

**biprobit** — Bivariate probit regression

<a href="#">Description</a>	<a href="#">Quick start</a>	<a href="#">Menu</a>	<a href="#">Syntax</a>
<a href="#">Options</a>	<a href="#">Remarks and examples</a>	<a href="#">Stored results</a>	<a href="#">Methods and formulas</a>
<a href="#">References</a>	<a href="#">Also see</a>		

## Description

`biprobit` fits maximum-likelihood two-equation probit models—either a bivariate probit or a seemingly unrelated probit (limited to two equations).

## Quick start

Bivariate probit regression of  $y_1$  and  $y_2$  on  $x_1$

```
biprobit y1 y2 x1
```

Bivariate probit regression of  $y_1$  and  $y_2$  on  $x_1$ ,  $x_2$ , and  $x_3$

```
biprobit y1 y2 x1 x2 x3
```

Constrain the coefficients for  $x_1$  to equality in both equations

```
constraint define 1 _b[y1:x1] = _b[y2:x1]  
biprobit y1 y2 x1 x2 x3, constraints(1)
```

Seemingly unrelated bivariate probit regression

```
biprobit (y1 = x1 x2 x3) (y2 = x1 x2)
```

With robust standard errors

```
biprobit (y1 = x1 x2 x3) (y2 = x1 x2), vce(robust)
```

Poirier partial observability model with `difficult` option

```
biprobit (y1 = x1 x2) (y2 = x2 x3), partial difficult
```

## Menu

### **biprobit**

Statistics > Binary outcomes > Bivariate probit regression

### **Seemingly unrelated biprobit**

Statistics > Binary outcomes > Seemingly unrelated bivariate probit regression

## Syntax

*Bivariate probit regression*

```
biprobit depvar1 depvar2 [indepvars] [if] [in] [weight] [, options]
```

*Seemingly unrelated bivariate probit regression*

```
biprobit equation1 equation2 [if] [in] [weight] [, su_options]
```

where *equation*<sub>1</sub> and *equation*<sub>2</sub> are specified as

```
( [eqname:] depvar [=] [indepvars] [, noconstant offset(varname)] )
```

<i>options</i>	Description
----------------	-------------

### Model

<u>noconstant</u>	suppress constant term
<u>partial</u>	fit partial observability model
<u>offset1</u> ( <i>varname</i> )	offset variable for first equation
<u>offset2</u> ( <i>varname</i> )	offset variable for second equation
<u>constraints</u> ( <i>constraints</i> )	apply specified linear constraints

### SE/Robust

<u>vce</u> ( <i>vcetype</i> )	<i>vcetype</i> may be <u>oim</u> , <u>robust</u> , <u>cluster</u> <i>clustvar</i> , <u>opg</u> , <u>bootstrap</u> , or <u>jackknife</u>
-------------------------------	---

### Reporting

<u>level</u> (#)	set confidence level; default is <u>level</u> (95)
<u>lrmodel</u>	perform the likelihood-ratio model test instead of the default Wald test
<u>nocnsreport</u>	do not display constraints
<u>display_options</u>	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling

### Maximization

<u>maximize_options</u>	control the maximization process; seldom used
<u>collinear</u>	keep collinear variables
<u>coeflegend</u>	display legend instead of statistics

<i>su_options</i>	Description
Model	
<code>partial</code>	fit partial observability model
<code>constraints(constraints)</code>	apply specified linear constraints
SE/Robust	
<code>vce(vctype)</code>	<i>vctype</i> may be <code>oim</code> , <code>robust</code> , <code>cluster clustvar</code> , <code>opg</code> , <code>bootstrap</code> , or <code>jackknife</code>
Reporting	
<code>level(#)</code>	set confidence level; default is <code>level(95)</code>
<code>lrmodel</code>	perform the likelihood-ratio model test instead of the default Wald test
<code>nocnsreport</code>	do not display constraints
<code>display_options</code>	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
Maximization	
<code>maximize_options</code>	control the maximization process; seldom used
<code>collinear</code>	keep collinear variables
<code>coeflegend</code>	display legend instead of statistics

*indepvars* may contain factor variables; see [U] 11.4.3 Factor variables.

*devar1*, *devar2*, *indepvars*, and *devar* may contain time-series operators; see [U] 11.4.4 Time-series varlists.

`bayes`, `bootstrap`, `by`, `collect`, `fp`, `jackknife`, `rolling`, `statsby`, and `svy` are allowed; see [U] 11.1.10 Prefix commands. For more details, see [BAYES] bayes: biprobit.

Weights are not allowed with the `bootstrap` prefix; see [R] bootstrap.

`vce()`, `lrmodel`, and `weights` are not allowed with the `svy` prefix; see [SVY] svy.

