LSQ

LSQ (DEBUG, HETERO, INST=*list of instrumental variables*,ITERU,
COVU=OWN or *name of residual covariance matrix*,nonlinear options)
list of equation names ;

Function:

LSQ is used to obtain least squares or minimum distance estimates of one or more linear or nonlinear equations. These estimates may optionally be instrumental variables estimates as well. LSQ can compute nonlinear least squares, nonlinear two stage least squares or instrumental variables, nonlinear multivariate regression with cross-equation constraints, seemingly unrelated regression, and nonlinear three stage least squares. The equations for any of these estimators may be linear or nonlinear in the variables and parameters, and there may be arbitrary cross-equation constraints. LSQ can also be invoked with the SUR, 3SLS, THSLS, and GMM commands.

Usage:

There are four basic estimators available in LSQ: single or multi-equation least squares and single or multi-equation instrumental variables. They are all iterative methods which minimize a distance function of the general form

$$f(y,x,b)' \, [S^{-1} \otimes H(H'H)^{-1}H'] \, f(y,x,b)$$

where f(y,x,b) is the (stacked) vector of residuals from the nonlinear model, S is the current estimate of the residual covariance matrix being used as a weighting matrix, and H is a matrix of instruments.

The form of f(y,x,b) is specified by the user as a FRML (see entry in this manual), which may be either unnormalized (in the form f(y,x,b) with no = sign), or normalized (the usual form of y=f(x,b)). The latter form will cause equation by equation statistics of the explanatory success of the model to be printed.

To get a particular estimator, various assumptions are made about the exact form of this distance function. These are described below.

Nonlinear single equation least squares: In this case, there are no instruments (H is identity) and S is assumed to be unity. This makes the objective function the sum of squared residuals of the model; minimizing this function is the same as obtaining maximum likelihood estimates of the parameters of the model under the assumption of normality of the disturbances.

The form of the LSQ statement for estimating this model is LSQ followed by options in parentheses, and then the name of the equation. Any of the standard Nonlinear options can be used (see NONLINEAR in this manual for a complete description of these options).

Nonlinear two stage least squares: for this estimator, H is the matrix of instrumental variables formed from the variables in the INST option, and S is again assumed to be unity. The estimator is described in Amemiya (1974). If the model is linear, conventional two stage least squares or instrumental variable estimates result.

Nonlinear multivariate regression: In this case, there are no instruments (H is the identity matrix) and S is either estimated or fixed. This estimator can be computed with two completely different objective functions. The default in TSP is to compute maximum likelihood estimates if the LSQ command is specified with no instruments and more than one equation. These estimates are obtained by concentrating variance parameters out of the multivariate likelihood and then maximizing the negative of the log determinant of the residual covariance matrix. They are

efficient if the disturbances are multivariate normal and identically distributed.

Using the option MAXITW=0, it is possible to obtain minimum distance estimates of a nonlinear multivariate regression model. For these estimates, the objective function is the distance function given above with the instrument matrix H equal to identity. The S matrix is given by the WNAME option: it can be identity, which is similar to estimating each equation separately (except now cross-equation constraints are binding and the parameter standard errors will be wrong unless the true residual variances are unity), it can be supplied by you from a previous estimation, or it can be computed from the parameters (the WNAME=OWN option). The S matrix is always a symmetric matrix of the order of the number of equations.

To obtain conventional seemingly unrelated regression estimates of a nonlinear multivariate regression model, use the SUR command, which is a special form of the LSQ command. This version of the procedure obtains single equation estimates of the parameters of the model, uses these to form a consistent estimate of the residual covariance matrix, and then minimizes the objective function shown above with respect to the parameters b. If the model is linear, this is a two stage procedure (only two iterations). The plain LSQ command will iterate simultaneously on the parameters and the residual covariance matrix. In this case, linear models may take more than one iteration to converge.

