# Title

stata.com

**xttobit** — Random-effects tobit models



## Syntax

```
xttobit depvar [indepvars] [if] [in] [weight] [, options]
```

| *options* | Description |
|---|---|
| **Model** | |
| noconstant | suppress constant term |
| ll(*varname*\|#) | left-censoring variable/limit |
| ul(*varname*\|#) | right-censoring variable/limit |
| offset(*varname*) | include *varname* in model with coefficient constrained to 1 |
| constraints(*constraints*) | apply specified linear constraints |
| collinear | keep collinear variables |
| **SE** | |
| vce(*vcetype*) | *vcetype* may be oim, bootstrap, or jackknife |
| **Reporting** | |
| level(#) | set confidence level; default is level(95) |
| tobit | perform likelihood-ratio test comparing against pooled tobit model |
| noskip | perform overall model test as a likelihood-ratio test |
| nocnsreport | do not display constraints |
| *display_options* | control column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling |
| **Integration** | |
| intmethod(*intmethod*) | integration method; *intmethod* may be mvaghermite (the default) or ghermite |
| intpoints(#) | use # quadrature points; default is intpoints(12) |
| **Maximization** | |
| *maximize_options* | control the maximization process; seldom used |
| coeflegend | display legend instead of statistics |

A panel variable must be specified; use xtset; see [XT] **xtset**.

*indepvars* may contain factor variables; see [U] **11.4.3 Factor variables**.

*depvar* and *indepvars* may contain time-series operators; see [U] **11.4.4 Time-series varlists**.

by, fp, and statsby are allowed; see [U] **11.1.10 Prefix commands**.

iweights are allowed; see [U] **11.1.6 weight**. Weights must be constant within panel.

coeflegend does not appear in the dialog box.

See [U] **20 Estimation and postestimation commands** for more capabilities of estimation commands.

# Menu

Statistics > Longitudinal/panel data > Censored outcomes > Tobit regression (RE)

# Description

`xttobit` fits random-effects tobit models. There is no command for a parametric conditional fixed-effects model, as there does not exist a sufficient statistic allowing the fixed effects to be conditioned out of the likelihood. Honoré (1992) has developed a semiparametric estimator for fixed-effect tobit models. Unconditional fixed-effects tobit models may be fit with the `tobit` command with indicator variables for the panels; the indicators can be created with the factor-variable syntax described in **[U] 11.4.3 Factor variables**. However, unconditional fixed-effects estimates are biased.

# Options

## Model

`noconstant`; see [R] **estimation options**.

`ll(`*varname*|*#*`)` and `ul(`*varname*|*#*`)` indicate the censoring points. You may specify one or both. `ll()` indicates the lower limit for left-censoring. Observations with *depvar* $\leq$ `ll()` are left-censored, observations with *depvar* $\geq$ `ul()` are right-censored, and remaining observations are not censored.

`offset(`*varname*`)`, `constraints(`*constraints*`)`, `collinear`; see [R] **estimation options**.

## SE

`vce(`*vcetype*`)` specifies the type of standard error reported, which includes types that are derived from asymptotic theory (`oim`) and that use bootstrap or jackknife methods (`bootstrap`, `jackknife`); see [XT] ***vce_options***.

## Reporting

`level(#)`; see [R] **estimation options**.

`tobit` specifies that a likelihood-ratio test comparing the random-effects model with the pooled (tobit) model be included in the output.

`noskip`; see [R] **estimation options**.

`nocnsreport`; see [R] **estimation options**.

*display_options*: `noomitted`, `vsquish`, `noemptycells`, `baselevels`, `allbaselevels`, `nofvlabel`, `fvwrap(#)`, `fvwrapon(`*style*`)`, `cformat(%`*fmt*`)`, `pformat(%`*fmt*`)`, `sformat(%`*fmt*`)`, and `nolstretch`; see [R] **estimation options**.

## Integration

`intmethod(`*intmethod*`)`, `intpoints(#)`; see [R] **estimation options**.

## Maximization

*maximize_options*: `difficult`, `technique(`*algorithm_spec*`)`, `iterate(#)`, `[no]log`, `trace`, `gradient`, `showstep`, `hessian`, `showtolerance`, `tolerance(#)`, `ltolerance(#)`, `nrtolerance(#)`, `nonrtolerance`, and `from(`*init_specs*`)`; see [R] **maximize**. These options are seldom used.

