

(DROPMISS, GRADCHEC, GRADIENT=*method*, HCOV=*method*,
 HESSCHEC, HITER=*method*, MAXIT=# of iterations, MAXSQZ=# of squeezes,
 PRINT, SILENT, STEP=*squeeze method*, SQZTOL=*squeeze tolerance*, SYMMETRIC,
 TERSE, TOL=*parameter convergence tolerance*, TOLG=*gradient/CRIT convergence tolerance*,
 TOLS=*squeezed parameter convergence tolerance*, VERBOSE)

Function:

These options are common to all of TSP's nonlinear estimation and simulation procedures of ARCH, BJEST, FIML, LSQ, PROBIT, TOBIT, LOGIT, SAMPSEL, SIML, SOLVE, ML, etc. See Chapter 10 of the *User's Guide* for further information.

Usage:

Include these options among any other special options which you supply within parentheses after the name of the command which invokes the estimation procedure: FIML (ENDO=(...), nonlinear options) *list of eq names* ;

Options:

DROPMISS/NODROPMISS drop observations with any missing values before estimation. The NODROP option can be used in commands which use FRMLs, where the user is already processing missing values by multiplying such observations by zero dummy variables. For example, the FIML command could be used to estimate an exact ML AR(1) model, where the FRML includes a lagged residual for observations 2 to T. NODROP would prevent the first observation from being dropped (it does not use the lagged residual).

GRADCHEC/NOGRADCHEC evaluates and compares the analytic and numerical gradient for the current model at the starting values. No actual estimation takes place. Useful for checking derivatives of a new likelihood function. The numeric gradient is evaluated in a time-consuming but accurate way. See the GRAD=C4 option.

GRADIENT= ANALYTIC or C2 or C4 or FORWARD specifies the method of calculating numeric first derivatives. GRAD=A is the default when analytic first derivatives are available (usually the case).

GRAD=FORWARD calculates the numeric derivatives for a given parameter B as

$$D = (F(B+EPS) - F(B))/EPS \quad (1 \text{ function evaluation per parameter})$$

GRAD=C2 (CENTRAL2) uses

$$D = (F(B+EPS) - F(B-EPS))/(2*EPS) \quad (2 \text{ function evaluations per parameter})$$

GRAD=C4 uses

$$D = (-F(B+2*EPS) + 8*F(B+EPS) - 8*F(B-EPS) + F(B-2*EPS))/(12*EPS) \quad (4 \text{ function evaluations per parameter})$$

In all cases, EPS = MAX(ABS(.001*B) , .0001)

HESSCHEC/NOHESSCHEC checks the analytic second derivatives against the "discrete Hessian" (first differences of analytic first derivatives). Normally this is only used when debugging second derivatives in new TSP commands.

HCOV=B or N or G or F or D or W or R or BNW, etc., specifies the method for calculating the asymptotic covariance

NONLINEAR

matrix of the parameter estimates (and standard errors). The default is usually N or B, depending on the procedure. Some procedures may not have N (and thus W) available. A label is printed below each table of standard errors and asymptotic t-statistics identifying the method of calculation used. More than one method may be specified for alternative VCOV matrices and standard errors. In this case, the first method is stored in @VCOV and @SES, and the VCOV for each method is stored under the name constructed by appending the letter to @VCOV. For example, HCOV=NB would store @VCOV, @VCOVN, and @VCOVB.

The methods available are:

- B = BHHH (Berndt-Hall-Hall-Hausman). Covariance of the analytic gradient. The default for ML.
- N = Newton (Analytic second derivatives). The default for PROBIT, TOBIT, LOGIT, SAMPSEL.
- G = GAUSS (Gauss-Newton). Quadratic form of the analytic gradient and the residual covariance matrix. The default for LSQ and for FIML iterations.
- F = BFGS (Broyden-Fletcher-Goldfarb-Shanno). Analytic or numeric first derivatives, and rank 1 update approximation of the Hessian from iterations. Usually HITER=F is superior to HITER=D. The HCOV=F option is valid only if HITER=F.
- D = DFP (Davidon-Fletcher-Powell). Analytic or numeric first derivatives, and rank 1 update approximation of the Hessian from iterations. This option is valid only if HITER=D. For upward compatibility, it implies a default of GRADIENT=C4.
- W = Eicker-White. A combination of analytic second derivatives and BHHH (see the White Reference). This option is not available for HITER=
- R = Robust. Robust to heteroscedasticity. This is equivalent to W and used in LSQ only. It is never used for iteration.
- NBW = print all three alternative standard errors.

