Chapter 7
ESTIMATION OF NONLINEAR SYSTEMS OF EQUATIONS

In Chapter 5, you learned how to estimate various types of linear single equation models in TSP. These models were specified implicitly by listing the dependent variable and independent variables after the name of the estimation method (OLSQ, 2SLS, LIML, or AR1). Although this shorthand method of specifying a model is convenient, you have to be more specific about the form of your models in order to estimate nonlinear equations.

Two procedures in TSP estimate general nonlinear regression models: LSQ and FIML. LSQ is a minimum distance estimator that can be used to compute nonlinear single equation least squares, nonlinear two-stage least squares, nonlinear multivariate regression, SUR (seemingly unrelated regressions), nonlinear three-stage least squares, and some Generalized Method of Moments (GMM) estimators.

FIML obtains full information maximum likelihood estimates for a nonlinear simultaneous equation model whose disturbances are jointly normally distributed. For maximum likelihood on models with alternative distributions, use the ML procedure discussed in Chapter 9. Both LSQ and FIML can be used on linear as well as nonlinear models. Except for the single equation estimates, even the linear models are nonlinear from the estimating point of view, since the covariance of the residuals is a nonlinear function of the parameters.

In this chapter we describe how to specify the equations of a model to be estimated and then discuss estimation with LSQ and FIML. In Chapter 8 we discuss hypothesis testing with applications for LSQ and FIML, and in Chapter 10 we describe the minimization techniques and convergence options used in all the nonlinear procedures.

7.1. Specifying the model: FRML, FORM, IDENT, PARAM, CONST

The FRML statement is used to define TSP equations for estimation or other computationss. To use FRML, you supply an equation in algebraic form (as in GENR) except that it is preceded by a name given to the equation when it is stored. Equations are referred to by their names when estimated and can be printed with the PRINT command.

For example, FRMLs from the complete illustrative model at the end of Chapter 3 are

```
FRML CONSEQ, CONS=A+B*GNP ;
FRML INVEQ, I=LAMBDA*I(-1) + ALPHA*GNP/(DELTA+R) ;
FRML INTRSTEQ, R=D+F*(LOG(GNP)+LP-LM) ;
FRML PRICEQ, LP=P(-1)+PSI*(LP(-1)-(LP(-2))+PHI*LGNP+TREND*TIME+P0 ;
```

The rules for composing FRMLs are the same as for GENRs (see Chapter 3). The first FRML, CONSEQ, could be used to estimate a simple linear regression of CONS on GNP and a constant term by specifying two additional statements:

```
PARAM A,B;  LSQ CONSEQ;
```

There is an implied additive disturbance term during estimation with LSQ or FIML; the model estimated above is:

$$CONS_t = \alpha + \beta GNP_t + \epsilon_t$$

If you are using simple linear equations, the FORM command can be used to choose parameter names automatically. For example, the command

```
FORM CONSEQ CONS C GNP;
```

would form an equation and declare parameters; this is equivalent to the two statements

```
FRML CONSEQ CONS = CONSEQ0 + CONSEQ1*GNP;
PARAM CONSEQ0 CONSEQ1;
```

FRMLs can also be used to specify functions of estimated parameters for testing (ANALYZ, Section 8.7), to specify formulas for equation substitution (EQSUB, Section 9.5) or differentiation (DIFFER, see the *Reference Manual*), and to specify log-likelihood equations (ML, Section 9.5). When used in this way, the equations do not have an implied additive disturbance.

Logical expressions may be included in FRMLs, but be aware that this will generally introduce a finite number of points at which your equation will not be differentiable. You can estimate or simulate with such an equation, but gradient methods may have difficulty converging if you happen to land on one of those points. The usual proofs of consistency, efficiency, and asymptotic normality will not go through in this case. However, such equations can be convenient when performing model simulations, and you may find them useful.