The following option is available with `xttobit` but is not shown in the dialog box:

`coeflegend`; see [R] **estimation options**.

## Remarks and examples

stata.com

Consider the linear regression model with panel-level random effects

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \nu_i + \epsilon_{it}$$

for $i = 1, \ldots, n$ panels, where $t = 1, \ldots, n_i$. The random effects, $\nu_i$, are i.i.d., $N(0, \sigma_\nu^2)$, and $\epsilon_{it}$ are i.i.d. $N(0, \sigma_\epsilon^2)$ independently of $\nu_i$.

The observed data, $y_{it}^o$, represent possibly censored versions of $y_{it}$. If they are left-censored, all that is known is that $y_{it} \leq y_{it}^o$. If they are right-censored, all that is known is that $y_{it} \geq y_{it}^o$. If they are uncensored, $y_{it} = y_{it}^o$. If they are left-censored, $y_{it}^o$ is determined by `ll()`. If they are right-censored, $y_{it}^o$ is determined by `ul()`. If they are uncensored, $y_{it}^o$ is determined by *depvar*.

### ▷ Example 1

Using the `nlswork` data described in [XT] **xt**, we fit a random-effects tobit model of adjusted (log) wages. We use the `ul()` option to impose an upper limit on the recorded log of wages. We use the `intpoints(25)` option to increase the number of integration points to 25 from 12, which aids convergence of this model.

```
. use http://www.stata-press.com/data/r13/nlswork3
(National Longitudinal Survey.  Young Women 14-26 years of age in 1968)
. xttobit ln_wage union age grade not_smsa south##c.year, ul(1.9)
> intpoints(25) tobit
  (output omitted)
Random-effects tobit regression                 Number of obs      =     19224
Group variable: idcode                          Number of groups   =      4148
Random effects u_i ~ Gaussian                   Obs per group: min =         1
                                                               avg =       4.6
                                                               max =        12
Integration method: mvaghermite                 Integration points =        25
                                                Wald chi2(7)       =   2924.91
Log likelihood  = -6814.4638                    Prob > chi2        =    0.0000
```

| ln_wage | Coef. | Std. Err. | z | P>\|z\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| union | .1430525 | .0069719 | 20.52 | 0.000 | .1293878 | .1567172 |
| age | .009913 | .0017517 | 5.66 | 0.000 | .0064797 | .0133463 |
| grade | .0784843 | .0022767 | 34.47 | 0.000 | .074022 | .0829466 |
| not_smsa | -.1339973 | .0092061 | -14.56 | 0.000 | -.1520409 | -.1159536 |
| 1.south | -.3507181 | .0695557 | -5.04 | 0.000 | -.4870447 | -.2143915 |
| year | -.0008283 | .0018372 | -0.45 | 0.652 | -.0044291 | .0027725 |
| south#c.year | | | | | | |
| 1 | .0031938 | .0008606 | 3.71 | 0.000 | .0015071 | .0048805 |
| _cons | .5101968 | .1006681 | 5.07 | 0.000 | .312891 | .7075025 |
| /sigma_u | .3045995 | .0048346 | 63.00 | 0.000 | .2951239 | .314075 |
| /sigma_e | .2488682 | .0018254 | 136.34 | 0.000 | .2452904 | .2524459 |
| rho | .599684 | .0084097 | | | .5831174 | .6160733 |

```
Likelihood-ratio test of sigma_u=0: chibar2(01)= 6650.63 Prob>=chibar2 = 0.000
  Observation summary:         0  left-censored observations
                           12334     uncensored observations
                            6890 right-censored observations
```

The output includes the overall and panel-level variance components (labeled `sigma_e` and `sigma_u`, respectively) together with $\rho$ (labeled `rho`)

$$\rho = \frac{\sigma_\nu^2}{\sigma_\epsilon^2 + \sigma_\nu^2}$$

which is the percent contribution to the total variance of the panel-level variance component.

