

bayes: heckprobit — Bayesian probit model with sample selection

[Description](#)[Remarks and examples](#)[Quick start](#)[Stored results](#)[Menu](#)[Methods and formulas](#)[Syntax](#)[Also see](#)

Description

`bayes: heckprobit` fits a Bayesian sample-selection probit regression to a partially observed binary outcome; see [\[BAYES\] bayes](#) and [\[R\] heckprobit](#) for details.

Quick start

Bayesian sample-selection probit regression of `y` on `x1` and `x2`, using `z1` and `z2` to model selection and using default normal priors for regression coefficients and atanh-correlation

```
bayes: heckprobit y x1 x2, select(z1 z2)
```

Use a standard deviation of 10 instead of 100 for the default normal priors

```
bayes, normalprior(10): heckprobit y x1 x2, select(z1 z2)
```

Use uniform priors for the slopes and a normal prior for the intercept of the main regression

```
bayes, prior({y: x1 x2}, uniform(-10,10)) ///
prior({y:_cons}, normal(0,10)): heckprobit y x1 x2, select(z1 z2)
```

Save simulation results to `simdata.dta` and use a random-number seed for reproducibility

```
bayes, saving(simdata) rseed(123):, ///
heckprobit y x1 x2, select(z1 z2)
```

Specify 20,000 MCMC samples, set length of the burn-in period to 5,000, and request that a dot be displayed every 500 simulations

```
bayes, mcmcsample(20000) burnin(5000) dots(500):, ///
heckprobit y x1 x2, select(z1 z2)
```

In the above, request that the 90% HPD credible interval be displayed instead of the default 95% equal-tailed credible interval

```
bayes, clevel(90) hpd
```

Also see [Quick start](#) in [\[BAYES\] bayes](#) and [Quick start](#) in [\[R\] heckprobit](#).

Menu

Statistics > Binary outcomes > Bayesian regression > Probit model with sample selection

Syntax

```
bayes [ , bayesopts ] : heckprobit depvar indepvars [ if ] [ in ] [ weight ] ,
      select( [ depvars = ] varlists [ , noconstant offset(varnameo) ] ) [ options ]
```

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

Model	
* <u>select</u> ()	specify selection equation: dependent and independent variables; whether to have constant term and offset variable
<u>noconstant</u>	suppress constant term
<u>offset</u> (<i>varname</i>)	include <i>varname</i> in model with coefficient constrained to 1
<u>collinear</u>	keep collinear variables

Reporting	
<u>display_options</u>	control spacing, line width, and base and empty cells
<u>level</u> (#)	set credible level; default is <code>level(95)</code>

* `select()` is required.

The full specification is `select([depvars =] varlists [, noconstant offset(varnameo)])`.
indepvars and *varlist_s* may contain factor variables; see [U] 11.4.3 **Factor variables**.

depvar, *indepvars*, *varlist_s*, and *depvar_s* may contain time-series operators; see [U] 11.4.4 **Time-series varlists**.

fweights are allowed; see [U] 11.1.6 **weight**.

`bayes: heckprobit`, `level()` is equivalent to `bayes, clevel(): heckprobit`.

For a detailed description of *options*, see *Options* in [R] **heckprobit**.

<i>bayesopts</i>	Description
------------------	-------------

Priors

* <u>normalprior</u> (#)	specify standard deviation of default normal priors for regression coefficients and atanh-correlation; default is <code>normalprior(100)</code>
<u>prior</u> (<i>priorspec</i>)	prior for model parameters; this option may be repeated
<u>dryrun</u>	show model summary without estimation

Simulation

<u>mcmcsize</u> (#)	MCMC sample size; default is <code>mcmcsize(10000)</code>
<u>burnin</u> (#)	burn-in period; default is <code>burnin(2500)</code>
<u>thinning</u> (#)	thinning interval; default is <code>thinning(1)</code>
<u>rseed</u> (#)	random-number seed
<u>exclude</u> (<i>paramref</i>)	specify model parameters to be excluded from the simulation results

