Title

eprobit postestimation — Postestimation tools for eprobit and xteprobit

Postestimation commands

The following postestimation command is of special interest after `eprobit` and `xteprobit`:

Command	Description
`estat teffects`	treatment effects and potential-outcome means

The following standard postestimation commands are also available after `eprobit` and `xteprobit`:

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
*`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`	means, probabilities, treatment effects, etc.
`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.

†`suest` and the survey data `estat` commands are not available after `xteprobit`.

predict

Predictions after `eprobit` and `xteprobit` are described in

[ERM] **eprobit predict**	predict after eprobit and xteprobit
[ERM] **predict treatment**	predict for treatment statistics
[ERM] **predict advanced**	predict's advanced features

[ERM] **eprobit 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	
`pr`	probability for binary or ordinal y_j; the default
`mean`	mean
`pomean`	potential-outcome mean
`te`	treatment effect
`tet`	treatment effect on the treated
`xb`	linear prediction excluding all complications
`pr(`*a*`,`*b*`)`	$\Pr(a < y_j < b)$ for continuous y_j
`e(`*a*`,`*b*`)`	$E(y_j \| a < y_j < b)$ for continuous y_j
`ystar(`*a*`,`*b*`)`	$E(y_j^*)$, $y_j^* = \max\{a, \min(y_j, b)\}$ for continuous y_j

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

See [ERM] **Intro 7** for an overview of using `margins` and `predict` after `eprobit` and `xteprobit`. For examples using `margins`, `predict`, and `estat teffects`, see *Interpreting effects* in [ERM] **Intro 9** and see [ERM] **Example 1a**.

Methods and formulas

These methods build on the discussions in *Methods and formulas* of [ERM] **eprobit**.

Methods and formulas are presented under the following headings:

Predictions and inferences using the default asf
General prediction framework

Predictions and inferences using the default asf

In the probit model, for exogenous covariates $\mathbf{x}_i$ and endogenous covariates $\mathbf{w}_i$, we have

$$y_i = \mathbf{1}(\mathbf{x_i}\boldsymbol{\beta} + \mathbf{w_i}\boldsymbol{\beta_2} + \boldsymbol{\epsilon_i} > \mathbf{0})$$

where ϵ_i is a standard normal error.

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

$$\epsilon_i = u_i + \psi_i$$

where $u_i = \boldsymbol{\gamma\epsilon}_{2i}$ is the unobserved heterogeneity that gives rise to the endogeneity and ψ_i is an idiosyncratic error term with variance σ_ψ^2. Conditional on the covariates and the unobserved heterogeneity, for one endogenous covariate, the probability that $y_i = 1$ is

$$\Pr(y_i = 1|\mathbf{x}_i, w_i, u_i) = \Phi\left(\frac{\mathbf{x}_i\boldsymbol{\beta} + \mathbf{w}_i\boldsymbol{\beta}_2 + u_i}{\sigma_\psi}\right)$$

Default predictions and effects are computed based on the expression above. Including u_i 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 $i = 1, \ldots, N$ has observations $t = 1, \ldots, N_i$, so we observe y_{it} with random effect α_i and observation-level error ϵ_{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 $1 + \sigma_\alpha^2$, where σ_α^2 is the variance of α_i. The results discussed earlier can then be applied using the combined error ξ_{it} rather than the cross-sectional error.

All predictions after `xteprobit` assume the panel-level random effects (α_i) are zero. Put another way, predictions condition on the random effects being set to their means.

General prediction framework

In this section, we discuss the general framework for predictions made after ERMs with multiple auxiliary equations and conditioned on both the covariates and the instruments. The predictions consider the total effect of all the covariates and instruments on the outcome.