When `rho` is zero, the panel-level variance component is unimportant, and the panel estimator is not different from the pooled estimator. A likelihood-ratio test of this is included at the bottom of the output. This test formally compares the pooled estimator (tobit) with the panel estimator.

◁

## ❑ Technical note

The random-effects model is calculated using quadrature, which is an approximation whose accuracy depends partially on the number of integration points used. We can use the `quadchk` command to see if changing the number of integration points affects the results. If the results change, the quadrature approximation is not accurate given the number of integration points. Try increasing the number of integration points using the `intpoints()` option and run `quadchk` again. Do not attempt to interpret the results of estimates when the coefficients reported by `quadchk` differ substantially. See [XT] **quadchk** for details and [XT] **xtprobit** for an example.

Because the `xttobit` likelihood function is calculated by Gauss–Hermite quadrature, on large problems the computations can be slow. Computation time is roughly proportional to the number of points used for the quadrature.

❑

# Stored results

xttobit stores the following in e():

Scalars

| | |
|---|---|
| e(N) | number of observations |
| e(N_g) | number of groups |
| e(N_unc) | number of uncensored observations |
| e(N_lc) | number of left-censored observations |
| e(N_rc) | number of right-censored observations |
| e(N_cd) | number of completely determined observations |
| e(k) | number of parameters |
| e(k_eq) | number of equations in e(b) |
| e(k_eq_model) | number of equations in overall model test |
| e(k_dv) | number of dependent variables |
| e(df_m) | model degrees of freedom |
| e(ll) | log likelihood |
| e(ll_0) | log likelihood, constant-only model |
| e(chi2) | $\chi^2$ |
| e(chi2_c) | $\chi^2$ for comparison test |
| e(rho) | $\rho$ |
| e(sigma_u) | panel-level standard deviation |
| e(sigma_e) | standard deviation of $\epsilon_{it}$ |
| e(n_quad) | number of quadrature points |
| e(g_min) | smallest group size |
| e(g_avg) | average group size |
| e(g_max) | largest group size |
| e(p) | significance |
| e(rank) | rank of e(V) |
| e(rank0) | rank of e(V) for constant-only model |
| e(ic) | number of iterations |
| e(rc) | return code |
| e(converged) | 1 if converged, 0 otherwise |

Macros

| | |
|---|---|
| e(cmd) | xttobit |
| e(cmdline) | command as typed |
| e(depvar) | names of dependent variables |
| e(ivar) | variable denoting groups |
| e(llopt) | contents of ll(), if specified |
| e(ulopt) | contents of ul(), if specified |
| e(k_aux) | number of auxiliary parameters |
| e(wtype) | weight type |
| e(wexp) | weight expression |
| e(title) | title in estimation output |
| e(offset1) | offset |
| e(chi2type) | Wald or LR; type of model $\chi^2$ test |
| e(chi2_ct) | Wald or LR; type of model $\chi^2$ test corresponding to e(chi2_c) |
| e(vce) | *vcetype* specified in vce() |
| e(vcetype) | title used to label Std. Err. |
| e(intmethod) | integration method |
| e(distrib) | Gaussian; the distribution of the random effect |
| e(opt) | type of optimization |
| e(which) | max or min; whether optimizer is to perform maximization or minimization |
| e(ml_method) | type of ml method |
| e(user) | name of likelihood-evaluator program |
| e(technique) | maximization technique |
| e(properties) | b V |
| e(predict) | program used to implement predict |
| e(asbalanced) | factor variables fvset as asbalanced |
| e(asobserved) | factor variables fvset as asobserved |

Matrices

| | |
|---|---|
| e(b) | coefficient vector |
| e(Cns) | constraints matrix |
| e(ilog) | iteration log |
| e(gradient) | gradient vector |
| e(V) | variance–covariance matrix of the estimator |