There are two ways to define a TSP equation: with the dependent variable on the left-hand side as in a GENR (a normalized equation), or as an unnormalized expression with no equal sign and no left-hand variable. An example of the latter would be

```
FRML CONSEQ2, CONS-A-B*GNP ;
```

This equation specifies the same model as above, but as though it were written

$$\epsilon_t = CONS_t - \alpha - \beta GNP_t .$$

Unnormalized FRMLs are used by FIML, LSQ, and SIML to handle models that are nonlinear in the endogenous variables; for example, to specify an orthogonality condition for GMM estimation (see section 7.2.5). Here is an example of this kind of equation, where the dependent variable is a general expression:

$$y_t/q_t = \alpha + \epsilon_t$$

This equation should be defined as:

```
FRML EQR, Y/Q - A;
```

If it were written as FRML EQR Y/Q = A; , the = sign would be interpreted as a logical operator: FRML EQR, Y/Q .EQ. A; , which means something entirely different from what was intended.

The primary difference between using normalized and unnormalized equations in LSQ or FIML is that unnormalized equations do not have a well-defined left-hand side variable associated with them, so certain goodness-of-fit statistics such as R^2 cannot be computed.

ANALYZ and EQSUB also accept unnormalized equations but for different purposes. SOLVE will not accept unnormalized equations and returns an error message if it encounters one.

A special form of FRML (IDENT) is available to specify identities in a FIML model. These identities do not contain parameters to be estimated, nor do they have disturbances, but they are necessary to complete the model, that is, to insure that the model has a square Jacobian matrix (as many equations as there are endogenous variables). This condition is necessary for both FIML and SIML. (SIML actually treats FRMLs and IDENTs identically, but it may be useful to use IDENT statements instead of FRMLs for documentation purposes.) Identities are given exactly like FRMLs except that they begin with the word IDENT instead of FRML. For example, the illustrative model is completed by a single identity relating GNP and consumption:

```
IDENT GNPID GNP=CONS+I+G ;
```

Information about the symbols in a FRML is provided in separate PARAM and CONST statements, for parameters to be estimated and constants, respectively. It is usually desirable to provide plausible numerical values for the parameters as starting points for estimations. Constants must be assigned values before estimation. For example, in the illustrative model (see Chapter 3),

```
PARAM A,-18,B,.62,LAMBDA,.6,ALPHA,1,D,-20,F,8,PSI,.3,PHI,.1,TREND,-.002,P0,-.6 ;
CONST DELTA,15 ;
```

The parameters need not be in order and may be defined in several PARAM and CONST statements. The only difference between a PARAM and a CONST is that PARAMs are estimated and CONSTs are not. Once parameters have been declared in a PARAM statement, they retain their estimability even if they are given new values by SET or UNMAKE commands. Parameters also retain their values from the most recent estimation. Parameters can be fixed at their current values by declaring them in a CONST statement. This will revise them to be treated as fixed until they are specified on a new PARAM statement.

7.2. Nonlinear least squares: LSQ

LSQ estimates single and multiple equation linear and nonlinear regression models. Depending on the number of equations and the specification of instrumental variables, several different econometric estimators are available with the LSQ command. Each estimator has a slightly different objective function which is minimized by means of iterative methods for nonlinear models. Details on the iterative techniques are in Chapter 10.

7.2.1. Single equation least squares

If LSQ is supplied with the name of just one FRML, and no instruments are specified, single equation least squares is the resulting estimator. For example,

```
LSQ CONSEQ;
```

estimates the consumption equation described earlier, assuming that A and B have been declared as PARAMs, and the series CONS and GNP have been properly defined over the current sample. In this example, the model is linear, so LSQ does not iterate, and the results are exactly the same as the command

```
OLSQ CONS C GNP;
```

OLSQ is easier to use for this model if you only want single equation estimation, but using LSQ automatically sets up starting values for a subsequent multiple equation estimation (which requires specifying the FRMLs).

For nonlinear least squares, the objective function is the sum of squared residuals (SSR). Minimizing SSR is equivalent to maximizing the likelihood function if the error in the equation is additive and normally distributed. The iteration technique is Gauss's method: derivatives of the equation residual with respect to each parameter are formed analytically. The current residual is regressed on the derivatives, and the resulting regression coefficients are the proposed changes in the parameters. It is possible to show that these coefficients will be zero if and only if the current parameters are at a local minimum of SSR. If the model was linear and the parameters had zero starting values, the first iteration would involve regressing the dependent variable (which equals the current residual) on the independent variables (which equal the derivatives of the residual with respect to the parameters, with a change of sign). For general nonlinear models, both the current residual and the derivatives will be functions of the current parameter values, but the procedure works the same, iterating until the derivatives are orthogonal to the residuals and the SSR is minimized.

