*xttobit* fits random-effects tobit models for panel data where the outcome variable is censored. Censoring limits may be fixed for all observations or vary across observations. The user can request that a likelihood-ratio test comparing the panel tobit model with the pooled tobit model be conducted at estimation time.

**Quick start**

Tobit model of \( y \) on \( x \) where \( y \) is censored at a lower limit of 5 using *xtset* data

```
xttobit y x, ll(5)
```

Add indicators for levels of categorical variable \( a \)

```
xttobit y x i.a, ll(5)
```

As above, but specify that censoring occurs at 5 and 25

```
xttobit y x i.a, ll(5) ul(25)
```

As above, but where lower and upper are variables containing the censoring limits

```
xttobit y x i.a, ll(lower) ul(upper)
```

Add likelihood-ratio test comparing the random-effects model with the pooled model

```
xttobit y x i.a, ll(lower) ul(upper) tobit
```

**Menu**

Statistics > Longitudinal/panel data > Censored outcomes > Tobit regression (RE)
Random-effects tobit models

**Syntax**

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

**Options**

<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>noconstant</td>
<td>suppress constant term</td>
</tr>
<tr>
<td>ll(varname</td>
<td>#)</td>
</tr>
<tr>
<td>ul(varname</td>
<td>#)</td>
</tr>
<tr>
<td>offset(varname)</td>
<td>include varname in model with coefficient constrained to 1</td>
</tr>
<tr>
<td>constraints(constraints)</td>
<td>apply specified linear constraints</td>
</tr>
</tbody>
</table>

**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
lrmodel perform the likelihood-ratio model test instead of the default Wald test
nocnsreport do not display constraints
display_options control columns and 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
collinear keep collinear variables
coefflegend 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.
collinear and coefflegend do not appear in the dialog box.

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

---

**Options**

- **Model**
  - `noconstant`; see [R] Estimation options.
  - `ll(varname | #)` and `ul(varname | #)` indicate the lower and upper limits for censoring, respectively. Observations with `depvar ≤ ll()` are left-censored; observations with `depvar ≥ ul()` are right-censored; and remaining observations are not censored. You do not have to specify the
censoring values. If you specify \texttt{ll}, the lower limit is the minimum of \texttt{depvar}. If you specify \texttt{ul}, the upper limit is the maximum of \texttt{depvar}.

\texttt{offset(varname)}, \texttt{constraints(constraints)}; see \texttt{[R] Estimation options}.

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

\texttt{level(#)}; see \texttt{[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.

\texttt{lrmodel}, \texttt{nocnsreport}; see \texttt{[R] Estimation options}.

\texttt{display_options}: \texttt{noci}, \texttt{nopvalues}, \texttt{noomitted}, \texttt{vsquish}, \texttt{noemptycells}, \texttt{baselevels}, \texttt{allbaselevels}, \texttt{nofvlabel}, \texttt{fvwrap(#)}, \texttt{fvwraption(style)}, \texttt{cformat(\%fmt)}, \texttt{pformat(\%fmt)}, \texttt{sformat(\%fmt)}, and \texttt{nolstretch}; see \texttt{[R] Estimation options}.

\texttt{Integration}\n
\texttt{intmethod(intmethod)}, \texttt{intpoints(#)}; see \texttt{[R] Estimation options}.

\texttt{Maximization}\n
\texttt{maximize_options}: \texttt{difficult}, \texttt{technique(algorithm_spec)}, \texttt{iterate(#)}, \texttt{[no]log}, \texttt{trace}, \texttt{gradient}, \texttt{showstep}, \texttt{hessian}, \texttt{showtolerance}, \texttt{tolerance(#)}, \texttt{ltolerance(#)}, \texttt{nrtolerance(#)}, \texttt{nonrtolerance}, and \texttt{from(init_specs)}; see \texttt{[R] Maximize}. These options are seldom used.

The following options are available with \texttt{xttobit} but are not shown in the dialog box:

\texttt{collinear}, \texttt{coeflegend}; see \texttt{[R] Estimation options}.

### Remarks and examples \href{stata.com}{stata.com}

\texttt{xttobit} fits random-effects tobit models. There is no command for a fixed-effects model, because there does not exist a sufficient statistic allowing the fixed effects to be conditioned out of the likelihood.

