Title

stata.com

eprobit postestimation — Postestimation tools for eprobit

Postestimation commands

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

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

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

Command	Description
`contrast`	contrasts and ANOVA-style joint tests of estimates
`estat ic`	Akaike's and Schwarz's Bayesian information criteria (AIC 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
* `forecast`	dynamic forecasts and simulations
`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`	predictions, residuals, influence statistics, and other diagnostic measures
`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` and `lrtest` are not appropriate with `svy` estimation results.

predict

Predictions after `eprobit` are described in

[ERM] **eprobit predict**	predict after eprobit
[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 margins of response for probabilities, means, potential-outcome means, treatment effects, and linear predictions.

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
`pr(`*a*`,`*b*`)`	$\Pr(a < y_j < b)$ for continuous y_j
`e(`*a*`,`*b*`)`	$E(y_j \mid a < y_j < b)$ 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 6** for an overview of using `margins` and `predict` after `eprobit`. For examples using `margins`, `predict`, and `estat teffects`, see *Interpreting effects* in [ERM] **intro 8** 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:

Counterfactual predictions and inferences

In *Methods and formulas* of [ERM] **eprobit**, we discussed how treatment effects are evaluated in extended probit regression models. Here, we discuss the counterfactual framework used to evaluate the effects of other changes to covariates.

In the extended probit regression model for y_i on exogenous covariates $\mathbf{x}_i$ and $\mathbf{w}_i$, we partition each set of covariates into two groups. The exogenous covariates $\mathbf{x}_i$ are partitioned into $\mathbf{x}_i^c$ and $\mathbf{w}_i^{nc}$, where we are interested in the effect of changes in $\mathbf{x}_i^c$. Similarly, the endogenous covariates $\mathbf{w}_i$ are partitioned into $\mathbf{w}_i^c$ and $\mathbf{w}_i^{nc}$, where the effect of changes in $\mathbf{w}_i^c$ is of interest. The superscripts indicate what is a counterfactual value (c) and what is not (nc).

If $\mathbf{x}_i^c = \mathbf{a}_0$ and $\mathbf{w}_i^c = \mathbf{a}_{20}$, for covariates $\mathbf{w}_i^{nc}$ and $\mathbf{x}_i^{nc}$ we would observe the outcome

$$\begin{aligned} y_{0i} &= \mathbf{1}(\boldsymbol{\beta}_{0nc}\mathbf{x}_i^{nc} + \boldsymbol{\beta}_{20nc}\mathbf{w}_i^{nc} + \boldsymbol{\beta}_c\mathbf{a}_0 + \boldsymbol{\beta}_{2c}\mathbf{a}_{20} + \epsilon_{0i} > 0) \\ &= \mathbf{1}(\boldsymbol{\beta}_{0nc}\mathbf{x}_i^{nc} + \boldsymbol{\beta}_{20nc}\mathbf{w}_i^{nc} + \beta_{c0} + \epsilon_{0i} > 0) \end{aligned}$$

where the unobserved error ϵ_{0i} is standard normal. We treat $\boldsymbol{\beta}_c\mathbf{a}_0 + \boldsymbol{\beta}_{2c}\mathbf{a}_{20} = \beta_{c0}$ as a constant intercept, because it is the same for each value combination of the covariates $\mathbf{w}_i^{nc}$ and $\mathbf{x}_i^{nc}$ and the error ϵ_{0i}.

Similarly, if $\mathbf{x}_i^c = \mathbf{a}_1$ and $\mathbf{w}_i^c = \mathbf{a}_{21}$, for covariates $\mathbf{w}_i^{nc}$ and $\mathbf{x}_i^{nc}$ we would observe the outcome

$$\begin{aligned} y_{1i} &= \mathbf{1}(\boldsymbol{\beta}_{1nc}\mathbf{x}_i^{nc} + \boldsymbol{\beta}_{21nc}\mathbf{w}_i^{nc} + \boldsymbol{\beta}_c\mathbf{a}_1 + \boldsymbol{\beta}_{2c}\mathbf{a}_{21} + \epsilon_{1i} > 0) \\ &= \mathbf{1}(\boldsymbol{\beta}_{1nc}\mathbf{x}_i^{nc} + \boldsymbol{\beta}_{21nc}\mathbf{w}_i^{nc} + \beta_{c1} + \epsilon_{1i} > 0) \end{aligned}$$

The effect of changing $\mathbf{x}_i^c$ and $\mathbf{w}_i^c$ from $\mathbf{a}_0$ and $\mathbf{a}_{20}$ to $\mathbf{a}_1$ and $\mathbf{a}_{21}$ on y_i is the expected difference between $\mathbf{y}_{1i}$ and $\mathbf{y}_{0i}$.