A simple example of a nonlinear equation is the direct estimation of an AR1 model:

```
FRML CONSAR1 CONS = A + B*GNP + RHO*(CONS(-1) - (A + B*GNP(-1)));
PARAM A,B,RHO;
LSQ CONSAR1;
```

Note that this method of estimating an AR(1) model drops the first observation, so that the AR1 procedure is more efficient than using LSQ in small samples. See Section 5.6 for an easier way to generate FRMLs for AR (p) estimation..

7.2.2. Multivariate regression and Seemingly Unrelated Regressions

If a model has two or more regression equations, it is likely that the disturbances from the two are correlated. If so, the technique of multivariate regression generally gives more efficient estimates than regression applied separately to each equation. Further, if two or more equations share the same parameter(s), they must be estimated jointly to impose these cross-equation constraints. This is particularly useful in estimating systems of demand equations derived from a utility function or a production function.

Multivariate regression is the simplest multiple equation estimator. It assumes there are no simultaneity problems with endogenous variables on the right-hand side of the equations. 3SLS and FIML are appropriate for joint estimation of equations with (or without) simultaneity; FIML will also estimate SUR models, often using less memory than LSQ.

An example of LSQ from the illustrative model (see Chapter 3) is

```
LSQ CONSEQ,INVEQ,INTRSTEQ,PRICEQ ;
```

This command specifies joint estimation of all four behavioral equations of the model.

The multivariate least squares method is a generalized least squares method: the disturbances of the model are assumed to be independent across observations, but to have free covariance across equations. A consistent estimate of this covariance matrix is formed in some way or supplied to the procedure, and this estimate is used to weight the observations when the equations are re-estimated. The objective function can be written as

$$Q(b) = e(b)' \, (S^{-1} \otimes I_T) \, e(b)$$

where e(b) is the vector of stacked residuals (a function of the parameters b), S is an estimated covariance matrix of the disturbances and I_T is the identity matrix of order of the number of observations. If S is recomputed from b(i) at each iteration, this estimator converges to the maximum likelihood estimator when the disturbances are multivariate normal.

Although any consistent estimator of S gives consistent parameter estimates for multivariate regression, the default method recomputes S at each iteration. In the case of demand systems, for example, this method yields estimates that are invariant with respect to which equation is dropped. However, S and b are not changed simultaneously in each iteration (as they are in FIML), so convergence is not guaranteed. Convergence may be especially difficult to obtain when the residuals are highly correlated and there are cross-equation constraints. FIML is recommended in this case.

If you want conventional Seemingly Unrelated Regressions estimates (not maximum likelihood), you can use SUR. SUR obtains a consistent estimate of S, and then iterates only on b until convergence is obtained. More precise control over the initial estimate of S and over iteration on S is possible with the WNAME= and MAXITW= LSQ options (explained in detail in the *Reference Manual*).

7.2.3. Nonlinear two-stage least squares: INST=

A nonlinear equation from a simultaneous model can be estimated by LSQ using a method developed by Amemiya (1974). The objective function for estimation is the sum of squared fitted residuals, where the fitted residuals are the fitted values from a regression of the true residuals on the instrumental variables. If the equation is linear in its parameters, this amounts to standard two-stage least squares, and LSQ will not iterate. If it is nonlinear, the estimates are consistent but not generally asymptotically efficient (relative to nonlinear three-stage least squares or FIML).

Nonlinear two-stage least squares in TSP is invoked by the INST= option. INST is followed by a list of instrumental variables in parentheses. For example, to estimate the investment function in the illustrative model (see Chapter 3),

```
LSQ(INST=(C,G,LM,TIME)) INVEQ ;
```

7.2.4. Linear or nonlinear three-stage least squares: 3SLS

Three-stage least squares is an instrumental variable method for estimating a system of simultaneous equations where there may be endogenous variables on the right-hand side as well as contemporaneous correlation of the disturbances. The advantage of 3SLS over FIML is that the model does not have to be completely specified; the estimates for the equations and parameters can be consistent even if the exact form of the rest of the model is unknown. For example, you may have a set of equations describing the quantities demanded of certain goods as a function of the prices of goods. Prices may be determined as part of a larger economy that you do not wish to model explicitly. With the choice of suitable instruments, you could estimate the demand equations consistently without specifying the complete model. FIML would require specification of the price equations.