Consider the linear regression model with panel-level random effects

\[
y_{it} = x_{it}\beta + \nu_i + \epsilon_{it}
\]

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

The observed data, \(y^o_{it}\), represent possibly censored versions of \(y_{it}\). If they are left-censored, all that is known is that \(y_{it} \leq y^o_{it}\). If they are right-censored, all that is known is that \(y_{it} \geq y^o_{it}\). If they are uncensored, \(y_{it} = y^o_{it}\). If they are left-censored, \(y^o_{it}\) is determined by \texttt{ll()}, If they are right-censored, \(y^o_{it}\) is determined by \texttt{ul()} If they are uncensored, \(y^o_{it}\) is determined by \texttt{depvar}.
Example 1: Random-effects tobit regression

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.

```
. use https://www.stata-press.com/data/r16/nlswork3
(National Longitudinal Survey. Young Women 14-26 years of age in 1968)
. xttobit ln_wage i.union age grade not_smsa south##c.year, ul(1.9) tobit
(output omitted)
```

| Estimate | Std. Error | z | P>|z| | 95% Conf. Interval |
|----------|------------|---|------|-------------------|
| ln_wage  |            |   |      |                   |
| 1.union  | .1430527   | .0069718 | 20.52| 0.000 | .1293883 -.1567171 |
| age      | .0099132   | .0017516 | 5.66 | 0.000 | .0064801 .0134646 |
| grade    | .0784855   | .0022764 | 34.48| 0.000 | .0740239 .0829472 |
| not_smsa | -.1339978  | .009206  | -14.56| 0.000 | -.1520413 -.1159544 |
| 1.south  | -.3507188  | .0695554 | -5.04| 0.000 | -.4870449 -.2143928 |
| year     | -.0008285  | .0018371 | -0.45| 0.652 | -.0044292 .0027721 |
| south#c.year 1 | .0031938 | .0008606 | 3.71 | 0.000 | .0015071 .0048805 |
| _cons    | .5101956   | .1006646 | 5.07 | 0.000 | .3128966 .7074946 |
| /sigma_u | .3045992   | .0048344 | 63.01| 0.000 | .2951239 .3140745 |
| /sigma_e | .2488678   | .0018254 | 136.34| 0.000 | .24529 .2524455 |
| rho      | .5996844   | .0084095 | 583118 | 0.6160734 |

LR test of sigma_u=0: chibar2(01) = 6650.63  Prob >= chibar2 = 0.000

The results from a tobit regression can be interpreted as we would those from a linear regression. Because the dependent variable is log transformed, the coefficients can be interpreted in terms of a percentage change. We see, for example, that on average, union members make 14.3% more than nonunion members.

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

\[
\rho = \frac{\sigma^2_v}{\sigma^2_v + \sigma^2_e}
\]

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. In this case, we reject the null hypothesis that there are no panel-level effects.
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(ll_c) log likelihood, comparison 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) p-value for model test
- 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)`  \( \texttt{xttobit} \)
- `e(cmdline)`  \text{command as typed}
- `e(depvar)`  \text{names of dependent variables}
- `e(ivar)`  \text{variable denoting groups}
- `e(llopt)`  \text{contents of } \texttt{ll()}
- `e(ulopt)`  \text{contents of } \texttt{ul()}
- `e(k_aux)`  \text{number of auxiliary parameters}
- `e(wtype)`  \text{weight type}
- `e(wexp)`  \text{weight expression}
- `e(title)`  \text{title in estimation output}
- `e(offset)`  \text{offset}
- `e(chi2type)`  \text{Wald or LR; type of model } \chi^2 \text{ test}
- `e(chi2_ct)`  \text{Wald or LR; type of model } \chi^2 \text{ test corresponding to } e(chi2_c)
- `e(vce)`  \text{vcetype specified in vce()}
- `e(intmethod)`  \text{integration method}
- `e(distrib)`  \text{Gaussian; the distribution of the random effect}
- `e(opt)`  \text{type of optimization}
- `e(which)`  \text{max or min; whether optimizer is to perform maximization or minimization}
- `e(ml_method)`  \text{type of ml method}
- `e(user)`  \text{name of likelihood-evaluator program}
- `e(technique)`  \text{maximization technique}
- `e(properties)`  \text{b V}
- `e(predict)`  \text{program used to implement predict}
- `e(marginsok)`  \text{predictions allowed by margins}
- `e(asbalanced)`  \text{factor variables fvset as asbalanced}
- `e(asobserved)`  \text{factor variables fvset as asobserved}

