and M.D. Intrilligator (eds.) Handbook of Econometrics, Vol. II, Amsterdam, North-Holland.

Farebrother, R.W., 1980, "The Durbin-Watson Test for Serial Correlation when There Is No Intercept in the Regression", Econometrica, 48, pp. 1553-1563.

Fischer, F.M., 1966, The Identification Problem In Economics, New York, McGraw-Hill.

Fletcher, R., 1970, "A New Approach to Variable Metric Algorithms", Computer Journal, 13, pp. 317-322.

Fletcher, R. and Powell, M.J.D., 1963, "A Rapidly Convergent Descent Method for Minimization", Computer Journal, pp. 163-168.

Goldstein, M. and Khan, M.S., 1978, "The Supply and Demand for Exports: A Simultaneous Approach", Review of Economics and Statistics, LX, pp. 275-286.

Gouriéroux, C., Holly, A. and Monfort, A., 1982, "Likelihood Ratio Test, Wald Test and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameter", Econometrica, 50, pp. 63-80.

Greenberg, E. and Webster, Jr, C.E., 1983, Advanced Econometrics: A Bridge to the Literature New York, John Wiley and Sons.

Haessel, W., 1978, "Measuring Goodness of Fit in Linear and Nonlinear Models", Southern Economic Journal, 44, pp. 648-652.

Hausman, J.A., 1983, "Specification and Estimation of Simultaneous Equations Models", in: Z.Griliches and M.D. Intrilligator (eds.), Handbook of Econometrics, Vol I, Amsterdam, Norh-Holland.

Hendry, D.F., 1971, "Maximum Likelihood Estimation of Systems of Simultaneous Regression Equations with Errors Generated by a Vector Autoregressive Process", International Economic Review, 12, pp. 257-271.

Hendry, D.F., 1974, Maximum Likelihood Estimation of Systems of Simultaneous Regression Equations with Errors Generated by a Vector Autoregressive Process: A Correction", International Economic Review, 15, p. 260.

Hendry, D.F., 1976, "The Structure of Simultaneous Equations Estimators", Journal of Econometrics, 4, pp. 51-78.

Jansson, L. and Mellander, E., 1983, "Optimization under Non-linear Constraints", Economic Letters, 11, pp. 343-346.

Judge, G.G., Griffiths, W.E., Hill, R.C. and Lee, T.C., 1980, The Theory and Practice of Econometrics, John Wiley and Sons, New York.

Lau, L.J., 1978, "Testing and Imposing Monotonicity, Convexity and Quasi-Convexity Constraints", in M. Fuss and D. McFadden (eds.) Production Economics: A Dual Approach to Theory and Applications, Vol. I, Amsterdam, North-Holland.

Maddala, G.S., 1978, Econometrics, New York, McGraw-Hill.

Magnus, J.R. and Neudecker, H., 1979, "The Commutation Matrix: Some Properties and Statistical Applications", Annals of Statistics, 7, pp. 381-394.

Mellander, E., 1984, "Härledning av informationsmatrisen när residualerna genereras av en vektorautoregressiv process och parametrarna lyder under icke-lineära restriktioner" (Derivation of the Information matrix when the errors are generated by a vector autoregressive process and the parameters obey non-linear constraints.) Mimeographed.

Neudecker, H., 1969, "Some Theorems on Matrix Differentiation with Special Reference to Kronecker Products", Journal of the American Statistical Association, 64, pp. 953-963.

Pollock, D.S.G., 1979, The Algebra of Econometrics Chichester, John Wiley and Sons.

Quandt, R.E., 1983, "Computational Problems and Methods", in: Z. Griliches and M.D. Intrilligator (eds.), Handbook of Econometrics, Vol I, Amsterdam, North-Holland.

Rothenberg, T.J., 1971, "Identification in Parametric Models", Econometrica, 39, pp. 577-591.

Silvey, S.D., 1975, Statistical Inference, London, Chapman and Hall.

Theil, H., 1983, "Linear Algebra and Matrix Methods in Econmetrics", in: Z. Griliches and M.D. Intrilligator (eds.), Handbook of Econometrics, Vol I, Amsterdam, North-Holland.

Wonnacott, R.J. and Wonnacott, T.H., 1970, Econometrics, New York, John Wiley and Sons.

WORKING PAPERS (Missing numbers indicate publication elsewhere)

1976

1. Corporate and Personal Taxation and the Growing Firm
by Ulf Jakobsson

8. Estimation and Analysis with a WDI Production Function
by Göran Eriksson, Ulf Jakobsson and Leif Jansson

1977

15. Pay Differentials between Government and Private Sector Employees in Sweden
by Siv Gustafsson

1980

25. On Unexplained Price Differences
by Bo Axell

34. Imperfect Information Equilibrium, Existence, Configuration and Stability
by Bo Axell

1981

36. Energi, stabilitet och tillväxt i svensk ekonomi (Energy, Stability and Growth in the Swedish Economy)
by Bengt-Christer Ysander

38. Utiliy in Local Government Budgeting
by Bengt-Christer Ysander

40. Wage Earners Funds and Rational Expectations
by Bo Axell

42. The Structure of the ISAC Model
by Leif Jansson, Tomas Nordström and Bengt-Christer Ysander

43. An Econometric Model of Local Government and Budgeting
by Bengt-Christer Ysander

44. Local Authorities, Economic Stability and the Efficiency of Fiscal Policy
by Tomas Nordström and Bengt-Christer Ysander

