

Postestimation commands  
Methods and formulas

predict  
References

margins  
Also see

Remarks and examples

## Postestimation commands

The following postestimation command is of special interest after `eregress` and `xteregress`:

Command	Description
<code>estat teffects</code>	treatment effects and potential-outcome means

The following standard postestimation commands are also available after `eregress` and `xteregress`:

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

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

<sup>†</sup> `suest` and the survey data `estat` commands are not available after `xteregress`.

# predict

Predictions after eregress and xteregress are described in

- [ERM] **eregress predict** predict after eregress
- [ERM] **predict treatment** predict for treatment statistics
- [ERM] **predict advanced** predict’s advanced features

[ERM] **eregress predict** describes the most commonly used predictions. If you fit a model with treatment effects, predictions specifically related to these models are detailed in [ERM] **predict treatment**. [ERM] **predict advanced** describes less commonly used predictions, such as predictions of outcomes in auxiliary equations.

## margins

### Description for margins

margins estimates statistics based on fitted models. These statistics include marginal means, marginal probabilities, potential-outcome means, average and conditional derivatives, average and conditional effects, and treatment effects.

### Menu for margins

Statistics > Postestimation

### Syntax for margins

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

statistic	Description
Main	
<u>m</u> ean	mean; the default
<u>p</u> r	probability for binary or ordinal $y_j$
<u>p</u> omean	potential-outcome mean
<u>t</u> e	treatment effect
<u>t</u> et	treatment effect on the treated
<u>x</u> b	linear prediction excluding all complications
<u>p</u> r( $a, b$ )	$\Pr(a < y_j < b)$ for continuous $y_j$
<u>e</u> ( $a, b$ )	$E(y_j   a < y_j < b)$ for continuous $y_j$
<u>y</u> star( $a, b$ )	$E(y_j^*), y_j^* = \max\{a, \min(y_j, b)\}$ for continuous $y_j$
<u>e</u> xp <u>m</u> ean	calculate $E\{\exp(y_i)\}$

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

For the full syntax, see [R] margins.

## Remarks and examples

See [ERM] [Intro 7](#) for an overview of using margins and predict after eregress. For examples using margins, predict, and estat teffects, see [Interpreting effects](#) in [ERM] [Intro 9](#) and see [ERM] [Example 1a](#).

## Methods and formulas

This section contains methods and formulas for predictions and inference for the default average structural function. Methods and formulas for all other predictions are given in [Methods and formulas](#) of [ERM] [eregress](#). We begin with the cross-sectional model and then extend our discussion to the random effect models that we use for panel data.

In the linear regression model, for exogenous covariates  $\mathbf{x}_i$  and  $C$  endogenous regressors  $\mathbf{w}_i$ , we have

$$y_i = \mathbf{x}_i\boldsymbol{\beta} + \mathbf{w}_i\boldsymbol{\beta}_2 + \epsilon_i$$

where the error  $\epsilon_i$  is normal and correlated with  $\mathbf{w}_i$ .

Because  $\epsilon_i$  is a normally distributed, mean 0, random variable, we can split it into two mean 0, normally distributed, independent parts,

$$\epsilon_i = \gamma\epsilon_{2i} + \psi_i$$

where  $\epsilon_{2i}$  is the unobserved heterogeneity that gives rise to the endogeneity and  $\psi_i$  is an idiosyncratic error term with variance  $\sigma_\psi^2$ .

Conditional on the covariates and the unobserved heterogeneity, the conditional mean of  $y_i$  is

$$E(y_i|\mathbf{x}_i, \mathbf{w}_i, \epsilon_i) = \mathbf{x}_i\boldsymbol{\beta} + \mathbf{w}_i\boldsymbol{\beta}_2 + \gamma\epsilon_{2i}$$

Predictions and effects are computed based on the expression above. Including  $\epsilon_{2i}$  controls for endogeneity. Thus, all effects computed using the expression above have a structural interpretation. See [Imbens and Newey \(2009\)](#) and [Wooldridge \(2010\)](#) for a detailed description of structural functions for models with endogeneity.

Our discussion easily extends to models for panel data with random effects. In this case, we have  $N$  panels. Panel  $i = 1, \dots, N$  has observations  $t = 1, \dots, N_i$ , so we observe  $y_{it}$  with random effect  $\alpha_i$  and observation-level error  $\epsilon_{it}$ . These errors are independent of each other. So the combined error  $\xi_{it} = \alpha_i + \epsilon_{it}$  is normal with mean 0 and variance  $\sigma^2 + \sigma_\alpha^2$ , where  $\sigma_\alpha^2$  is the variance of  $\alpha_i$ . The results discussed earlier can then be applied using the combined error  $\xi_{it}$  rather than the cross-sectional error.

All predictions after xt eregress assume the panel-level random effects ( $\alpha_i$ ) are zero. Put another way, predictions condition on the random effects being set to their means.

## References

- Imbens, G. W., and W. K. Newey. 2009. Identification and estimation of triangular simultaneous equations models without additivity. *Econometrica* 77: 1481–1512. <https://doi.org/10.3982/ECTA7108>.
- Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

## Also see

[ERM] **eregress** — Extended linear regression

[ERM] **eregress predict** — predict after eregress and xteregress

[ERM] **predict treatment** — predict for treatment statistics

[ERM] **predict advanced** — predict's advanced features

[ERM] **eprobit postestimation** — Postestimation tools for eprobit and xteprobit

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

Stata, Stata Press, and Mata are registered trademarks of StataCorp LLC. Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations. StataNow and NetCourseNow are trademarks of StataCorp LLC. Other brand and product names are registered trademarks or trademarks of their respective companies. Copyright © 1985–2025 StataCorp LLC, College Station, TX, USA. All rights reserved.

For suggested citations, see the FAQ on [citing Stata documentation](#).