**Matrices**

- `e(b)`  \text{coefficient vector}
- `e(Cns)`  \text{constraints matrix}
- `e(ilog)`  \text{iteration log}
- `e(gradient)`  \text{gradient vector}
- `e(V)`  \text{variance–covariance matrix of the estimators}

**Functions**

- `e(sample)`  \text{marks estimation sample}

**Methods and formulas**

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

\[
f(y_{i1}^o, \ldots, y_{in_i}^o | x_{i1}, \ldots, x_{in_i}) = \int_{-\infty}^{\infty} \frac{e^{-\nu_i^2/2\sigma^2}}{\sqrt{2\pi}\sigma} \left\{ \prod_{t=1}^{n_i} F(y_{it}^o, x_{it}\beta + \nu_i) \right\} d\nu_i
\]

where

\[
F(y_{it}^o, \Delta_{it}) = \begin{cases} 
(\sqrt{2\pi}\sigma) \Phi \left( \frac{y_{it}^o - \Delta_{it}}{\sigma} \right) & \text{if } y_{it}^o \in C \\
1 - \Phi \left( \frac{y_{it}^o - \Delta_{it}}{\sigma} \right) & \text{if } y_{it}^o \in L \\
\end{cases}
\]

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

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

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

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

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

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

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

where $\hat{\sigma}_i$ and $\hat{\mu}_i$ 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

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

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 $\hat{\mu}_i$ and $\hat{\sigma}_i$ by following the method of Naylor and Smith (1982), further discussed in Skrondal and Rabe-Hesketh (2004). We start with $\hat{\sigma}_{i,0} = 1$ and $\hat{\mu}_{i,0} = 0$, 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

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

\[
\hat{\mu}_{i,k} = \sum_{m=1}^{M} \left( \tau_{i,m,k} - 1 \right) \frac{\sqrt{2} \hat{\sigma}_{i,k} - 1 w_m^* \exp\{ (a_m^*)^2 \} g(y_{it}^0, x_{it}, \tau_{i,m,k} - 1) }{l_{i,k}}
\]

and

\[
\hat{\sigma}_{i,k} = \sum_{m=1}^{M} \left( \tau_{i,m,k} - 1 \right)^2 \frac{\sqrt{2} \hat{\sigma}_{i,k} - 1 w_m^* \exp\{ (a_m^*)^2 \} g(y_{it}^0, x_{it}, \tau_{i,m,k} - 1) }{l_{i,k} - (\hat{\mu}_{i,k})^2}
\]

and this is repeated until \( \hat{\mu}_{i,k} \) and \( \hat{\sigma}_{i,k} \) 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 if the `intmethod(ghermite)` option is specified. For nonadaptive Gauss–Hermite quadrature, the following formula for the log likelihood is used in place of (1).

\[
L = \sum_{i=1}^{n} w_i \log \left\{ \Pr(y_{i1}, \ldots, y_{in_i} | x_{i1}, \ldots, x_{in_i}) \right\}
\]

\[\approx \sum_{i=1}^{n} w_i \log \left[ \frac{1}{\sqrt{\pi}} \prod_{m=1}^{M} w_m^* \prod_{t=1}^{n_i} F \left( y_{it}^0, x_{it} \beta + \sqrt{2} \nu_{i} \right \} \right] \]

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

\[
\prod_{t=1}^{n_i} F(y_{it}^0, x_{it} \beta + \nu_{i})
\]

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


Also see

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

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

[XT] **xteintreg** — Extended random-effects interval regression

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

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

[XT] **xtset** — Declare data to be panel data

[ME] **metobit** — Multilevel mixed-effects tobit regression

[R] **tobit** — Tobit regression

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