HITER=B or N or G or F or D specifies the method of Hessian (second derivative matrix) approximation to be used during the parameter iterations. The options are the same as those described above for the estimate of the covariance matrix of the parameter estimates. When HITER=N is used, a special "modified Cholesky factorization" (Gill, Murray and Wright (1981)) is used to iterate past regions where the Hessian is locally indefinite.

MAXIT= maximum number of iterations. The default is 20.

MAXSQZ= maximum number of "squeezes" in the stepsize search. The default depends on STEP:

STEP	MAXSQZ
BARD	10
BARDB	10
CEA	10
CEAB	10
GOLDEN	20

PRINT/NOPRINT produces short diagnostic output at each iteration, including a table of the current parameter estimates and their change vector. The value of the objective function for each squeeze on the change vector is also printed. **CRIT** is the norm of the gradient in the metric of the Hessian, which approaches zero at convergence.

SILENT/NOSILENT suppresses all output, even non-convergence warnings (check @IFCONV). See also **TERSE**.

NONLINEAR

STEP= BARD or BARDB or CEA or CEAB or GOLDEN specifies the stepsize method for squeezing. The default depends on HITER and the procedure:

HITER	STEP
Newton	CEA
BHHH	CEA
Gauss	BARD(LSQ), CEAB(FIML)
DFP	GOLDEN

SQZTOL= tolerance of determining stepsize. Used for STEP=GOLDEN. The default is .1.

SYMMETRIC/NOSYMMETRIC is an old option which has been replaced with **GRADIENT=method**. **SYMMETRIC** is the same as **GRAD=C4**; **NOSYM** is equivalent to **GRAD=FORWARD**.

TERSE/NOTERSE yields minimal output. It suppresses all output except the objective function, number of observations, and the table of parameter estimates, standard errors, and asymptotic t-statistics. Any non-convergence warning is printed. See also **SILENT**.

TOL= tolerance of determining convergence of the parameters, using a unit stepsize. The default is .001.

TOLG= tolerance of determining convergence of the norm of the gradient (printed as CRIT in the output). The default is .001. $CRIT = g'H^1g$, which is usually many orders of magnitude smaller than .001.

TOLS= tolerance of determining convergence of the parameters, using the squeezed step. The default is 0, i.e. ignore the squeezed change in parameters and use the regular TOL instead.

VERBOSE/NOVERBOSE produces lots of diagnostic output, including the gradient, Hessian, and inverse Hessian at each iteration, and the non-inverted Hessian for each output VCOV.

Starting values:

The default values depend on the procedure. For the standard TSP models which are (potentially) nonlinear in the parameters (LSQ, 3SLS, GMM, FIML, ML), the user provides them with **PARAM** and **SET**. **PROBIT** and **LOGIT** use zeros. **TOBIT** uses a regression and formulas from Greene (1981). **SAMPSEL** uses probit and a regression.

The default is overridden in the linear model procedures (**ARCH**, **BJEST**, **PROBIT**, **TOBIT**, **LOGIT**, and **SAMPSEL**) if the user supplies a matrix named **@START**. The length of **@START** must be equal to the number of parameters in the estimation (otherwise it is ignored). The easiest way to create **@START** is with a statement like:

```
MMAKE @START 12.3 4.56 33 44 55;
```

The order of the parameters in **@START** is obvious for most of the linear models, except for the following:

TOBIT: **SIGMA** comes last.

SAMPSEL: the probit equation is first, then the regression equation, and then **SIGMA** and **RHO**.

Method:

The method used is a standard gradient method, explained in more detail in Chapter 10 of the *User's Guide*. Briefly, at each iteration, a new parameter vector is computed by moving in the direction specified by the gradient of the likelihood (uphill), weighting this gradient by an approximation to the matrix of second derivatives at that point (in order to adjust for the curvature). Convergence is declared when the changes in the parameters are all "small", where "small" is defined by the **TOL**= option.

NONLINEAR

References:

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.

Fletcher, R., **Practical Methods of Optimization, Volume I: Unconstrained Optimization**, John Wiley and Sons, New York, 1980.

Gill, Philip E., Walter Murray, and Margaret H. Wright, **Practical Optimization**, Academic Press, New York, 1981.

Goldfeld, S. M. and R. E. Quandt, **Nonlinear Methods in Econometrics**, North-Holland, 1972.

Greene, William H., "On the Asymptotic Bias of the Ordinary Least Squares Estimator of the Tobit Model," **Econometrica** **49**, March 1981, pp. 505-513.

Quandt, Richard E., "Computational Problems and Methods," in Griliches and Intriligator, eds., **Handbook of Econometrics, Volume I**, North-Holland Publishing Company, Amsterdam, 1983.

White, Halbert, "Maximum Likelihood Estimation of Misspecified Models", **Econometrica** **50** (1982), pp. 1-26.