**bayes: heckprobit** — Bayesian probit model with sample selection

### 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 \( x_1 \) and \( x_2 \), using \( z_1 \) and \( z_2 \) to model selection and using default normal priors for regression coefficients and atanh-correlation

\[
\text{bayes: heckprobit } y \ x_1 \ x_2, \text{ select}(z_1 \ z_2)
\]

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

\[
\text{bayes, normalprior(10): heckprobit } y \ x_1 \ x_2, \text{ select}(z_1 \ z_2)
\]

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

\[
\text{bayes, prior({y: x1 x2}, uniform(-10,10)): heckprobit y x1 x2, select(z1 z2)}
\]

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

\[
\text{bayes, saving(simdata) rseed(123): heckprobit y x1 x2, select(z1 z2)}
\]

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

\[
\text{bayes, mcmcsize(20000) burnin(5000) dots(500): heckprobit y x1 x2, select(z1 z2)}
\]

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

\[
\text{bayes, clevel(90) hpd}
\]

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

```plaintext
bayes [, bayesopts] : heckprobit depvar indepvars [if] [in] [weight],

   select([depvar_s =] varlist_s [, noconstant offset(varname_o)]) [options]
```

<table>
<thead>
<tr>
<th>options</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td></td>
</tr>
<tr>
<td>*select()</td>
<td>specify selection equation: dependent and independent</td>
</tr>
<tr>
<td></td>
<td>variables; whether to have constant term and offset</td>
</tr>
<tr>
<td></td>
<td>variable</td>
</tr>
<tr>
<td>noconstant</td>
<td>suppress constant term</td>
</tr>
<tr>
<td>offset(varname)</td>
<td>include varname in model with coefficient constrained</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Reporting</td>
<td></td>
</tr>
<tr>
<td>display_options</td>
<td>control spacing, line width, and base and empty</td>
</tr>
<tr>
<td></td>
<td>cells</td>
</tr>
<tr>
<td>level(#)</td>
<td>set credible level; default is level(95)</td>
</tr>
</tbody>
</table>

*select() is required.

The full specification is `select([depvar_s =] varlist_s [, noconstant offset(varname_o)])`.

`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.

Weights 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.

### bayesopts

<table>
<thead>
<tr>
<th>options</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priors</td>
<td></td>
</tr>
<tr>
<td>*normalprior(#)</td>
<td>specify standard deviation of default normal priors</td>
</tr>
<tr>
<td></td>
<td>for regression coefficients and atanh-correlation;</td>
</tr>
<tr>
<td></td>
<td>default is <code>normalprior(100)</code></td>
</tr>
<tr>
<td>prior(priorspec)</td>
<td>prior for model parameters; this option may be repeated</td>
</tr>
<tr>
<td>dryrun</td>
<td>show model summary without estimation</td>
</tr>
<tr>
<td>Simulation</td>
<td></td>
</tr>
<tr>
<td>nchains(#)</td>
<td>number of chains; default is to simulate one chain</td>
</tr>
<tr>
<td>mcmcsize(#)</td>
<td>MCMC sample size; default is <code>mcmcsize(10000)</code></td>
</tr>
<tr>
<td>burnin(#)</td>
<td>burn-in period; default is <code>burnin(2500)</code></td>
</tr>
<tr>
<td>thinning(#)</td>
<td>thinning interval; default is <code>thinning(1)</code></td>
</tr>
<tr>
<td>rseed(#)</td>
<td>random-number seed</td>
</tr>
<tr>
<td>exclude(paramref)</td>
<td>specify model parameters to be excluded from the</td>
</tr>
<tr>
<td></td>
<td>simulation results</td>
</tr>
<tr>
<td>Blocking</td>
<td></td>
</tr>
<tr>
<td>*blocksize(#)</td>
<td>maximum block size; default is <code>blocksize(50)</code></td>
</tr>
<tr>
<td>block(paramref[, blockopts])</td>
<td>specify a block of model parameters; this option</td>
</tr>
<tr>
<td></td>
<td>may be repeated</td>
</tr>
<tr>
<td>blocksummary</td>
<td>display block summary</td>
</tr>
<tr>
<td>*noblocking</td>
<td>do not block parameters by default</td>
</tr>
</tbody>
</table>
Initialization

invalidate \( \text{init}(\text{initspec}) \) specify initial values for model parameters with a single chain
invalidate \( \text{init}(\#; \text{initspec}) \) specify initial values for \#th chain; requires \text{nchains()}
invalidate \( \text{initall}(\text{initspec}) \) specify initial values for all chains; requires \text{nchains()}

\text{nomleinit} \) suppress the use of maximum likelihood estimates as starting values

\text{inirandom} \) specify random initial values

\text{inisummary} \) display initial values used for simulation

\*noisily \) display output from the estimation command during initialization

Adaptation

\text{adaptation}(\text{adaptopts}) \) control the adaptive MCMC procedure

\text{scale}(\#) \) initial multiplier for scale factor; default is \text{scale}(2.38)

\text{covariance}(\text{cov}) \) initial proposal covariance; default is the identity matrix

Reporting

\text{clevel}(\#) \) set credible interval level; default is \text{clevel}(95)

\text{hpd} \) display HPD credible intervals instead of the default equal-tailed credible intervals

\text{eform}[\text{\text{(string)}}] \) report exponentiated coefficients and, optionally, label as \text{string}

\text{batch}(\#) \) specify length of block for batch-means calculations; default is \text{batch}(0)

\text{saving}(\text{filename}, \text{replace}) \) save simulation results to \text{filename}.dta

\text{nomodelsummary} \) suppress model summary

\text{chainsdetail} \) display detailed simulation summary for each chain

\[\text{no}\]\text{dots} \) suppress dots or display dots every 100 iterations and iteration numbers every 1,000 iterations; default is \text{nodots}

\text{dots}(\#[, \text{every}(\#)]) \) display dots as simulation is performed

\[\text{no}\]\text{show}(\text{paramref}) \) specify model parameters to be excluded from or included in the output

\text{notable} \) suppress estimation table

\text{noheader} \) suppress output header

\text{title}(\text{string}) \) display \text{string} as title above the table of parameter estimates

\text{display_options} \) control spacing, line width, and base and empty cells

Advanced

\text{search}(\text{search_options}) \) control the search for feasible initial values

\text{corrlag}(\#) \) specify maximum autocorrelation lag; default varies

\text{corrtol}(\#) \) specify autocorrelation tolerance; default is \text{corrtol}(0.01)

\*Starred options are specific to the \text{bayes} prefix; other options are common between \text{bayes} and \text{bayesmh}.

Options \text{prior()} and \text{block()} may be repeated.

\text{priorspec} and \text{paramref} are defined in \text{[BAYES] bayesmh}.

\text{paramref} may contain factor variables; see \text{[U] 11.4.3 Factor variables}.

See \text{[U] 20 Estimation and postestimation commands} for more capabilities of estimation commands.

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

For a detailed description of \text{bayesopts}, see \text{Options} in \text{[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] 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] bayes.

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