# Title

stata.com

**tobit postestimation —** Postestimation tools for tobit

Postestimation commands predict margins Remarks and examples
Reference Also see

## Postestimation commands

The following postestimation commands are available after `tobit`:

| Command | Description |
|---|---|
| `contrast` | contrasts and ANOVA-style joint tests of estimates |
| `estat ic` | Akaike's, consistent Akaike's, corrected Akaike's, and Schwarz's Bayesian information criteria (AIC, CAIC, AICc, and BIC) |
| `estat summarize` | summary statistics for the estimation sample |
| `estat vce` | variance–covariance matrix of the estimators (VCE) |
| `estat (svy)` | postestimation statistics for survey data |
| `estimates` | cataloging estimation results |
| `etable` | table of estimation results |
| * `forecast` | dynamic forecasts and simulations |
| * `hausman` | Hausman's specification test |
| `lincom` | point estimates, standard errors, testing, and inference for linear combinations of coefficients |
| `linktest` | link test for model specification |
| * `lrtest` | likelihood-ratio test |
| `margins` | marginal means, predictive margins, marginal effects, and average marginal effects |
| `marginsplot` | graph the results from margins (profile plots, interaction plots, etc.) |
| `nlcom` | point estimates, standard errors, testing, and inference for nonlinear combinations of coefficients |
| `predict` | linear, censored, and truncated predictions |
| `predictnl` | point estimates, standard errors, testing, and inference for generalized predictions |
| `pwcompare` | pairwise comparisons of estimates |
| `suest` | seemingly unrelated estimation |
| `test` | Wald tests of simple and composite linear hypotheses |
| `testnl` | Wald tests of nonlinear hypotheses |

*`forecast`, `hausman`, and `lrtest` are not appropriate with `svy` estimation results.

# predict

## Description for predict

`predict` creates a new variable containing predictions such as linear predictions, standard errors, probabilities, and expected values.

## Menu for predict

Statistics > Postestimation

## Syntax for predict

```
predict [type] newvar [if] [in] [, statistic nooffset]

predict [type] stub* [if] [in], scores
```

| *statistic* | Description |
|---|---|
| Main | |
| `xb` | linear prediction; the default |
| `stdp` | standard error of the linear prediction |
| `stdf` | standard error of the forecast |
| `pr(`$a$`,`$b$`)` | $\Pr(a < y_j < b)$ |
| `e(`$a$`,`$b$`)` | $E(y_j \mid a < y_j < b)$ |
| `ystar(`$a$`,`$b$`)` | $E(y_j^*)$, $y_j^* = \max\{a, \min(y_j, b)\}$ |

These statistics are available both in and out of sample; type `predict ... if e(sample) ...` if wanted only for the estimation sample.

`stdf` is not allowed with `svy` estimation results.

where $a$ and $b$ may be numbers or variables; $a$ missing ($a \geq .$) means $-\infty$, and $b$ missing ($b \geq .$) means $+\infty$; see **[U] 12.2.1 Missing values**.

## Options for predict

Main

`xb`, the default, calculates the linear prediction.

`stdp` calculates the standard error of the prediction, which can be thought of as the standard error of the predicted expected value or mean for the observation's covariate pattern. The standard error of the prediction is also referred to as the standard error of the fitted value.

`stdf` calculates the standard error of the forecast, which is the standard error of the point prediction for 1 observation. It is commonly referred to as the standard error of the future or forecast value. By construction, the standard errors produced by `stdf` are always larger than those produced by `stdp`; see *Methods and formulas* in [R] **regress postestimation**.

`pr(`*a*`,`*b*`)` calculates $\Pr(a < \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j < b)$, the probability that $y_j|\mathbf{x}_j$ would be observed in the interval $(a, b)$.

*a* and *b* may be specified as numbers or variable names; *lb* and *ub* are variable names;
`pr(20,30)` calculates $\Pr(20 < \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j < 30)$;
`pr(`*lb*`,`*ub*`)` calculates $\Pr(lb < \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j < ub)$; and
`pr(20,`*ub*`)` calculates $\Pr(20 < \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j < ub)$.

*a* missing ($a \geq .$) means $-\infty$; `pr(.,30)` calculates $\Pr(-\infty < \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j < 30)$;
`pr(`*lb*`,30)` calculates $\Pr(-\infty < \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j < 30)$ in observations for which $lb \geq .$
and calculates $\Pr(lb < \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j < 30)$ elsewhere.

*b* missing ($b \geq .$) means $+\infty$; `pr(20,.)` calculates $\Pr(+\infty > \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j > 20)$;
`pr(20,`*ub*`)` calculates $\Pr(+\infty > \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j > 20)$ in observations for which $ub \geq .$
and calculates $\Pr(20 < \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j < ub)$ elsewhere.

`e(`*a*`,`*b*`)` calculates $E(\mathbf{x}_j\boldsymbol{\beta} + \epsilon_j \mid a < \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j < b)$, the expected value of $y_j|\mathbf{x}_j$ conditional on $y_j|\mathbf{x}_j$ being in the interval $(a, b)$, meaning that $y_j|\mathbf{x}_j$ is truncated.
*a* and *b* are specified as they are for `pr()`.