Functions

| | |
|---|---|
| e(sample) | marks estimation sample |

## Methods and formulas

Assuming a normal distribution, $N(0, \sigma_\nu^2)$, for the random effects $\nu_i$, we have the joint (unconditional of $\nu_i$) density of the observed data from the $i$th panel

$$f(y_{i1}^o, \ldots, y_{in_i}^o | \mathbf{x}_{i1}, \ldots, \mathbf{x}_{in_i}) = \int_{-\infty}^{\infty} \frac{e^{-\nu_i^2/2\sigma_\nu^2}}{\sqrt{2\pi}\sigma_\nu} \left\{ \prod_{t=1}^{n_i} F(y_{it}^o, \mathbf{x}_{it}\boldsymbol{\beta} + \nu_i) \right\} d\nu_i$$

where

$$F(y_{it}^o, \Delta_{it}) = \begin{cases} \left(\sqrt{2\pi}\sigma_\epsilon\right)^{-1} e^{-(y_{it}^o - \Delta_{it})^2/(2\sigma_\epsilon^2)} & \text{if } y_{it}^o \in C \\ \Phi\left(\frac{y_{it}^o - \Delta_{it}}{\sigma_\epsilon}\right) & \text{if } y_{it}^o \in L \\ 1 - \Phi\left(\frac{y_{it}^o - \Delta_{it}}{\sigma_\epsilon}\right) & \text{if } y_{it}^o \in R \end{cases}$$

where $C$ is the set of noncensored observations, $L$ is the set of left-censored observations, $R$ is the set of right-censored observations, and $\Phi()$ is the cumulative normal distribution.

The panel level likelihood $l_i$ is given by

$$l_i = \int_{-\infty}^{\infty} \frac{e^{-\nu_i^2/2\sigma_\nu^2}}{\sqrt{2\pi}\sigma_\nu} \left\{ \prod_{t=1}^{n_i} F(y_{it}^o, \mathbf{x}_{it}\boldsymbol{\beta} + \nu_i) \right\} d\nu_i$$

$$\equiv \int_{-\infty}^{\infty} g(y_{it}^o, x_{it}, \nu_i) d\nu_i$$

This integral can be approximated with $M$-point Gauss–Hermite quadrature

$$\int_{-\infty}^{\infty} e^{-x^2} h(x) dx \approx \sum_{m=1}^{M} w_m^* h(a_m^*)$$

This is equivalent to

$$\int_{-\infty}^{\infty} f(x) dx \approx \sum_{m=1}^{M} w_m^* \exp\left\{(a_m^*)^2\right\} f(a_m^*)$$

where the $w_m^*$ denote the quadrature weights and the $a_m^*$ denote the quadrature abscissas. The log likelihood, $L$, is the sum of the logs of the panel level likelihoods $l_i$.

The default approximation of the log likelihood is by adaptive Gauss–Hermite quadrature, which approximates the panel level likelihood with

$$l_i \approx \sqrt{2}\widehat{\sigma}_i \sum_{m=1}^{M} w_m^* \exp\left\{(a_m^*)^2\right\} g(y_{it}^o, x_{it}, \sqrt{2}\widehat{\sigma}_i a_m^* + \widehat{\mu}_i)$$

where $\widehat{\sigma}_i$ and $\widehat{\mu}_i$ are the adaptive parameters for panel $i$. Therefore, with the definition of $g(y_{it}^o, x_{it}, \nu_i)$, the total log likelihood is approximated by

$$L \approx \sum_{i=1}^{n} w_i \log\left[\sqrt{2}\widehat{\sigma}_i \sum_{m=1}^{M} w_m^* \exp\{(a_m^*)^2\} \frac{\exp\{-(\sqrt{2}\widehat{\sigma}_i a_m^* + \widehat{\mu}_i)^2/2\sigma_\nu^2\}}{\sqrt{2\pi}\sigma_\nu} \prod_{t=1}^{n_i} F(y_{it}^o, x_{it}\boldsymbol{\beta} + \sqrt{2}\widehat{\sigma}_i a_m^* + \widehat{\mu}_i)\right]$$

where $w_i$ is the user-specified weight for panel $i$; if no weights are specified, $w_i = 1$.

