# Title

stata.com

**ivprobit postestimation —** Postestimation tools for ivprobit



## Description

The following postestimation commands are of special interest after `ivprobit`:

| Command | Description |
|---|---|
| `estat classification` | report various summary statistics, including the classification table |
| `lroc` | compute area under ROC curve and graph the curve |
| `lsens` | graph sensitivity and specificity versus probability cutoff |

These commands are not appropriate after the two-step estimator or the `svy` prefix.

The following standard postestimation commands are also available:

| Command | Description |
|---|---|
| `contrast` | contrasts and ANOVA-style joint tests of estimates |
| `estat ic`[1] | 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`[2] | dynamic forecasts and simulations |
| `hausman` | Hausman's specification test |
| `lincom` | point estimates, standard errors, testing, and inference for linear combinations of coefficients |
| `lrtest`[3] | likelihood-ratio test; not available with two-step estimator |
| `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`[1] | seemingly unrelated estimation |
| `test` | Wald tests of simple and composite linear hypotheses |
| `testnl` | Wald tests of nonlinear hypotheses |

[1] `estat ic` and `suest` are not appropriate after `ivprobit, twostep`.

[2] `forecast` is not appropriate with `svy` estimation results or after `ivprobit, twostep`.

[3] `lrtest` is not appropriate with `svy` estimation results.

# Syntax for predict

*After ML or twostep*

```
predict [type] newvar [if] [in] [, statistic rules asif]
```

*After ML*

```
predict [type] {stub* | newvarlist} [if] [in], scores
```

| *statistic* | Description |
|---|---|
| Main | |
| `xb` | linear prediction; the default |
| `stdp` | standard error of the linear prediction |
| `pr` | probability of a positive outcome; not available with two-step estimator |

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

# Menu for predict

Statistics > Postestimation > Predictions, residuals, etc.

# Options for predict

Main

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

`stdp` calculates the standard error of the linear prediction.

`pr` calculates the probability of a positive outcome. `pr` is not available with the two-step estimator.

`rules` requests that Stata use any rules that were used to identify the model when making the prediction. By default, Stata calculates missing for excluded observations. `rules` is not available with the two-step estimator.

`asif` requests that Stata ignore the rules and the exclusion criteria and calculate predictions for all observations possible using the estimated parameters from the model. `asif` is not available with the two-step estimator.

`scores`, not available with `twostep`, calculates equation-level score variables.

For models with one endogenous regressor, four new variables are created.

The first new variable will contain $\partial \ln L/\partial(\boldsymbol{z}_i\boldsymbol{\delta})$.

The second new variable will contain $\partial \ln L/\partial(\boldsymbol{x}_i\boldsymbol{\Pi})$.

The third new variable will contain $\partial \ln L/\partial \operatorname{atanh} \rho$.

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

For models with $p$ endogenous regressors, $p + \{(p+1)(p+2)\}/2$ new variables are created.

The first new variable will contain $\partial \ln L/\partial(\boldsymbol{z}_i\boldsymbol{\delta})$.

The second through $(p+1)$th new variables will contain $\partial \ln L/\partial(\boldsymbol{x}_i\boldsymbol{\Pi}_k)$, $k = 1, \ldots, p$, where $\boldsymbol{\Pi}_k$ is the $k$th column of $\boldsymbol{\Pi}$.

The remaining score variables will contain the partial derivatives of $\ln L$ with respect to $s_{21}$, $s_{31}$, ..., $s_{p+1,1}$, $s_{22}$, ..., $s_{p+1,2}$, ..., $s_{p+1,p+1}$, where $s_{m,n}$ denotes the $(m,n)$ element of the Cholesky decomposition of the error covariance matrix.

## Remarks and examples

stata.com

Remarks are presented under the following headings:



### Marginal effects

▷ Example 1

We can obtain marginal effects by using the `margins` command after `ivprobit`. We will calculate average marginal effects by using the labor-supply model of example 1 in [R] **ivprobit**.

```
. use http://www.stata-press.com/data/r13/laborsup
. ivprobit fem_work fem_educ kids (other_inc = male_educ)
  (output omitted)
. margins, dydx(*) predict(pr)
Average marginal effects                        Number of obs   =        500
Model VCE    : OIM
Expression   : Probability of positive outcome, predict(pr)
dy/dx w.r.t. : other_inc fem_educ kids male_educ
```

| | dy/dx | Delta-method Std. Err. | z | P>\|z\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| other_inc | -.014015 | .0009836 | -14.25 | 0.000 | -.0159428 | -.0120872 |
| fem_educ | .0545129 | .0066007 | 8.26 | 0.000 | .0415758 | .06745 |
| kids | -.0470199 | .0123397 | -3.81 | 0.000 | -.0712052 | -.0228346 |
| male_educ | 0 | (omitted) | | | | |

Here we see that a \$1,000 increase in `other_inc` leads to an average decrease of 0.014 in the probability that the woman has a job. `male_edu` has no effect because it appears only as an instrument. ◁

### Obtaining predicted values

After fitting your model with `ivprobit`, you can obtain the linear prediction and its standard error for both the estimation sample and other samples by using the `predict` command; see **[U] 20 Estimation and postestimation commands** and [R] **predict**. If you had used the maximum likelihood estimator, you could also obtain the probability of a positive outcome.

`predict`'s `pr` option calculates the probability of a positive outcome, remembering any rules used to identify the model, and calculates missing for excluded observations. `predict`'s `rules` option uses the rules in predicting probabilities, whereas `predict`'s `asif` option ignores both the rules and the exclusion criteria and calculates probabilities for all possible observations by using the estimated parameters from the model. See *Obtaining predicted values* in [R] **probit postestimation** for an example.

## Methods and formulas

The linear prediction is calculated as $\boldsymbol{z}_i\widehat{\boldsymbol{\delta}}$, where $\widehat{\boldsymbol{\delta}}$ is the estimated value of $\boldsymbol{\delta}$, and $\boldsymbol{z}_i$ and $\boldsymbol{\delta}$ are defined in (1*a*) of [R] **ivprobit**. The probability of a positive outcome is $\Phi(\boldsymbol{z}_i\widehat{\boldsymbol{\delta}})$, where $\Phi(\cdot)$ is the standard normal distribution function.

## Also see

[R] **ivprobit** — Probit model with continuous endogenous regressors

[R] **estat classification** — Classification statistics and table

[R] **lroc** — Compute area under ROC curve and graph the curve

[R] **lsens** — Graph sensitivity and specificity versus probability cutoff

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