First, assume that we have a model with random effects in each equation and a panel-data structure. We have N panels. For panel $i = 1, \ldots, N$, there are N_i observations, and for $t = 1, \ldots, N_i$, we have

$$\begin{aligned} y_{1it} &= g_{1it}(\mathbf{w}_{1it}\boldsymbol{\beta}_1 + v_{1it} + u_{1i}) \\ &\vdots \\ y_{Hit} &= g_{Hit}(\mathbf{w}_{Hit}\boldsymbol{\beta}_H + v_{Hit} + u_{Hi}) \\ y_{it} = y_{Jit} &= g_{Jit}(\mathbf{w}_{Jit}\boldsymbol{\beta}_J + v_{Jit} + u_{Ji}) \end{aligned}$$

The observation-level errors v_{1it}, …, v_{Jit} are multivariate normal with mean 0 and covariance $\boldsymbol{\Sigma}$. They are independent of the panel-level errors, or random effects u_{1i}, …, u_{Ji}, which are multivariate normal with mean 0 and covariance $\boldsymbol{\Sigma}_u$. We further assume that the observation-level errors are independent within panels.

We will perform prediction conditional on the observed covariates, so we can collapse the random effects and observation-level errors. The new observation-level errors are $\xi_{jit} = v_{jit} + u_{ji}$. These errors, $\xi_{1it}, \ldots, \xi_{Jit}$, are multivariate normal with mean 0 and variance $\boldsymbol{\Sigma}_\xi = \boldsymbol{\Sigma} + \boldsymbol{\Sigma}_u$.

In the following, we will derive prediction formulas for the cross-sectional case without a panel structure, but our results will apply to the random-effects model we have just discussed, using the combined covariance $\boldsymbol{\Sigma}_\xi$ rather than the cross-sectional covariance matrix $\boldsymbol{\Sigma}$.

In the cross-sectional case, we have H auxiliary equations with endogenous outcomes $y_{1i}, \ldots, y_{Hi}$. We will treat the main outcome y_{it} as stage $J = H + 1$, so $y_{Ji} = y_{it}$. The ERMs that we fit with `eintreg`, `eoprobit`, `eprobit`, and `eregress` are triangular, so we can order the equations such that the first depends only on exogenous covariates and instruments—say, $\mathbf{w}_{1i} = \mathbf{z}_i$—and for $j = 2$, …, J, equation j depends only on the exogenous covariates and instruments $\mathbf{z}_i$ and the endogenous covariates from equation $h = j - 1$ and $y_{1i}, \ldots, y_{hi}$ below. These are stored together in $\mathbf{w}_{ji}$.

When we predict conditional probabilities for binary and ordinal outcomes, we condition on all the endogenous and exogenous covariates and instruments that affect y_{ji}. Conditional probabilities are calculated as the ratio of the joint density over the marginal density of the conditioning covariates. For binary or ordinal outcome y_{ji}, we have

$$\Pr(y_{ji} = Y | y_{1i}, \ldots, y_{(j-1)i}, \mathbf{z}_i) = \frac{f(Y, y_{1i}, \ldots, y_{(j-1)i} | \mathbf{z}_i)}{f(y_{1i}, \ldots, y_{(j-1)i} | \mathbf{z}_i)}$$

where the densities can be computed as described in [ERM] **eprobit**.

Now, suppose instead that y_{ji} is continuous. We can predict the probability that y_{ji} lies in the range (l_{ji}, u_{ji}):

$$\begin{aligned} \Pr(l_{ji}, u_{ji}) &= \Pr(l_{ji} < y_{ji} < u_{ji} | y_{1i}, \ldots, y_{(j-1)i}, \mathbf{z}_i) \\ &= \int_{(l_{ji}, u_{ji}) \times \mathbf{V}^\star_{(j-1)i}} \phi_j(v_{1i}, \ldots, v_{ji}, \boldsymbol{\Sigma}_j) dv_{ji} d\mathbf{v}^\star_{(j-1)i} \end{aligned}$$

This integral can be evaluated using the methods discussed in *Likelihood for multiequation models* in [ERM] **eprobit**.