To obtain this difference, we average the conditional probabilities of $\mathbf{y}_{1i}$ and $\mathbf{y}_{0i}$ as a predictive margin.

For $j = 0, 1$, we can predict the counterfactual probability for group j by using the tools discussed in *Predictions using the full model*,

$$\text{CP}_j(\mathbf{w}_i^{nc}, \mathbf{x}_i^{nc}, \mathbf{z}_i) = \Pr(y_{ji} = 1 | \mathbf{w}_i^{nc}, \mathbf{x}_i^{nc}, \mathbf{x}_i^c = \mathbf{a}_j, \mathbf{z}_i)$$

where $\mathbf{z}_i$ are instruments necessary for modeling the endogenous regressors $\mathbf{w}_i^{nc}$. By the law of iterated expectations, we have

$$E(y_{1i} - y_{0i}) = E\left\{\text{CP}_1(\mathbf{w}_i^{nc}, \mathbf{x}_i^{nc}, \mathbf{z}_i)\right\} - E\left\{\text{CP}_0(\mathbf{w}_i^{nc}, \mathbf{x}_i^{nc}, \mathbf{z}_i)\right\}$$

So the effect of changing $\mathbf{x}_i^c$ and $\mathbf{w}_i^c$ from $\mathbf{a}_0$ and $\mathbf{a}_{20}$ to $\mathbf{a}_1$ and $\mathbf{a}_{21}$ can be estimated as a predictive margin on the counterfactual probabilities.

We can use `predict` with the `fix()` and `target()` options to predict the counterfactual probabilities. The `fix()` option is used to indicate the endogenous covariates in $\mathbf{w}_i^c$. The `target()` option can be used to set the counterfactual values a_j and a_{2j} of $\mathbf{x}_i^c$ and $\mathbf{w}_i^c$.

When $\mathbf{w}_i^c$ corresponds to a single ordinal or binary regressor, the difference in counterfactual probabilities corresponds to a treatment effect of $\mathbf{w}_i^c$. We can also evaluate the structural effect of a change in $\mathbf{w}_i^c$ and $\mathbf{x}_i^c$, conditioned on $\mathbf{w}_i^c$. This effect is analogous to the treatment effect on the treated discussed in *Methods and formulas* of [ERM] **eprobit**. We are conditioning the effect on some base value for $\mathbf{w}_i^c$, $\mathbf{w}_i^c = \mathbf{b}$.

Now, the counterfactual probabilities are conditioned on $\mathbf{w}_i^c = \mathbf{b}$. So for $j = 0, 1$, we have

$$\text{CP}_{bj}(\mathbf{w}_i^{nc}, \mathbf{b}, \mathbf{x}_i^{nc}, \mathbf{z}_i) = \Pr(y_{ji} = 1|\mathbf{w}_i^{nc}, \mathbf{w}_i^c = \mathbf{b}, \mathbf{x}_i^{nc}, \mathbf{x}_i^c = \mathbf{a}_j, \mathbf{z}_{bi})$$

where $\mathbf{z}_{bi}$ are instruments necessary for modeling the endogenous regressors $\mathbf{w}_i^{nc}$ and $\mathbf{w}_i^c$. This counterfactual probability can be evaluated using the tools discussed in *Predictions using the full model*.