The default method of adaptive Gauss–Hermite quadrature is to calculate the posterior mean and variance and use those parameters for $\widehat{\mu}_i$ and $\widehat{\sigma}_i$ by following the method of Naylor and Smith (1982), further discussed in Skrondal and Rabe-Hesketh (2004). We start with $\widehat{\sigma}_{i,0} = 1$ and $\widehat{\mu}_{i,0} = 0$, and the posterior means and variances are updated in the $k$th iteration. That is, at the $k$th iteration of the optimization for $l_i$ we use

$$l_{i,k} \approx \sum_{m=1}^{M} \sqrt{2}\widehat{\sigma}_{i,k-1} w_m^* \exp\{a_m^*)^2\} g(y_{it}^o, x_{it}, \sqrt{2}\widehat{\sigma}_{i,k-1} a_m^* + \widehat{\mu}_{i,k-1})$$

Letting

$$\tau_{i,m,k-1} = \sqrt{2}\widehat{\sigma}_{i,k-1} a_m^* + \widehat{\mu}_{i,k-1}$$

$$\widehat{\mu}_{i,k} = \sum_{m=1}^{M} (\tau_{i,m,k-1}) \frac{\sqrt{2}\widehat{\sigma}_{i,k-1} w_m^* \exp\{(a_m^*)^2\} g(y_{it}^o, x_{it}, \tau_{i,m,k-1})}{l_{i,k}}$$

and

$$\widehat{\sigma}_{i,k} = \sum_{m=1}^{M} (\tau_{i,m,k-1})^2 \frac{\sqrt{2}\widehat{\sigma}_{i,k-1} w_m^* \exp\{(a_m^*)^2\} g(y_{it}^o, x_{it}, \tau_{i,m,k-1})}{l_{i,k}} - (\widehat{\mu}_{i,k})^2$$

and this is repeated until $\widehat{\mu}_{i,k}$ and $\widehat{\sigma}_{i,k}$ have converged for this iteration of the maximization algorithm. This adaptation is applied on every iteration until the log-likelihood change from the preceding iteration is less than a relative difference of 1e–6; after this, the quadrature parameters are fixed.

The log likelihood can also be calculated by nonadaptive Gauss–Hermite quadrature, the `intmethod(ghermite)` option:

$$L = \sum_{i=1}^{n} w_i \log\Big\{\Pr(y_{i1},\ldots,y_{in_i}|\mathbf{x}_{i1},\ldots,\mathbf{x}_{in_i})\Big\}$$
$$\approx \sum_{i=1}^{n} w_i \log\left[\frac{1}{\sqrt{\pi}}\sum_{m=1}^{M} w_m^* \prod_{t=1}^{n_i} F\left\{y_{it}^o, \mathbf{x}_{it}\boldsymbol{\beta} + \sqrt{2}\sigma_\nu a_m^*\right\}\right]$$

Both quadrature formulas require that the integrated function be well approximated by a polynomial of degree equal to the number of quadrature points. The number of periods (panel size) can affect whether

$$\prod_{t=1}^{n_i} F(y_{it}^o, \mathbf{x}_{it}\boldsymbol{\beta} + \nu_i)$$

is well approximated by a polynomial. As panel size and $\rho$ increase, the quadrature approximation can become less accurate. For large $\rho$, the random-effects model can also become unidentified. Adaptive quadrature gives better results for correlated data and large panels than nonadaptive quadrature; however, we recommend that you use the `quadchk` command (see [XT] **quadchk**) to verify the quadrature approximation used in this command, whichever approximation you choose.

## References

Honoré, B. E. 1992. Trimmed LAD and least squares estimation of truncated and censored regression models with fixed effects. *Econometrica* 60: 533–565.

Naylor, J. C., and A. F. M. Smith. 1982. Applications of a method for the efficient computation of posterior distributions. *Journal of the Royal Statistical Society, Series C* 31: 214–225.

Pendergast, J. F., S. J. Gange, M. A. Newton, M. J. Lindstrom, M. Palta, and M. R. Fisher. 1996. A survey of methods for analyzing clustered binary response data. *International Statistical Review* 64: 89–118.

Skrondal, A., and S. Rabe-Hesketh. 2004. *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. Boca Raton, FL: Chapman & Hall/CRC.

## Also see

[XT] **xttobit postestimation** — Postestimation tools for xttobit

[XT] **quadchk** — Check sensitivity of quadrature approximation

[XT] **xtintreg** — Random-effects interval-data regression models

[XT] **xtreg** — Fixed-, between-, and random-effects and population-averaged linear models

[R] **tobit** — Tobit regression

**[U] 20 Estimation and postestimation commands**