The conditional mean of continuous outcome y_{ji} is

$$E(y_{ji}|\mathbf{w}_{ji}) = \mathbf{w}_{ji}\boldsymbol{\beta}_j + E(v_{ji}|\mathbf{w}_{ji})$$

where $\mathbf{w}_{ji}$ contains the endogenous covariates $y_{1i}, \ldots, y_{(j-1)i}$ and exogenous covariates $\mathbf{z}_i$ that affect y_{ji}.

By conditioning on the binary and ordinal endogenous covariates $y_{1i}, \ldots, y_{(j-1)i}$, the errors $v_{hi}, \ldots, v_{Ji}$ become truncated normal. Together with v_{ji}, they have a truncated multivariate distribution. So the mean of the continuous endogenous covariate is calculated using the moment formulas for the truncated multivariate normal. The first and second moments of the doubly truncated multivariate normal were derived in Manjunath and Wilhelm (2012). Tallis (1961) derived the first and second moments of the multivariate normal with one-sided truncation.

A key result in Manjunath and Wilhelm (2012) is that

$$\int_{l_1}^{u_1} \ldots \int_{l_d}^{u_d} \epsilon_f \phi_d(\boldsymbol{\epsilon}, \boldsymbol{\Sigma}) \quad d\epsilon_1 \ldots d\epsilon_d = \sum_{k=1}^{d} \sigma_{fk} \left\{ F_k(l_k) - F_k(u_k) \right\} \tag{1}$$

where the functions $F_k(\cdot)$ are defined as

$$F_k(e) = \int_{l_1}^{u_1} \ldots \int_{l_{k-1}}^{u_{k-1}} \int_{l_{k+1}}^{u_{k+1}} \phi_d(e_1, \ldots, e_{k-1}, e, e_{k+1}, \ldots, e_k, \boldsymbol{\Sigma}) de_1 \ldots de_{k-1} de_{k+1} \ldots de_d$$

The $F_k(\cdot)$ functions can be computed like the joint density in *Likelihood for multiequation models* in [ERM] **eprobit**. So we have

$$E(v_{ji}|\mathbf{w}_{ji}) = \frac{\sum_{k=j}^{J} \sigma_{jk} \left\{ F_k(l_{ki}) - F_k(u_{ki}) \right\}}{\Phi_J^*(\mathbf{l}_i, \mathbf{u}_i, \boldsymbol{\Sigma}_j)}$$

where $l_{ji} = -\infty$ and $u_{ji} = \infty$.

If there are continuous endogenous regressors in y_{1i}, ..., y_{ji}, we condition on them in calculating (1). As in the calculation of the joint density in *Likelihood for multiequation models* in [ERM] **eprobit**, we multiply by the marginal density and adjust the cutpoints and variance.

The constrained mean of continuous outcome y_{ji}, the mean of y_{ji} when y_{ji} falls between l_{ji} and u_{ji}, is

$$\begin{aligned} E(l_{ji}, u_{ji}) &= E(y_{ji}|\mathbf{w}_{ji}, l_{ji} < y_{ji} < u_{ji}) \\ &= \mathbf{w}_{ji}\boldsymbol{\beta}_j + E(v_{ji}|\mathbf{w}_{ji}, l_{ji} - \mathbf{w}_{ji}\boldsymbol{\beta}_j < \epsilon_{ji} < v_{ji} - \mathbf{w}_{ji}\boldsymbol{\beta}_j) \end{aligned}$$

We use the same method as for the unconstrained mean, with cutpoints $l_{ji} - \mathbf{w}_{ji}\boldsymbol{\beta}_j$ and $u_{ji} - \mathbf{w}_{ji}\boldsymbol{\beta}_j$ instead of $-\infty$ and ∞.