Nonlinear three stage least squares: this estimator uses the distance function as shown, with S equal to a consistent estimate of the residual covariance (either supplied or computed), and H equal to the Kronecker product of an identity matrix of the order of the number of equations and the matrix of instruments. This means that all the instruments are used for all the equations.

Three stage least squares estimates can be obtained in two ways: LSQ with the WNAME option, the INST option, and more than one equation name will give three stage least squares estimates using the S matrix you specify. Alternatively, if you use the 3SLS form of the LSQ command with the INST option, LSQ automatically computes consistent nonlinear two stage least squares estimates of the parameters, uses them to form an estimate of the residual covariance matrix S, and then computes three stage least squares estimates.

To use any of these estimators, first specify the equations to be estimated using FRML statements and name the parameters and supply starting values with PARAM statements (an alternative to this is the FORM(PARAM) command after a linear estimation procedure). Any parameters which appear in more than one equation are assumed to be the same parameter and the equality constraint is automatically imposed.

LSQ always determines the linearity or nonlinearity of the model; if the model is linear in the parameters, it prints a message to that effect, and uses just one iteration.

Options:

COVU= residual covariance matrix (same as the old WNAME= option below).

DEBUG/**NODEBUG** specifies whether detailed computations of the model and its derivatives are to be printed out at every iteration. This option produces extremely voluminous output and is not recommended for use except by systems programmers maintaining TSP.

INST= (*list of instrumental variables*). If this option is included, the LSQ estimator becomes nonlinear two stage least squares or nonlinear IV (if there is one equation) and nonlinear three stage least squares (if there is more than one equation). The list of instrumental variables supplied is used for all the equations. See the INST section of this manual and the references for further information on the choice of instruments.

ITERU/**NOITERU** specifies iteration on the COVU matrix; provides the same function as the old MAXITW= option.

MAXITW= the number of iterations to be performed on the parameters of the residual covariance matrix estimate. If MAXITW is zero the covariance matrix of the residuals is held fixed at the initial estimate (which is specified by

WNAME). This option can be used to obtain estimates that are invariant to which equation is dropped in a shares model like translog.

HETERO/**NOHETERO** causes heteroscedastic-consistent standard errors to be used. See the GMM(NMA=) command for autocorrelation-consistent standard errors. Same as the old ROBUST option, or HCOV=R.

WNAME= the name of a matrix to be used as the starting value of the covariance matrix of the residuals.

WNAME=OWN specifies that the initial covariance matrix of the residuals is to be obtained from the residuals corresponding to the initial parameter values. If neither form of WNAME= is used, the initial covariance matrix is an identity matrix.

Nonlinear options control the iteration methods and printing. They are explained in the NONLINEAR section of this manual. Some of the common options are MAXIT, MAXSQZ, PRINT/**NOPRINT**, and SILENT/**NOSILENT**.

The only legal choice for HITER= is G (Gauss). HCOV=G is also the default method for calculating standard errors; R (Robust) is also a valid option. D is not allowed at present.

Examples:

Assume that the following equations have been specified for the illustrative model of the U.S. economy:

```
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 = LP(-1)+PSI*(LP(-1)-LP(-2))+PHI*LOG(GNP)+TREND*TIME+P0 ;
PARAM A B LAMBDA ALPHA D F PSI PHI TREND P0 ;
CONST DELTA 15 ;
```

The model as specified has four equations: the parameters to be estimated are A, B, LAMBDA, ALPHA, D, F, PSI, PHI, TREND, and P0. There are 7 variables in the model, CONS, GNP, I, R, LP, LM, and TIME, and one additional instrument, G. To estimate the investment equation by nonlinear least squares, use the following command:

```
LSQ (NOPRINT,TOL=.0001) INVEQ ;
```

We can obtain multivariate regression estimates of the whole model with the following command, although these estimates are probably not consistent due to the simultaneity of the model (there are endogenous variables on the right hand side of the equations):

```
LSQ (MAXIT=50) CONSEQ INVEQ PRICEQ INTRSTEQ ;
```

