# TOBIT

---

TOBIT (MILLS=*name for output inverse Mills ratio*,
WEIGHT=*weighting series*,nonlinear options)
*dependent variable  list of independent variables* ;

---

**Function:**

TOBIT obtains estimates of the linear Tobit model, where the dependent variable is either zero or positive. The method used is maximum likelihood under the assumption of homoskedastic normal disturbances.

**Usage:**

The basic TOBIT statement is like the PROBIT or OLSQ statements: first list the dependent variable and then the independent variables. If you wish to have an intercept term in the regression (usually recommended), include the special variable C or CONSTANT in your list of independent variables. You may have as many independent variables as you like subject to the overall limits on the number of arguments per statement and the amount of working space, but of course the number is limited by the number of data observations you have available.

The observations over which the regression is computed are determined by the current sample. If any of the observations have missing values within the current sample, TOBIT will print a warning message and will drop those observations.

The list of independent variables on the TOBIT command may include variables with explicit lags and leads as well as PDL (Polynomial Distributed Lag) variables. These distributed lag variables are a way to reduce the number of free coefficients when you are entering a large number of lagged variables in a regression by imposing smoothness on the coefficients. See PDL for a description of how to specify such variables.

The dependent variable need not be a strictly zero/positive variable. Negative values are treated as zero. The standard Tobit model involves truncation of the dependent variable below zero. Models with either upper or lower truncation at any constant point can be estimated by transforming the dependent variable to the standard case. See the Examples for more details.

**Options:**

**MILLS=** the name of a series used to store the inverse Mills ratio series evaluated at the estimated parameters. The default is @MILLS.

**WEIGHT=** the name of a weighting series. The weights are applied directly to the likelihood function, and no normalization is performed.

Nonlinear options - see NONLINEAR in this manual.

**Examples:**

Standard Tobit model, truncation below zero:

```
TOBIT CAR C INCOME RURAL MSTAT;
```

# TOBIT

Truncation below two:

```
GENR CAR2 = CAR-2;
TOBIT CAR2 C INCOME RURAL MSTAT;
SET @COEF(1) = @COEF(1)+2;  ?CORRECT FOR TRANSFORMATION
PRINT @COEF;
```

Truncation above ten:

```
GENR CAR10 = 10-CAR;
TOBIT CAR10 C INCOME RURAL MSTAT;
MMULT -1 @COEF @COEF;       ?CORRECT FOR TRANSFORMATION
SET @COEF(1) = @COEF(1)+10;  PRINT @COEF;
```

**Output:**

The output of TOBIT begins with an equation title and the name of the dependent variable. Then the starting values and diagnostic output from the iterations are printed, followed by the convergence status.

The results printed are the mean of the dependent variable, the number of positive observations, and a table of right hand side variable names, estimated coefficients, standard errors and associated t-statistics. The estimated standard deviation of the residual, SIGMA, is listed last in this table.

TOBIT also stores some of these results in data storage for your later use. The table below lists the results available after a TOBIT command.

| Name | Type | Length | Variable Description |
|---|---|---|---|
| @LHV | list | 1 | Name of dependent variable. |
| @YMEAN | scalar | 1 | Fraction of positive observations. |
| @NOB | scalar | 1 | Number of observations. |
| @NPOS | scalar | 1 | Number of positive observations. |
| @LOGL | scalar | 1 | Log of likelihood function. |
| @IFCONV | scalar | 1 | 1 if convergence achieved, 0 otherwise. |
| @NCOEF | scalar | 1 | Number of rhs variables (#vars). |
| @NCID | scalar | 1 | Number of identified coefficients. |
| @RNMS | list | #vars | Names of right hand side variables. |
| @COEF | vector | #vars | Coefficient estimates. |
| @SES | vector | #vars | Standard errors. |
| @T | vector | #vars | T-statistics. |
| %T | vector | #vars | p-values for T-statistics. |
| @GRAD | vector | #vars | Gradient of log L at convergence. |
| @VCOV | matrix | #vars*#vars | Variance-covariance of estimated coeffecients. |
| @DPDX | matrix | #varx*2 | Mean of probability derivatives. |
| @RES | series | #obs | Residuals for non-truncated obs. |
| @MILLS | series | #obs | Inverse Mills ratios. |

If the regression includes a PDL variable, the following will also be stored:

| | | | |
|---|---|---|---|
| @SLAG | scalar | 1 | Sum of the lag coefficients. |
| @MLAG | scalar | 1 | Mean lag coefficient (number of time periods). |
| @LAGF | vector | #lags | Estimated lag coefficients, after "unscrambling". |

**Method:**

TOBIT uses analytic first and second derivatives to obtain maximum likelihood estimates via the Newton-Raphson algorithm. This algorithm usually converges fairly quickly. Starting values for the parameters are obtained from a regression on the observations with positive y values. See Greene (1981), p. 508, formula (13) and footnote 5 for the details. Alternative starting values may be supplied in @START (see NONLINEAR in this manual). A globally concave parameterization of the likelihood function is used for the iteration.

If there are more independent variables than the number of observations with positive y values, the estimation is halted. Multicollinearity of the independent variables is handled with generalized inverses, like TSP regression procedures.

The numerical implementation involves evaluating the normal density and cumulative normal distribution functions. The cumulative normal distribution function is computed from an asymptotic expansion, since it has no closed form. See the references under CDF for the actual method used to evaluate CNORM(). The ratio of the density to the distribution function is also known as the inverse Mills ratio. This is used in the derivatives and with the MILLS= option.

**References:**

Amemiya, Takeshi, "Tobit Models: A Survey," **Journal of Econometrics 24**, December 1981, pp. 3-61.

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

Maddala, G. S., **Limited-dependent and Qualitative Variables in Econometrics**, Cambridge University Press, New York, 1983, pp. 151-155.