The expected value of continuous y_{ji} with censoring at l_{ji} and u_{ji} is

$$\begin{aligned} E(y_{ji}^{\star}|\mathbf{w}_{ji}) = {} & l_{ji}\mathbf{1}(\mathbf{w}_{ji}\boldsymbol{\beta}_j + \epsilon_{ji} < l_{ji}) + u_{ji}\mathbf{1}(\mathbf{w}_{ji}\boldsymbol{\beta}_j + \epsilon_{ji} > u_{ji}) \\ & + (\mathbf{w}_{ji}\boldsymbol{\beta}_j + \epsilon_{ji})\mathbf{1}(l_{ji} \le \mathbf{w}_{ji}\boldsymbol{\beta}_j + \epsilon_{ji} \le u_{ji}) \end{aligned}$$

where $y_{ji}^{\star} = \max\left\{l_{ij}, \min(y_{ij}, u_{ij})\right\}$. This can be calculated using predictions we have already discussed:

$$E(y_{ji}^{\star}|\mathbf{w}_{ji}) = \Pr(-\infty, l_{ji})l_{ji} + \Pr(l_{ji}, u_{ji})E(l_{ji}, u_{ji}) + \Pr(u_{ji}, \infty)u_{ji}$$

Sometimes, we model a continuous outcome y_{ji} that is the natural logarithm of another outcome y^e_{ji}. In this case, the conditional mean of y^e_{ji} is

$$\begin{aligned} E(y^e_{ji}|\mathbf{w}_{ji}) = E\left\{\exp(y_{ji})|\mathbf{w}_{ji}\right\} &= E\left\{\exp\left(\mathbf{w}_{ji}\boldsymbol{\beta}_j + v_{ji}\right)\mathbf{w}_{ji}\right\} \\ &= \exp\left(\mathbf{w}_{ji}\boldsymbol{\beta}_j\right) E\left\{\exp\left(v_{ji}\right)|\mathbf{w}_{ji}\right\} \end{aligned}$$

As discussed earlier, v_{ji} can be truncated normal when we condition on $\mathbf{w}_{ji}$. So the conditional expectation above is the moment-generating function of a truncated normal random variable. This function was also derived in Manjunath and Wilhelm (2012). Letting $\boldsymbol{\sigma}_j$ be the jth column of $\boldsymbol{\Sigma}_j$, we have

$$E\left\{\exp\left(v_{ji}\right)|\mathbf{w}_{ji}\right\} = \exp\left(\frac{\sigma_j^2}{2}\right)\frac{\Phi_j^*(\mathbf{l}_i - \boldsymbol{\sigma}_j, \mathbf{u}_i - \boldsymbol{\sigma}_j, \boldsymbol{\Sigma}_j)}{\Phi_j^*(\mathbf{l}_i, \mathbf{u}_i, \boldsymbol{\Sigma}_j)}$$

All the predictions above can be made after estimation by using `predict`. By also specifying either the `pr` or the `pr(`l_{ji}`,`u_{ji}`)` option in `predict`, we can obtain conditional probabilities for a binary or ordinal outcome or the conditional probability that a continuous outcome lies in the specified range (l_{ji}, u_{ji}).

By also specifying the `mean` option, we obtain the conditional mean of a continuous endogenous covariate. The `e(`l_{ji}`,`u_{ji}`)` option is used to obtain the constrained mean, and `ystar(`l_{ji}`,`u_{ji}`)` is used to obtain the expected value with censoring.

Prediction of treatment effects and potential-outcome means in models with endogenous covariates use the above formulas for the conditional mean and probabilities applied to the potential outcomes $y_{1i}, \ldots, y_{Ti}$ rather than the observed y_i. Methods and formulas for other predictions are given in the *Methods and formulas* sections of [ERM] **eoprobit**, [ERM] **eintreg**, and [ERM] **eregress**.

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.

Manjunath, B. G., and S. Wilhelm. 2012. Moments calculation for the doubly truncated multivariate normal density. Unpublished manuscript. https://arxiv.org/pdf/1206.5387.pdf.

Tallis, G. M. 1961. The moment generating function of the truncated multi-normal distribution. *Journal of the Royal Statistical Society, Series B* 23: 223–229. https://doi.org/10.1111/j.2517-6161.1961.tb00408.x.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

Also see

[ERM] **eprobit** — Extended probit regression

[ERM] **eprobit predict** — predict after eprobit and xteprobit

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

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

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

For suggested citations, see the FAQ on citing Stata documentation.