

**bayes: heckman** — Bayesian Heckman selection model

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

## Description

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

## Quick start

Bayesian Heckman model of  $y$  on  $x_1$  and  $x_2$ , using  $z_1$  and  $z_2$  to model selection and using default normal priors for regression coefficients, log-standard-deviation, and atanh-correlation

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

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

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

Specify 20,000 Markov chain Monte Carlo (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):, ///
heckman y x1 x2, select(z1 z2)
```

In the above, request that the 90% highest posterior density (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 \[R\] heckman](#).

## Menu

Statistics > Linear models and related > Bayesian regression > Heckman selection model

## Syntax

```
bayes [ , bayesopts ] : heckman 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

## 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<sub>s</sub>* may contain factor variables; see [U] 11.4.3 **Factor variables**.

*depvar*, *indepvars*, *varlist<sub>s</sub>*, and *depvars* may contain time-series operators; see [U] 11.4.4 **Time-series varlists**.

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

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

For a detailed description of *options*, see *Options for Heckman selection model (ML)* and *Options for Heckman selection model (two-step)* in [R] **heckman**.

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

## Priors

* <u>normalprior</u> (#)	specify standard deviation of default normal priors for regression coefficients, log-standard-deviation, 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>nchains</u> (#)	number of chains; default is to simulate one chain
<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><u>initial</u>(<i>initspec</i>)</code>	specify initial values for model parameters with a single chain
<code>init#(<i>initspec</i>)</code>	specify initial values for #th chain; requires <code>nchains()</code>
<code>initall(<i>initspec</i>)</code>	specify initial values for all chains; requires <code>nchains()</code>
<code>nomleinitial</code>	suppress the use of maximum likelihood estimates as starting values
<code><u>initrandom</u></code>	specify random initial values
<code><u>initsummary</u></code>	display initial values used for simulation
* <code>noisily</code>	display output from the estimation command during initialization

Adaptation

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

Reporting

<code><u>clevel</u>(#)</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><u>eform</u>[ (<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><u>saving</u>(<i>filename</i>[ , <i>replace</i> ])</code>	save simulation results to <i>filename.dta</i>
<code><u>nomodelsummary</u></code>	suppress model summary
<code>chainsdetail</code>	display detailed simulation summary for each chain
<code>[ <i>no</i> ]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>[ <i>no</i> ]show(<i>paramref</i>)</code>	specify model parameters to be excluded from or included in the output
<code><u>notable</u></code>	suppress estimation table
<code><u>noheader</u></code>	suppress output header
<code>title(<i>string</i>)</code>	display <i>string</i> as title above the table of parameter estimates
<code><u>display_options</u></code>	control spacing, line width, and base and empty cells

Advanced

<code><u>search</u>(<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()` may be repeated.

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

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

`collect` is allowed; see [U] 11.1.10 Prefix commands.

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, atanh-transformed correlation `{athrho}`, and log-standard deviation `{lnsigma}`. 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

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] **heckman**.

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] **bayes**.

## Methods and formulas

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

## Also see

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

[R] **heckman** — Heckman selection model

[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**

Stata, Stata Press, and Mata are registered trademarks of StataCorp LLC. Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations. Other brand and product names are registered trademarks or trademarks of their respective companies. Copyright © 1985–2023 StataCorp LLC, College Station, TX, USA. All rights reserved.