`pweights`, `fweights`, and `iweights` are allowed; see [U] 11.1.6 weight.

`collinear` and `coeflegend` do not appear in the dialog box.

See [U] 20 Estimation and postestimation commands for more capabilities of estimation commands.

## Options

Model

`noconstant`; see [R] Estimation options.

`partial` specifies that the partial observability model be fit. This particular model commonly has poor convergence properties, so we recommend that you use the `difficult` option if you want to fit the Poirier partial observability model; see [R] Maximize.

This model computes the product of the two dependent variables so that you do not have to replace each with the product.

`offset1(varname)`, `offset2(varname)`, `constraints(constraints)`; see [R] Estimation options.

## SE/Robust

`vce(vctype)` specifies the type of standard error reported, which includes types that are derived from asymptotic theory (`oim`, `opg`), that are robust to some kinds of misspecification (`robust`), that allow for intragroup correlation (`cluster clustvar`), and that use bootstrap or jackknife methods (`bootstrap`, `jackknife`); see [R] [vce\\_option](#).

## Reporting

`level(#)`, `lrmodel`, `nocnsreport`; see [R] [Estimation options](#).

`display_options`: `nocl`, `nopvalues`, `noomitted`, `vsquish`, `noemptycells`, `baselevels`, `allbaselevels`, `nofvlabel`, `fvwrap(#)`, `fvwrapon(style)`, `cformat(%fmt)`, `pformat(%fmt)`, `sformat(%fmt)`, and `nolstretch`; see [R] [Estimation options](#).

## Maximization

`maximize_options`: `difficult`, `technique(algorithm_spec)`, `iterate(#)`, `[no]log`, `trace`, `gradient`, `showstep`, `hessian`, `showtolerance`, `tolerance(#)`, `ltolerance(#)`, `nrtolerance(#)`, `nonrtolerance`, and `from(init_specs)`; see [R] [Maximize](#). These options are seldom used.

Setting the optimization type to `technique(bhhh)` resets the default `vctype` to `vce(opg)`.

The following options are available with `biprobit` but are not shown in the dialog box:

`collinear`, `coeflegend`; see [R] [Estimation options](#).

## Remarks and examples

[stata.com](http://www.stata.com)

For a good introduction to the bivariate probit models, see [Greene \(2018, sec. 17.9\)](#) and [Pindyck and Rubinfeld \(1998\)](#). [Poirier \(1980\)](#) explains the partial observability model. [Van de Ven and Van Pragg \(1981\)](#) explain the probit model with sample selection; see [R] [heckprobit](#) for details.

### ► Example 1

We use the data from [Pindyck and Rubinfeld \(1998, 332\)](#). In this dataset, the variables are whether children attend private school (`private`), number of years the family has been at the present residence (`years`), log of property tax (`logptax`), log of income (`loginc`), and whether the head of the household voted for an increase in property taxes (`vote`).

We wish to model the bivariate outcomes of whether children attend private school and whether the head of the household voted for an increase in property tax based on the other covariates.

```

. use https://www.stata-press.com/data/r17/school
. biprobit private vote years logptax loginc
Fitting comparison equation 1:
Iteration 0:  log likelihood = -31.967097
Iteration 1:  log likelihood = -31.452424
Iteration 2:  log likelihood = -31.448958
Iteration 3:  log likelihood = -31.448958
Fitting comparison equation 2:
Iteration 0:  log likelihood = -63.036914
Iteration 1:  log likelihood = -58.534843
Iteration 2:  log likelihood = -58.497292
Iteration 3:  log likelihood = -58.497288
Comparison:   log likelihood = -89.946246
Fitting full model:
Iteration 0:  log likelihood = -89.946246
Iteration 1:  log likelihood = -89.258897
Iteration 2:  log likelihood = -89.254028
Iteration 3:  log likelihood = -89.254028
Bivariate probit regression
Log likelihood = -89.254028
Number of obs = 95
Wald chi2(6) = 9.59
Prob > chi2 = 0.1431

```

	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
<b>private</b>						
years	-.0118884	.0256778	-0.46	0.643	-.0622159	.0384391
logptax	-.1066962	.6669782	-0.16	0.873	-1.413949	1.200557
loginc	.3762037	.5306484	0.71	0.478	-.663848	1.416255
_cons	-4.184694	4.837817	-0.86	0.387	-13.66664	5.297253
<b>vote</b>						
years	-.0168561	.0147834	-1.14	0.254	-.0458309	.0121188
logptax	-1.288707	.5752266	-2.24	0.025	-2.416131	-.1612839
loginc	.998286	.4403565	2.27	0.023	.1352031	1.861369
_cons	-.5360573	4.068509	-0.13	0.895	-8.510188	7.438073
/athrho	-.2764525	.2412099	-1.15	0.252	-.7492153	.1963102
rho	-.2696186	.2236753			-.6346806	.1938267

```
LR test of rho=0: chi2(1) = 1.38444 Prob > chi2 = 0.2393
```

The output shows several iteration logs. The first iteration log corresponds to running the univariate probit model for the first equation, and the second log corresponds to running the univariate probit for the second model. If  $\rho = 0$ , the sum of the log likelihoods from these two models will equal the log likelihood of the bivariate probit model; this sum is printed in the iteration log as the comparison log likelihood.