By the law of iterated expectations, we have

$$\begin{aligned} E(y_{1i} - y_{0i}|\mathbf{w}_i^c = \mathbf{b}) =& E\left\{\text{CP}_{b1}(\mathbf{w}_i^{nc}, \mathbf{w}_i^c = \mathbf{b}, \mathbf{x}_i^{nc}, \mathbf{z}_i)|\mathbf{w}_i^c = \mathbf{b}\right\} - \\ & E\left\{\text{CP}_{b0}(\mathbf{w}_i^{nc}, \mathbf{w}_i^c = \mathbf{b}, \mathbf{x}_i^{nc}, \mathbf{z}_i)|\mathbf{w}_i^c = \mathbf{b}\right\} \end{aligned}$$

So the effect of changing $\mathbf{x}_i^c$ and $\mathbf{w}_i^c$ from $\mathbf{a}_0$ and $\mathbf{a}_{20}$ to $\mathbf{a}_1$ and $\mathbf{a}_{21}$ conditioned on $\mathbf{w}_i^c = \mathbf{b}$ can be estimated as a predictive margin on the counterfactual probabilities.

The base values $\mathbf{b}$ for $\mathbf{w}_i^c$ are specified in the `base()` option. As before, `target()` can be used to specify the counterfactual values for $\mathbf{x}_i^c$ and $\mathbf{w}_i^c$.

When $\mathbf{x}_i^c = \mathbf{x}_i$ and $\mathbf{w}_i^c = \mathbf{w}_i$, the counterfactual probability matches the average structural probability (ASP). Applying the average structural function (ASF) discussed by Blundell and Powell (2003), Blundell and Powell (2004), Wooldridge (2005), and Wooldridge (2014) to a conditional probability on the covariates and unobserved endogenous error produces the ASP.

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

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

where ϵ_i is a standard normal error.

The ASP provides a structural interpretation of $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_2$ when the $\mathbf{w}_i$ are correlated with ϵ_i. 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 $u_i = \boldsymbol{\gamma}\boldsymbol{\epsilon}_{2i}$ is the unobserved heterogeneity that gives rise to the endogeneity and ψ_i is an error term with variance σ_ψ^2. Conditional on the covariates and the unobserved heterogeneity, the probability that $y_i = 1$ is

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

Because u_i is an unobserved random variable, this conditional probability is not observable. Integrating out the u_i, just like we do with random effects in panel-data models, produces the ASP,

$$\text{ASP}(\mathbf{x}_i^0, \mathbf{w}_i^0) = \int \Pr(y_i = 1 | \mathbf{x}_i^0, \mathbf{w}_i^0, u_i) f(u_i) du_i$$

where $f(u_i)$ is the marginal distribution of u_i, and $\mathbf{x}_i^0$ and $\mathbf{w}_i^0$ are given covariate values.

Predictions using the full model

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. See *Counterfactual predictions and inferences* for a discussion of predictions that may not involve all the covariates and instruments.

Suppose that we have H auxiliary equations with endogenous outcomes $y_{1i}, \ldots, y_{Hi}$. We will treat the main outcome y_i as stage $J = H + 1$, so $y_{Ji} = y_i$. 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 $j = 2, \ldots, J$, equation j depends only on the exogenous covariates and instruments $\mathbf{z}_i$ and the endogenous covariates from equation $h = j - 1$ and below $y_{1i}, \ldots, y_{hi}$. 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 $y_{1i},\dots,y_{(j-1)i}$, the errors $v_{hi},\dots,v_{Ji}$ 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

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

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

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

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 $y_{1i},\dots,y_{ji}$, 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 ∞.

Finally, 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:

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

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

Blundell, R. W., and J. L. Powell. 2003. Endogeity in nonparametric and semiparametric regression models. In *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, ed. M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, vol. 2, 312–357. Cambridge: Cambridge University Press.

——. 2004. Endogeneity in semiparametric binary response models. *Review of Economic Studies* 71: 655–679.

Manjunath, B. G., and S. Wilhelm. 2012. Moments calculation for the doubly truncated multivariate normal density. 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.

Wooldridge, J. M. 2005. Unobserved heterogeneity and estimation of average partial effects. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. D. W. K. Andrews and J. H. Stock, 27–55. New York: Cambridge University Press.

——. 2014. Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables. *Journal of Econometrics* 182: 226–234.

Also see

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

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

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

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

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