The example below obtains three stage least squares estimates of the model, using a weighting matrix based on the starting values of the parameters (which are obtained by nonlinear two stage least squares):

```
LSQ (INST=(C,LM,G,TIME)) CONSEQ ;
LSQ (INST=(C,LM,G,TIME)) INVEQ ;
LSQ (INST=(C,LM,G,TIME)) INTRSTEQ ;
LSQ (INST=(C,LM,G,TIME)) PRICEQ ;
LSQ (INST=(C,LM,G,TIME),WNAME=OWN) CONSEQ, INVEQ, INTRSTEQ, PRICEQ ;
```

You can get the same three stage least squares estimates without the intermediate two stage least squares printout by using this command:

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

LSQ

See the description of the LIST command for an example of using cross- equation restrictions.

Output:

LSQ stores its results in data storage. The estimated values of the parameters are stored under the parameter names. The fitted values and residuals will only be stored if the RESID option is on (the default). In addition, the following results are stored:

Name	Type	Length	Variable Description
@RNMS	list	#params	Parameter names.
@LOGL	scalar	1	Log of likelihood function (if valid).
@TR	scalar	1	Trace of COVT (if the minimum distance estimator is used).
@PHI	scalar	1	E'HH'E, the objective function for instrumental variable estimation.
@IFCONV	scalar	1	Convergence status (1 = success).
@GRAD	vector	#params	Gradient of objective function at the convergence.
@COEF	vector	#params	Estimated values of parameters,also stored under their names.
@SES	vector	#params	Standard Errors of estimated parameters.
@T	vector	#params	T-statistics.
@SSR	vector	#eqs	Sum of the squared residuals for each of the equations, stored in a vector.
@YMEAN	vector	#eqs	Mean of the dependent variable for each of the equations, stored in a vector.
@SDEV	vector	#eqs	Standard deviation of the dependent variable for each of the equations, stored in a vector.
@S	vector	#eqs	Standard error of each of the equations, stored in a vector.
@DW	vector	#eqs	Durbin-Watson statistic for each equation, stored in a vector.
@RSQ	vector	#eqs	R-squared for each equation.
@ARSQ	vector	#eqs	Adjusted R-squared for each equation.
@COVU	matrix	#eqs*#eqs	Residual covarince matrix.
@W	matrix	#eqs*#eqs	The inverse square root of COVU, the upper triangular weighting matrix.
@COVT	matrix	#eqs*#eqs	The covariance matrix of the transformed (weighted) residuals. This is equal to the number of obs. times the identity matrix if maximum likelihood is being used.
@VCOV	matrix	#par*#par	Estimated variance-covariance of estimated parameters.
@RES	matrix	#obs*#eqs	Residuals=actual - fitted values of the dependent variable.
@FIT	matrix	#obs* #choices	Fitted values of the dependent variables,stored as a matrix.

Normal LSQ output begins with a listing of the equations. The model is checked for linearity in the parameters (which simplifies the computations). A message is printed if linearity is found and LSQ does not iterate because it is unnecessary. The amount of working space used by LSQ is also printed - this number can be compared with the amount printed at the end of the run to see how much extra room you have if you wish to expand the model.

Next LSQ prints the values of constants and the starting conditions for the parameters, and then iteration-by-iteration output. If the print option is off, this output consists of only one line, showing the beginning value of the log likelihood, the ending value, the number of squeezes in the stepsize search (ISQZ), the final stepsize, and a criterion which should go to zero rapidly if the iterations are well-behaved. This criterion is the norm of the gradient in the metric of the Hessian approximation. It will be close to zero at convergence.

When the print option is on, LSQ also prints the value of the parameters at the beginning of the iteration and their direction vector. These are shown in a convenient table so that you can easily spot parameters with which you are

having difficulty.