The final iteration log is for fitting the full bivariate probit model. A likelihood-ratio test of the log likelihood for this model and the comparison log likelihood is presented at the end of the output. If we had specified the `vce(robust)` option, this test would be presented as a Wald test instead of as a likelihood-ratio test.

We could have fit the same model by using the seemingly unrelated syntax as

```
. biprobit (private=years logptax loginc) (vote=years logptax loginc)
```

## Stored results

`biprobit` stores the following in `e()`:

### Scalars

<code>e(N)</code>	number of observations
<code>e(k)</code>	number of parameters
<code>e(k_eq)</code>	number of equations in <code>e(b)</code>
<code>e(k_aux)</code>	number of auxiliary parameters
<code>e(k_eq_model)</code>	number of equations in overall model test
<code>e(k_dv)</code>	number of dependent variables
<code>e(df_m)</code>	model degrees of freedom
<code>e(ll)</code>	log likelihood
<code>e(ll_0)</code>	log likelihood, constant-only model ( <code>lrmmodel</code> only)
<code>e(ll_c)</code>	log likelihood, comparison model
<code>e(N_clust)</code>	number of clusters
<code>e(chi2)</code>	$\chi^2$
<code>e(chi2_c)</code>	$\chi^2$ for comparison test
<code>e(p)</code>	$p$ -value for model test
<code>e(rho)</code>	$\rho$
<code>e(rank)</code>	rank of <code>e(V)</code>
<code>e(rank0)</code>	rank of <code>e(V)</code> for constant-only model
<code>e(ic)</code>	number of iterations
<code>e(rc)</code>	return code
<code>e(converged)</code>	1 if converged, 0 otherwise

### Macros

<code>e(cmd)</code>	<code>biprobit</code>
<code>e(cmdline)</code>	command as typed
<code>e(depvar)</code>	names of dependent variables
<code>e(wtype)</code>	weight type
<code>e(wexp)</code>	weight expression
<code>e(title)</code>	title in estimation output
<code>e(clustvar)</code>	name of cluster variable
<code>e(offset1)</code>	offset for first equation
<code>e(offset2)</code>	offset for second equation
<code>e(chi2type)</code>	Wald or LR; type of model $\chi^2$ test
<code>e(chi2_ct)</code>	Wald or LR; type of model $\chi^2$ test corresponding to <code>e(chi2_c)</code>
<code>e(vce)</code>	<code>vcetype</code> specified in <code>vce()</code>
<code>e(vcetype)</code>	title used to label Std. err.
<code>e(opt)</code>	type of optimization
<code>e(which)</code>	max or min; whether optimizer is to perform maximization or minimization
<code>e(ml_method)</code>	type of <code>ml</code> method
<code>e(user)</code>	name of likelihood-evaluator program
<code>e(technique)</code>	maximization technique
<code>e(properties)</code>	<code>b V</code>
<code>d(predict)</code>	program used to implement <code>predict</code>
<code>e(marginsok)</code>	predictions allowed by <code>margins</code>
<code>e(marginsnotok)</code>	predictions disallowed by <code>margins</code>
<code>e(asbalanced)</code>	factor variables <code>fvset</code> as <code>asbalanced</code>
<code>e(asobserved)</code>	factor variables <code>fvset</code> as <code>asobserved</code>

### Matrices

<code>e(b)</code>	coefficient vector
<code>e(Cns)</code>	constraints matrix
<code>e(ilog)</code>	iteration log (up to 20 iterations)
<code>e(gradient)</code>	gradient vector
<code>e(V)</code>	variance-covariance matrix of the estimators
<code>e(V_modelbased)</code>	model-based variance

### Functions

<code>e(sample)</code>	marks estimation sample
------------------------	-------------------------

In addition to the above, the following is stored in `r()`:

Matrices	
<code>r(table)</code>	matrix containing the coefficients with their standard errors, test statistics, $p$ -values, and confidence intervals

Note that results stored in `r()` are updated when the command is replayed and will be replaced when any `r-class` command is run after the estimation command.