To specify three-stage estimation use the INST= option and a list of equation names with 3SLS (or LSQ). For example, to estimate part of the illustrative model by three-stage least squares,

```
3SLS(INST=(C,G,LM,TIME)) CONSEQ,INTRSTEQ,PRICEQ;
```

The output from this example is shown on the following pages. (**Example 7.1**)

The options for three-stage least squares are the same as those for univariate and multivariate regression. To request iteration over the covariance matrix of the residuals, use the MAXITW= option (note: iteration has no demonstrated statistical value in this case nor is convergence guaranteed). MAXITW=0 is the default.

3SLS estimates a set of equations by the same technique described for nonlinear two stage least squares, but considers the covariances across equations as well. The criterion for estimation is the sum of squared transformed fitted residuals. For each observation, fitted residuals are formed as the fitted values from regressions on instrumental variables. These are transformed by multiplying by the square root of the covariance matrix of the residuals. The contribution of the observation to the criterion is then the sum of squared values of these transformed fitted residuals.

For further details on the properties of this estimator, see Jorgenson and Laffont (1974) and Amemiya (1977). The NL3SLS estimator discussed by Amemiya is slightly more general in its choice of instruments than Jorgenson-Laffont; TSP uses the form specified by the latter where the same set of instruments is used for all equations. See GMM and the MASK option in the *Reference Manual* for a way to specify different instruments for each equation. The method of estimation is described in Berndt, Hall, Hall, and Hausman (1975). If the model is linear in its parameters and variables, three-stage least squares estimates are asymptotically efficient.

The three-stage estimation criterion requires an estimate of the residual covariance matrix. TSP obtains this by carrying out an initial estimation with the covariance matrix set equal to an identity matrix. If there are no parameters in common among the equations, these initial estimates are just the two-stage estimates. Then the covariance matrix is estimated from the true (not fitted) residuals from the initial estimates. Unless the user specifies otherwise, this estimate of the covariance matrix will be held fixed while the parameters are re-estimated to obtain three-stage least squares estimates.

7.2.5. Generalized Method of Moments

3SLS coincides with the GMM estimator of Hansen (1982) when the errors are serially independent and the same instruments are used for each equation. In Hansen's notation, the GMM estimator sets the orthogonality conditions $u_t(b,y,X) \otimes z_t$ as close to zero as possible using the estimated variance of this vector as the metric. To perform this type of estimation in LSQ, define each element of the vector u as a normalized or unnormalized equation using FRML. The vector of z's are specified as instruments in the INST list. The nonlinear three stage least squares estimates obtained are consistent and asymptotically efficient, and are also numerically identical to those obtained by the corresponding GMM estimator (for the default NOHET and NMA=0 options).