Blocking

* <u>blocksize</u> (#)	maximum block size; default is <code>blocksize(50)</code>
<u>block</u> (<i>paramref</i> [, <i>blockopts</i>])	specify a block of model parameters; this option may be repeated
<u>blocksummary</u>	display block summary
* <u>noblocking</u>	do not block parameters by default

Initialization

<code>initial(<i>initspec</i>)</code>	initial values for model parameters
<code>nomleinitial</code>	suppress the use of maximum likelihood estimates as starting values
<code>initransom</code>	specify random initial values
<code>initsummary</code>	display initial values used for simulation
* <code>noisily</code>	display output from the estimation command during initialization

Adaptation

<code>adaptation(<i>adaptopts</i>)</code>	control the adaptive MCMC procedure
<code>scale(#)</code>	initial multiplier for scale factor; default is <code>scale(2.38)</code>
<code>covariance(<i>cov</i>)</code>	initial proposal covariance; default is the identity matrix

Reporting

<code>clevel(#)</code>	set credible interval level; default is <code>clevel(95)</code>
<code>hpd</code>	display HPD credible intervals instead of the default equal-tailed credible intervals
<code>eform[(<i>string</i>)]</code>	report exponentiated coefficients and, optionally, label as <i>string</i>
<code>batch(#)</code>	specify length of block for batch-means calculations; default is <code>batch(0)</code>
<code>saving(<i>filename</i>[, <i>replace</i>])</code>	save simulation results to <i>filename.dta</i>
<code>nomodelsummary</code>	suppress model summary
<code>[no]dots</code>	suppress dots or display dots every 100 iterations and iteration numbers every 1,000 iterations; default is <code>nodots</code>
<code>dots(#[, <i>every</i>(#)])</code>	display dots as simulation is performed
<code>[no]show(<i>paramref</i>)</code>	specify model parameters to be excluded from or included in the output
<code>notable</code>	suppress estimation table
<code>noheader</code>	suppress output header
<code>title(<i>string</i>)</code>	display <i>string</i> as title above the table of parameter estimates
<code>display_options</code>	control spacing, line width, and base and empty cells

Advanced

<code>search(<i>search_options</i>)</code>	control the search for feasible initial values
<code>corrlag(#)</code>	specify maximum autocorrelation lag; default varies
<code>corrtol(#)</code>	specify autocorrelation tolerance; default is <code>corrtol(0.01)</code>

*Starred options are specific to the `bayes` prefix; other options are common between `bayes` and `bayesmh`.
Options `prior()` and `block()` can be repeated.

priorspec and *paramref* are defined in [BAYES] `bayesmh`.

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

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

Model parameters are regression coefficients `{depvar: indepvars}` for the main regression and `{select: varlist_s}` for the selection equation, and atanh-transformed correlation `{athrho}`. Use the `dryrun` option to see the definitions of model parameters prior to estimation.

For a detailed description of *bayesopts*, see *Options* in [BAYES] `bayes`.

Remarks and examples

stata.com

For a general introduction to Bayesian analysis, see [BAYES] `intro`. For a general introduction to Bayesian estimation using an adaptive Metropolis–Hastings algorithm, see [BAYES] `bayesmh`. For

remarks and examples specific to the `bayes` prefix, see [BAYES] **bayes**. For details about the estimation command, see [R] **heckprobit**.

For a simple example of the `bayes` prefix, see *Introductory example* in [BAYES] **bayes**. Also see *Heckman selection model* in [BAYES] **bayes**.

Stored results

See *Stored results* in [BAYES] **bayesmh**.

Methods and formulas

See *Methods and formulas* in [BAYES] **bayesmh**.

Also see

[BAYES] **bayes** — Bayesian regression models using the `bayes` prefix

[R] **heckprobit** — Probit model with sample selection

[BAYES] **bayesian postestimation** — Postestimation tools for `bayesmh` and the `bayes` prefix

[BAYES] **bayesian estimation** — Bayesian estimation commands

[BAYES] **bayesian commands** — Introduction to commands for Bayesian analysis

[BAYES] **intro** — Introduction to Bayesian analysis

[BAYES] **Glossary**