## Methods and formulas

The log likelihood,  $\ln L$ , is given by

$$\begin{aligned}\xi_j^\beta &= x_j\beta + \text{offset}_j^\beta \\ \xi_j^\gamma &= z_j\gamma + \text{offset}_j^\gamma \\ q_{1j} &= \begin{cases} 1 & \text{if } y_{1j} \neq 0 \\ -1 & \text{otherwise} \end{cases} \\ q_{2j} &= \begin{cases} 1 & \text{if } y_{2j} \neq 0 \\ -1 & \text{otherwise} \end{cases} \\ \rho_j^* &= q_{1j}q_{2j}\rho \\ \ln L &= \sum_{j=1}^n w_j \ln \Phi_2 \left( q_{1j}\xi_j^\beta, q_{2j}\xi_j^\gamma, \rho_j^* \right)\end{aligned}$$

where  $\Phi_2()$  is the cumulative bivariate normal distribution function (with mean  $[0 \ 0]'$ ) and  $w_j$  is an optional weight for observation  $j$ . This derivation assumes that

$$\begin{aligned}y_{1j}^* &= x_j\beta + \epsilon_{1j} + \text{offset}_j^\beta \\ y_{2j}^* &= z_j\gamma + \epsilon_{2j} + \text{offset}_j^\gamma \\ E(\epsilon_1) &= E(\epsilon_2) = 0 \\ \text{Var}(\epsilon_1) &= \text{Var}(\epsilon_2) = 1 \\ \text{Cov}(\epsilon_1, \epsilon_2) &= \rho\end{aligned}$$

where  $y_{1j}^*$  and  $y_{2j}^*$  are the unobserved latent variables; instead, we observe only  $y_{ij} = 1$  if  $y_{ij}^* > 0$  and  $y_{ij} = 0$  otherwise (for  $i = 1, 2$ ).

In the maximum likelihood estimation,  $\rho$  is not directly estimated, but  $\text{atanh } \rho$  is

$$\text{atanh } \rho = \frac{1}{2} \ln \left( \frac{1 + \rho}{1 - \rho} \right)$$

From the form of the likelihood, if  $\rho = 0$ , then the log likelihood for the bivariate probit models is equal to the sum of the log likelihoods of the two univariate probit models. A likelihood-ratio test may therefore be performed by comparing the likelihood of the full bivariate model with the sum of the log likelihoods for the univariate probit models.

This command supports the Huber/White/sandwich estimator of the variance and its clustered version using `vce(robust)` and `vce(cluster clustvar)`, respectively. See [P] [\\_robust](#), particularly [Maximum likelihood estimators](#) and [Methods and formulas](#).

`biprobit` also supports estimation with survey data. For details on VCEs with survey data, see [SVY] [Variance estimation](#).

## References

- De Luca, G. 2008. SNP and SML estimation of univariate and bivariate binary-choice models. *Stata Journal* 8: 190–220.
- Greene, W. H. 2018. *Econometric Analysis*. 8th ed. New York: Pearson.
- Heckman, J. 1979. Sample selection bias as a specification error. *Econometrica* 47: 153–161. <https://doi.org/10.2307/1912352>.
- Hernández-Alava, M., and S. Pudney. 2016. bicop: A command for fitting bivariate ordinal regressions with residual dependence characterized by a copula function and normal mixture marginals. *Stata Journal* 16: 159–184.
- Lokshin, M., and Z. Sajaia. 2011. Impact of interventions on discrete outcomes: Maximum likelihood estimation of the binary choice models with binary endogenous regressors. *Stata Journal* 11: 368–385.
- Mullahy, J. 2016. Estimation of multivariate probit models via bivariate probit. *Stata Journal* 16: 37–51.
- Pindyck, R. S., and D. L. Rubinfeld. 1998. *Econometric Models and Economic Forecasts*. 4th ed. New York: McGraw–Hill.
- Poirier, D. J. 1980. Partial observability in bivariate probit models. *Journal of Econometrics* 12: 209–217. [https://doi.org/10.1016/0304-4076\(80\)90007-X](https://doi.org/10.1016/0304-4076(80)90007-X).
- Van de Ven, W. P. M. M., and B. M. S. Van Pragg. 1981. The demand for deductibles in private health insurance: A probit model with sample selection. *Journal of Econometrics* 17: 229–252. [https://doi.org/10.1016/0304-4076\(81\)90028-2](https://doi.org/10.1016/0304-4076(81)90028-2).

## Also see

- [R] **biprobit postestimation** — Postestimation tools for biprobit
- [R] **mprobit** — Multinomial probit regression
- [R] **probit** — Probit regression
- [BAYES] **bayes: biprobit** — Bayesian bivariate probit regression
- [SVY] **svy estimation** — Estimation commands for survey data
- [U] **20 Estimation and postestimation commands**