As an example, consider the simplest version of the Hansen-Singleton model [Hansen and Singleton (1982); for a simple presentation of the Euler equation for this type of model, see Hall(1978)]: a consumption-based asset pricing model where investors have a utility function of the constant relative risk aversion form. Denote consumption in period t as

```
                         THREE STAGE LEAST SQUARES
                         =========================

EQUATIONS: CONSEQ INVEQ INTRSTEQ PRICEQ

INSTRUMENTS: C G LM TIME


                         OPTIONS FOR THIS ROUTINE
                         ========================

  COVOC    =                COVU     =                DEBUG    = FALSE
  HETERO   = FALSE          INST     = 0 0001         ITEROC   = FALSE
  ITERU    = FALSE          KERNEL   =                LSQSTART = FALSE
  MAXITW   = 0              NMA      = 0              ROBUST   = FALSE
  WNAME    =
MAXIMUM NUMBER OF ITERATIONS ON V-COV MATRIX OF RESIDUALS = 0

                                CONSTANTS:

                 DELTA
VALUE         15.00000

NOTE => The model is linear in the parameters.

                         OPTIONS FOR THIS ROUTINE
                         ========================

  GRADCHEC = FALSE          HCOV     = G              HITER    = G
  MARQUARD = FALSE          MAXIT    = 20             MAXSQZ   = 10
  METHOD   =                PRINT    = TRUE           SILENT   = FALSE
  SQZTOL   = 0.10000        STEP     = BARD           SYMMETRI = FALSE
  TOL      = 0.0010000      VERBOSE  = FALSE

Working space used: 4437
                              STARTING VALUES

                    B0            B1       LAMBDA        ALPHA           D0
VALUE        -23.20011       0.63999      0.62323      1.14676     -6.39086

                    D1           PSI          PHI        TREND           P0
VALUE          8.20009      -0.43942      0.63001    -0.018798     -3.85964



ITERATION NUMBER   1
====================
F=   3621.7     FNEW=   3561.9     ISQZ=  0 STEP=  1.0000     CRIT=  59.806

                    B0            B1       LAMBDA        ALPHA           D0
ESTIMATE     -23.20011       0.63999      0.62323      1.14676     -6.39086
CHANGES        3.48525    -0.0038124     0.071956     -0.21414     -0.10849

                    D1           PSI          PHI        TREND           P0
ESTIMATE       8.20009      -0.43942      0.63001    -0.018798     -3.85964
CHANGES       0.080951       1.40209     -0.66676     0.020838      4.07317

CONVERGENCE ACHIEVED AFTER   1 ITERATIONS

    2 FUNCTION EVALUATIONS.
```

Example 7.1: Three Stage Least Squares

```
END OF TWO STAGE LEAST SQUARES ITERATIONS (SIGMA=IDENTITY). THREE STAGE
LEAST SQUARES ESTIMATES WILL BE OBTAINED USING THIS ESTIMATE OF SIGMA:
                          RESIDUAL COVARIANCE MATRIX

                CONSEQ        INVEQ     INTRSTEQ       PRICEQ
CONSEQ       146.86262
INVEQ        -66.47420    185.26175
INTRSTEQ       1.42782     -0.25296      1.12389
PRICEQ       -0.077880      0.13425    0.0040156   0.00050308

                               WEIGHTING MATRIX

                CONSEQ        INVEQ     INTRSTEQ       PRICEQ
CONSEQ        0.082517     0.036336    -0.010321     0.014659
INVEQ                      0.080277   -0.0024068    -0.032191
INTRSTEQ                                 0.94958     -0.20911
PRICEQ                                               51.28644

Working space used: 4437


ITERATION NUMBER   1
====================
F=   33.574     FNEW=   33.029     ISQZ=  0 STEP=  1.0000     CRIT= 0.54463

                    B0           B1       LAMBDA        ALPHA           D0
ESTIMATE     -19.71487      0.63618      0.69519      0.93262     -6.49934
CHANGES       -1.84942    0.0020574     0.013003    -0.038817    -0.075091


                    D1          PSI          PHI        TREND           P0
ESTIMATE       8.28104      0.96267    -0.036751    0.0020400      0.21353
CHANGES       0.057388      0.15946    -0.053669    0.0014703      0.33005

CONVERGENCE ACHIEVED AFTER   1 ITERATIONS


    4 FUNCTION EVALUATIONS.



                           THREE STAGE LEAST SQUARES
                           =========================

                          RESIDUAL COVARIANCE MATRIX

                CONSEQ        INVEQ     INTRSTEQ       PRICEQ
CONSEQ       147.57391
INVEQ        -67.94115    186.02693
INTRSTEQ       1.40383     -0.47859      1.12388
PRICEQ        -0.11551      0.17869    0.0033926   0.00064413


                               WEIGHTING MATRIX

                CONSEQ        INVEQ     INTRSTEQ       PRICEQ
CONSEQ        0.082517     0.036336    -0.010321     0.014659
INVEQ                      0.080277   -0.0024068    -0.032191
INTRSTEQ                                 0.94958     -0.20911
PRICEQ                                               51.28644
```