45. Growth, Exit and Entry of Firms
by Göran Eriksson

56. Central Control of the Local Government Sector in Sweden
by Richard Murray

59. Longitudinal Lessons from the Panel Study of Income Dynamics
by Greg J. Duncan and James N. Morgan

1982

61. Var står den nationalekonomiska centralteorin idag?
av Bo Axell

63. General Search Market Equilibrium
by James W. Albrecht and Bo Axell
General Equilibrium without an Auctioneer
by James W. Albrecht, Bo Axell and Harald Lang

64. The Structure and Working of the ISAC Model
by Leif Jansson, Thomas Nordström and Bengt-Christer Ysander

65. Comparative Advantage and Development Policy Twenty Years Later
by Anne O. Krueger

67. Computable Multi-Country Models of Production and Trade
by James M. Henderson

69. Relative Competitiveness of Foreign Subsidiary Operations of a Multinational Company 1962-77
by Anders Grufman

71. Technology, Pricing and Investment in Telecommunications
by Tomas Pousette

72. The Micro Initialization of MOSES
by James W Albrecht and Thomas Lindberg

75. The MOSES Manual
by Fredrik Bergholm

76. Differential Patterns of Unemployment in Sweden
by Linda Leighton and Siv Gustafsson

77. Household Market and a Nonmarket Activities (HUS) - A Pilot Study
by Anders Klevmarken

1983

79. Energy Prices, Industrial Structure and Choice of Technology; An International Comparison with Special Emphasis on the Cement Industry
by Bo Carlsson

81. ELIAS - A Model of Multisectoral Economic Growth in a Small Open Economy
by Lars Bergman

84. Monopoly and Allocative Efficiency with Stochastic Demand
by Tomas Pousette

86. The Micro (Firm) Foundations of Industrial Policy
by Gunnar Eliasson

87. Excessive Government Spending in the U.S.: Facts and Theories
by Edward M. Gramlich

88. Control of Local Authority Expenditure - The Use of Cash Limits
by Noel Hepworth

89. Fiscal Containment and Local Government Finance in the U.K.
by Peter Jackson

90. Fiscal Limitations: An Assessment of the U.S. Experience
by Wallace E. Oates

91. Pricing and Privatization of Public Services
by George E. Peterson

94. Job Mobility and Wage Growth:
A Study of Selection Rules and Rewards
by Bertil Holmlund

96. The Machine Tool Industry - Problems and Prospects in an International Perspective
by Bo Carlsson

97. The Development and Use of Machine Tools in Historical Perspective
by Bo Carlsson

99. An Equilibrium Model of Search Unemployment
by James W. Albrecht and Bo Axell

100. Quit Behavior under Imperfect Information: Searching, Moving, Learning
by Bertil Holmlund and Harald Lang

102. Collecting Data for Micro Analysis.
Experiences from the Hus-Pilot Study
by N. Anders Klevmarken

103. The Swedish Micro-to-Macro Model
- Idea, Design and Application
by Gunnar Eliasson

105. Estimation of Wage Gains and Welfare Gains from
Self-Selection Models
by Anders Björklund and Robert Moffitt

106. Public Policy Evaluation in Sweden
by Bengt-Christer Ysander

108. Entry, Industry Growth and the Microdynamics
of Industry Supply
by John C. Hause and Gunnar Du Rietz

112. Schumpeterian Competition in Alternative Technological
Regimes
by Sidney Winter

113. The Industrial Finance Systems; Europe, U.S. and Japan
by Tad Rybczynski

117. Micro Heterogeneity of Firms and the Stability of Industrial
Growth
by Gunnar Eliasson

118. The Initialization Process - The Moses Manual, Part 2
by Fredrik Bergholm

1984

119. Forestalling the Demise of Empirical Economics: The Role
of Microdata in Labor Economics Research
by Frank Stafford

120. Unemployment and Mental Health - Some Evidence from
Panel Data
by Anders Björklund

121. Comparative Keynesian Dynamics
by Richard H. Day

122. The Firm and Financial Markets in the Swedish
Micro-to-Macro Model (MOSES)
- Theory, Model and Verification
by Gunnar Eliasson

123. What Can Be Learned from the Chrysler Experience?
by Frank Stafford

125. Estimates of Bounds for Export Price Elasticities in an Underidentified Simultaneous Two Equation System by Eva Christina Horwitz

126. An Easy Fixup of the Flaw in Lucas' "Neutrality of Money" by Harald Lang

127. Comment on Feldstein and Poterba's "Unemployment Insurance and Reservation Wages" by Harald Lang

128. Information Criterion and Estimation of Misspecified Qualitative Choice Models by David Brownstone

129. Organisation och styrsystem i ett modernt storföretag av Gunnar Eliasson

130. Fertility and the Opportunities of Children: Evidence on Quantity-Quality Interactions in Sweden by Bertil Holmlund

131. CONRAD - A Maximum Likelihood Program for Estimation of Simultaneous Equations Models that are Non-Linear in the Parameters by Leif Jansson and Erik Mellander