`ystar(`*a*`,`*b*`)` calculates $E(y_j^*)$, where $y_j^* = a$ if $\mathbf{x}_j\boldsymbol{\beta} + \epsilon_j \leq a$, $y_j^* = b$ if $\mathbf{x}_j\boldsymbol{\beta} + \epsilon_j \geq b$, and $y_j^* = \mathbf{x}_j\boldsymbol{\beta} + \epsilon_j$ otherwise, meaning that $y_j^*$ is censored. *a* and *b* are specified as they are for `pr()`.

`nooffset` is relevant only if you specified `offset(`*varname*`)`. It modifies the calculations made by `predict` so that they ignore the offset variable; the linear prediction is treated as $\mathbf{x}_j\boldsymbol{\beta}$ rather than as $\mathbf{x}_j\boldsymbol{\beta} + \text{offset}_j$.

`scores` calculates equation-level score variables.

The first new variable will contain $\partial \ln L/\partial(\mathbf{x}_j\boldsymbol{\beta})$.

The second new variable will contain $\partial \ln L/\partial\sigma$.

# margins

## Description for margins

`margins` estimates margins of response for linear predictions, probabilities, and expected values.

## Menu for margins

Statistics > Postestimation

## Syntax for margins

`margins` [*marginlist*] [`,` *options*]

`margins` [*marginlist*] `,` `predict(`*statistic* ...`)` [`predict(`*statistic* ...`)` ...] [*options*]

| *statistic* | Description |
|---|---|
| `xb` | linear prediction; the default |
| `pr(`*a*`,`*b*`)` | $\Pr(a < y_j < b)$ |
| `e(`*a*`,`*b*`)` | $E(y_j \mid a < y_j < b)$ |
| `ystar(`*a*`,`*b*`)` | $E(y_j^*)$, $y_j^* = \max\{a, \min(y_j, b)\}$ |
| `stdp` | not allowed with `margins` |
| `stdf` | not allowed with `margins` |

Statistics not allowed with `margins` are functions of stochastic quantities other than `e(b)`.

For the full syntax, see [R] **margins**.

# Remarks and examples

stata.com

## ▷ Example 1: Marginal predictions

In example 2 of [R] **tobit**, we fit a tobit model of annual number of hours that married women spend working and then obtained estimated average marginal effect of 47.47 for years of education on observed hours worked.

```
. use https://www.stata-press.com/data/r18/mroz87
(1975 PSID data from Mroz, 1987)
. tobit whrs75 nwinc wedyrs wexper c.wexper#c.wexper wifeage kl6 k618, ll(0)
  (output omitted)
. margins, dydx(wedyrs) predict(ystar(0,.))
  (output omitted)
```

However, we may not want this overall effect. To obtain marginal effects for specific alternative scenarios, we use `margins` with the `at()` option. For example, continuing with example 2 of [R] **tobit** to estimate the means of the marginal effects on the expected value of the censored outcome conditional on education ranging from 8 years to 17 years, we type

```
. margins, dydx(wedyrs) predict(ystar(0,.)) at(wedyrs=(8(1)17))
Average marginal effects                                  Number of obs = 753
Model VCE: OIM
Expression: E(whrs75*|whrs75>0), predict(ystar(0,.))
dy/dx wrt:  wedyrs
1._at:  wedyrs =  8
2._at:  wedyrs =  9
3._at:  wedyrs = 10
4._at:  wedyrs = 11
5._at:  wedyrs = 12
6._at:  wedyrs = 13
7._at:  wedyrs = 14
8._at:  wedyrs = 15
9._at:  wedyrs = 16
10._at: wedyrs = 17
```

| | dy/dx | Delta-method std. err. | z | P>\|z\| | [95% conf. | interval] |
|---|---|---|---|---|---|---|
| wedyrs | | | | | | |
| _at | | | | | | |
| 1 | 39.58775 | 8.432006 | 4.69 | 0.000 | 23.06132 | 56.11418 |
| 2 | 41.4497 | 9.421414 | 4.40 | 0.000 | 22.98407 | 59.91533 |
| 3 | 43.30531 | 10.41233 | 4.16 | 0.000 | 22.89752 | 63.71309 |
| 4 | 45.14859 | 11.39804 | 3.96 | 0.000 | 22.80885 | 67.48833 |
| 5 | 46.97371 | 12.37208 | 3.80 | 0.000 | 22.72489 | 71.22254 |
| 6 | 48.77504 | 13.32825 | 3.66 | 0.000 | 22.65216 | 74.89793 |
| 7 | 50.54717 | 14.26071 | 3.54 | 0.000 | 22.5967 | 78.49765 |
| 8 | 52.28499 | 15.16403 | 3.45 | 0.001 | 22.56403 | 82.00594 |
| 9 | 53.98369 | 16.03324 | 3.37 | 0.001 | 22.55912 | 85.40827 |
| 10 | 55.63887 | 16.8639 | 3.30 | 0.001 | 22.58624 | 88.6915 |

The estimated mean of the marginal effects is about 39.59 hours for 8 years of schooling, about 41.45 hours for 9 years of schooling, and so on.

◁

## Reference

McDonald, J. F., and R. A. Moffitt. 1980. The use of tobit analysis. *Review of Economics and Statistics* 62: 318–321. https://doi.org/10.2307/1924766.

## Also see

[R] **tobit** — Tobit regression

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