Example 7.1: (Continued, page 2).

```
                  COVARIANCE MATRIX OF TRANSFORMED RESIDUALS

                CONSEQ        INVEQ     INTRSTEQ       PRICEQ
CONSEQ        27.13077
INVEQ         -0.20479     26.92743
INTRSTEQ     -0.059263     -0.46166     27.04051
PRICEQ        -4.15961      3.11028     -0.26982     31.87350

                                    E'HH'E =     33.0294
                    Number of Observations =     27

                         Standard
Parameter  Estimate        Error       t-statistic
B0         -21.5643      8.92636       -2.41580
B1         .638235       .010190       62.6309
LAMBDA     .708188       .157971       4.48301
ALPHA      .893806       .466509       1.91595
D0         -6.57444      1.18326       -5.55619
D1         8.33843       .895622       9.31021
PSI        1.12213       .319314       3.51421
PHI        -.090421      .194599       -.464651
TREND      .351033E-02   .673604E-02   .521127
P0         .543581       1.18972       .456899

Standard Errors computed from   quadratic form of analytic first
derivatives (Gauss)


                               Equation    CONSEQ
                               ==================

Dependent variable: CONS

 Mean of dependent variable = 519.033       Std. error of regression = 12.1480
Std. dev. of dependent var. = 147.459                      R-squared = .992999
   Sum of squared residuals = 3984.50        Durbin-Watson statistic = .466537
      Variance of residuals = 147.574

                                Equation    INVEQ
                                =================

Dependent variable: I

 Mean of dependent variable = 126.504       Std. error of regression = 13.6392
Std. dev. of dependent var. = 37.8154                      R-squared = .866469
   Sum of squared residuals = 5022.73        Durbin-Watson statistic = 1.52125
      Variance of residuals = 186.027

                              Equation    INTRSTEQ
                              ====================

Dependent variable: R

 Mean of dependent variable = 4.27741       Std. error of regression = 1.06013
Std. dev. of dependent var. = 2.21592                      R-squared = .762323
   Sum of squared residuals = 30.3448        Durbin-Watson statistic = 1.46411
      Variance of residuals = 1.12388

                               Equation    PRICEQ
                               ==================

Dependent variable: LP

 Mean of dependent variable = -.302748      Std. error of regression = .025380
Std. dev. of dependent var. = .233969                      R-squared = .987859
   Sum of squared residuals = .017392        Durbin-Watson statistic = 2.10372
      Variance of residuals = .644131E-03
```

Example 7.1: (Continued, page 3).

C_t and the one period return on asset j as x_{jt}. Then the representative agent model of intertemporal utility maximization implies the following population Euler equations in equilibrium:

$$E_t \{[\beta(C_t/C_{t-1})^\alpha x_{jt} - 1] z_{mt}\} = 0$$

where b is the discount rate and the z_{mt}, m=1,...,M are in the agent's information set at time t (they may include such things as lagged asset prices and consumption). In Hansen's notation u_t is the expression in the square brackets and the z's are the instruments.

This version of the Hansen-Singleton model can be easily estimated in the 3SLS procedure of TSP; the estimates coincide with GMM estimates, provided there is no serial correlation in the u's (which will be true if the assets in u are stocks or other one period assets). Here is how to set up the problem when there are two assets and you wish to use four lags as instruments:

```
LIST LAGXS X1(-1)-X1(-4) X2(-1)-X2(-4) ;
LIST UEQS U1EQ U2EQ ;
FRML U1EQ BETA*(CONS/CONS(-1))**ALPHA * X1(-1) - 1 ;
FRML U2EQ BETA*(CONS/CONS(-1))**ALPHA * X2(-1) - 1 ;
PARAM BETA 1 ALPHA -1 ;
GMM(HET,INST=LAGXS) UEQS ;
```

The preceding GMM example allows for conditional heteroskedasticity of the disturbances (the option HET for the GMM distance matrix), but not for serial correlation. If the asset returns in the previous example were based on multi-period rather than one-period returns, there is no reason to expect that the covariance of the marginal utility of consumption with these returns will not be correlated across the periods comprising the multi-period returns (see Hansen and Singleton, section 2 for a further discussion of this point). In this case the estimates obtained by GMM will be consistent but not asymptotically efficient, since they use the "wrong" covariance matrix of the orthogonality conditions as a weighting matrix. For this case, options for the GMM command allow you to compute the correct weighting matrix automatically.

