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),