Finally LSQ prints the results of the estimation (whether or not it converged); these results are printed no matter what print option is set (PRINT, NOPRINT, or SILENT). The names of the equations and endogenous variables are printed, the value of the objective function at the optimum, and the corresponding estimate of the covariance of the structural disturbances. If instrumental variable estimation was used, the objective function is labelled E'HH'E. If minimum distance estimation was used, the trace of the weighted residual covariance matrix is the objective function $(f(y,b)'[S^{-1} \otimes I]f(y,b))$. Otherwise the objective function is the negative of the log of the likelihood function.

Following this is a table of parameter estimates and asymptotic standard errors, as well as their estimated variance-covariance (unless it has been suppressed). For each equation, LSQ prints a few goodness-of-fit statistics: the sum of squared residuals, standard error, mean and standard deviation of the dependent variable, number of observations, and the Durbin- Watson statistic. The computation of these statistics is described in the regression output section of the User's Manual. If the equations are unnormalized, only the standard error, sum of squared residuals, and Durbin-Watson are printed.

Method:

The method used by LSQ is a generalized Gauss-Newton method. The Gauss-Newton method is Newton's method applied to a sum of squares problem where advantage is taken of the fact that the squared residuals are very small near the minimum of the objective function. This enables the Hessian of the objective function to be well approximated by the outer product of the gradient of the equations of the model. "Generalized" refers to the fact that the objective function contains a fixed weighting matrix also, rather than being a simple sum of squares.

This implementation of the Gauss-Newton method in TSP uses analytic first derivatives of the model, which implies that the estimating equations must be differentiable in the parameters (TSP defines the derivatives of discontinuous functions like SIGN() to be zero, so this will always be true). The method is one of the simplest and fastest for well-behaved equations where the starting values of the parameters are reasonably good. When the equation is highly nonlinear, or the parameters are far away from the answers, this method often has numerical difficulties, since it is fundamentally based on the local properties of the function. These problems are usually indicated by numerical error messages from TSP; the program tries to continue executing for a while, but if things do not improve, the estimation will be terminated. When you encounter a problem like this, you can often get around it by estimating only a few parameters at a time to obtain better starting values. Use CONST to fix the others at reasonable values.

For details on the estimation method, see the Berndt, Hall, Hall, and Hausman article.

References:

Amemiya, Takeshi, "The Nonlinear Two-Stage Least-Squares Estimator," **Journal of Econometrics,** July 1974, pp. 105-110.

Amemiya, Takeshi, "The Maximum Likelihood and the Nonlinear Three-Stage Least Squares Estimator in the General Nonlinear Simultaneous Equation Model," **Econometrica**, May 1977, pp. 955-966.

Berndt, E. K., B. H. Hall, R. E. Hall, and J. A. Hausman, "Estimation and Inference in Nonlinear Structural Models," **Annals of Economic and Social Measurement**, October 1974, pp. 653-665.

Chamberlain, Gary, "Multivariate Regression Models for Panel Data," **Journal of Econometrics 18,** 1982, pp. 5-46.

Jorgenson, Dale W. and Jean-Jacques Laffont, "Efficient Estimation of Nonlinear Simultaneous Equations with Additive Disturbances," **Annals of Economic and Social Measurement**, October 1974, pp. 615-640.

Judge et al, **The Theory and Practice of Econometrics,** 1980, John Wiley and Sons, New York, Chapter 7.

LSQ

Maddala, G. S., **Econometrics,** 1982, McGraw-Hill Book Co., New York, pp. 174- 175, 470-492.

Theil, Henri, **Principles of Econometrics,** John Wiley and Sons, New York, 1971, pp. 294-311.

White, Halbert, "Instrumental Variables Regression with Independent Observations," **Econometrica 50,** March 1982, pp. 483-500.

Zellner, Arnold, "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests of Aggregation Bias," **JASA 57,** (1962), pp. 348-368.