Suppose that in the previous problem you wanted to use a covariance estimate that incorporates moving average disturbances of second order. You would use the NMA option to specify this:

```
GMM(HET,NMA=2,INST=LAGXS) UEQS ;
```

Among others, Newey and West (1987) pointed out that the original estimate proposed by Hansen and Singleton in this case is frequently not positive definite in finite samples and proposed the use of declining weights to guarantee positive semi-definiteness. TSP offers two choices of spectral density kernels (KERNEL=BARTLETT or PARZEN) to compute these weights; the default choice is BARTLETT.

7.3. Full information maximum likelihood: FIML

FIML is the asymptotically efficient estimator for linear and nonlinear simultaneous models, under the assumption that the disturbances are multivariate normal. When this assumption fails, FIML may still be *asymptotically* efficient; see White (1982) or Gourieroux, Montfort, and Trognon (1984) for a discussion of when this will be true.

Because FIML operates on the model as a whole, the model must be complete -- it must have as many equations as endogenous variables. Thus in addition to the behavioral equations containing unknown parameters, FIML must be supplied with any identities that involve the endogenous variables. Identities provide a convenient way of entering repeated sums and differences of endogenous variables into several equations; another way is the EQSUB command.

In FIML, the endogenous variables are listed in parentheses after the ENDOG= keyword. The corresponding instruments are then defined implicitly. For example, the illustrative model is estimated by FIML:

```
FIML(ENDOG=(CONS,I,R,LP,GNP)) CONSEQ,INVEQ,INTRSTEQ,PRICEQ,GNPID;
```

The objective function for FIML is the negative of the log likelihood function, which involves the log of the determinant of the residual covariance matrix, and the log of the determinant of the Jacobian (the derivatives of the residuals with respect to the endogenous variables). If there no simultaneity and no nonlinear functions of the endogenous variables appear in the equations, the Jacobian term drops out, and the model is equivalent to multivariate regression.

FIML can be used to estimate nonlinear LIML models (see Section 5.5 for linear LIML). In addition to the nonlinear equation of interest, FRMLs must be specified for each of the remaining endogenous variables. To make this a LIML model, these FRMLs are linear in the instruments and there are no constraints among their parameters. For example, to estimate the nonlinear AR1 equation of Section 7.2.1 with LIML:

```
FRML CONSAR1  CONS= A + B*GNP + RHO*(CONS(-1) - (A+B*GNP(-1));
PARAM A,B,RHO;
FORM GNPL GNP C G LM TIME CONS(-1) GNP(-1);          ? Equation for GNP as f (instruments)
LSQ(SILENT) CONSAR1;                                 ? To obtain starting values for parameters
LSQ(SILENT) GNPL;
FIML(ENDOG=(CONS,GNP)) CONSAR1,GNPL;                 ? Estimate LIML model
```

The standard errors for the FIML structural parameters are computed from the matrix of sums of squares of the outer products of the gradient of the likelihood function with respect to both the structural parameters and the unique elements of the inverse residual covariance matrix (the BHHH matrix). These standard errors are consistent and generally larger than those calculated in versions of TSP prior to 4.1, which were computed from the submatrix for the structural parameters only. For instance, the example above has nine structural parameters and three covariance parameters (NEQ*(NEQ+1)/2).

Note: Calzolari and Panattoni (1988) studied eight alternate FIML standard error formulas and demonstrated the consistency and good small-sample performance of the TSP Version 4.1 (and later) FIML standard errors. The 4.0 and earlier standard errors are not technically consistent, but they were shown to have reasonably good small-sample properties. Also, the R ("Gauss") matrix was shown to be inconsistent for nonlinear models. This matrix was used by TSP 4.0 when the number of parameters was larger than the number of observations (in which case the BHHH matrix will be singular, since its rank is less than its order). The R matrix is still used for iterations, but it is not available for standard errors because of this inconsistency. The best solution to the singularity of the BHHH matrix in small samples would be computation of the analytic second derivatives, and this will be available in future versions of TSP. If you can't wait for this, try doing FIML explicitly using the ML procedure (you will have to program the likelihood function of the multivariate normal simultaneous equations model).