Postestimation commands

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

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>estat correlation</code></td>
<td>report the correlation matrix of the errors of the dependent variable and the endogenous variables</td>
</tr>
<tr>
<td><code>estat covariance</code></td>
<td>report the covariance matrix of the errors of the dependent variable and the endogenous variables</td>
</tr>
</tbody>
</table>

These commands are not appropriate after the `svy` prefix.

The following standard postestimation commands are also available:

<table>
<thead>
<tr>
<th>Command</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>contrast</code></td>
<td>contrasts and ANOVA-style joint tests of estimates</td>
</tr>
<tr>
<td><code>estat ic</code></td>
<td>Akaike’s, consistent Akaike’s, corrected Akaike’s, and Schwarz’s Bayesian information criteria (AIC, CAIC, AICc, and BIC)</td>
</tr>
<tr>
<td><code>estat summarize</code></td>
<td>summary statistics for the estimation sample</td>
</tr>
<tr>
<td><code>estat vce</code></td>
<td>variance–covariance matrix of the estimators (VCE)</td>
</tr>
<tr>
<td><code>estat (svy)</code></td>
<td>postestimation statistics for survey data</td>
</tr>
<tr>
<td><code>estimates</code></td>
<td>cataloging estimation results</td>
</tr>
<tr>
<td><code>etable</code></td>
<td>table of estimation results</td>
</tr>
<tr>
<td><code>*forecast</code></td>
<td>dynamic forecasts and simulations</td>
</tr>
<tr>
<td><code>lincom</code></td>
<td>point estimates, standard errors, testing, and inference for linear combinations of coefficients</td>
</tr>
<tr>
<td><code>margins</code></td>
<td>marginal means, predictive margins, marginal effects, and average marginal effects</td>
</tr>
<tr>
<td><code>marginsplot</code></td>
<td>graph the results from margins (profile plots, interaction plots, etc.)</td>
</tr>
<tr>
<td><code>nlncom</code></td>
<td>point estimates, standard errors, testing, and inference for nonlinear combinations of coefficients</td>
</tr>
<tr>
<td><code>predict</code></td>
<td>conditional means, linear predictions, etc.</td>
</tr>
<tr>
<td><code>predictnl</code></td>
<td>point estimates, standard errors, testing, and inference for generalized predictions</td>
</tr>
<tr>
<td><code>pwcompare</code></td>
<td>pairwise comparisons of estimates</td>
</tr>
<tr>
<td><code>test</code></td>
<td>Wald tests of simple and composite linear hypotheses</td>
</tr>
<tr>
<td><code>testnl</code></td>
<td>Wald tests of nonlinear hypotheses</td>
</tr>
</tbody>
</table>

`*forecast` is not appropriate with `svy` estimation results.
predict

Description for predict

predict creates a new variable containing predictions such as conditional means, linear predictions, and standard errors.

Menu for predict

Statistics > Postestimation

Syntax for predict

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

<table>
<thead>
<tr>
<th>statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cm</td>
<td>conditional mean; the default</td>
</tr>
<tr>
<td>xb</td>
<td>linear prediction</td>
</tr>
<tr>
<td>stdp</td>
<td>standard error of the linear prediction</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>asfmethod</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>asf</td>
<td>average structural function; the default</td>
</tr>
<tr>
<td>fixedasf</td>
<td>fixed average structural function</td>
</tr>
</tbody>
</table>

Options for predict

- `cm`, the default, calculates the conditional mean.
- `xb` calculates the linear prediction.
- `stddp` calculates the standard error of the linear prediction.
- `asf` and `fixedasf` determine how the average structural function (ASF) of the conditional mean is computed. These options are not allowed with `xb` or `stddp`.

`asf` is the default estimator when the `cm` statistic is specified. `asf` computes the ASF of the conditional mean. It is the prediction conditional on the errors of the endogenous variable equations. Put another way, it is the conditional mean accounting for the correlation of the endogenous covariates with the errors of the main equation. Derivatives and contrasts based on `asf` have a structural interpretation. See `margins` below for computing derivatives and contrasts.
fixedasf calculates a fixed ASF. It is the prediction using only the coefficients and variables of the outcome equation. fixedasf does not condition on the errors of the endogenous variable equations. Contrasts between two fixed counterfactuals averaged over the whole sample have a potential-outcome interpretation. Intuitively, it is as if the values of the covariates were fixed at a value exogenously by fiat. See margins below for computing derivatives and contrasts.

To be clear, derivatives and contrasts between two fixed counterfactuals using the default asf option also have a potential-outcome interpretation. And, unlike fixedasf, they retain that interpretation when computed over subpopulations for both linear and nonlinear models.

scores calculates the equation-level score variables.

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

   The first new variable will contain $\partial \ln L / \partial (z_i \delta)$.
   The second new variable will contain $\partial \ln L / \partial (x_i \Pi)$.
   The third new variable will contain $\partial \ln L / \partial \text{atanh } \rho$.
   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 $\partial \ln L / \partial (z_i \delta)$.
   The second through $(p + 1)$th new variables will contain $\partial \ln L / \partial (x_i \Pi_k)$, $k = 1, \ldots, p$, where $\Pi_k$ is the $k$th column of $\Pi$.
   The remaining score variables will contain the partial derivatives of $\ln L$ with respect to $s_{21}, s_{31}, \ldots, s_{p+1,1}, s_{22}, \ldots, s_{p+1,2}, \ldots, s_{p+1,p+1}$, where $s_{m,n}$ denotes the $(m, n)$ element of the Cholesky decomposition of the error covariance matrix.
margins

Description for margins

margins estimates margins of response for linear predictions and probabilities.

Menu for margins

Statistics > Postestimation

Syntax for margins

margins [marginlist] [ , options ]  
margins [marginlist] , predict(statistic ...) [ predict(statistic ...) ... ] [ options ]

<table>
<thead>
<tr>
<th>statistic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cm</td>
<td>conditional mean; the default</td>
</tr>
<tr>
<td>xb</td>
<td>linear prediction</td>
</tr>
<tr>
<td>stdp</td>
<td>not allowed with margins</td>
</tr>
</tbody>
</table>

Statistics not allowed with margins are functions of stochastic quantities other than e(b).

For the full syntax, see [R] margins.
estat

Description for estat

estat correlation displays the correlation matrix of the errors of the dependent variable and the endogenous variables.
estat covariance displays the covariance matrix of the errors of the dependent variable and the endogenous variables.

Menu for estat

Statistics > Postestimation

Syntax for estat

Correlation matrix

estat correlation [, border(bspec) left(#) format(%fmt)]

Covariance matrix

estat covariance [, border(bspec) left(#) format(%fmt)]

Options for estat

border(bspec) sets the border style of the matrix display. The default is border(all).
left(#) sets the left indent of the matrix display. The default is left(2).
format(%fmt) specifies the format for displaying the individual elements of the matrix. The default is format(%9.0g).
Remarks and examples

Remarks are presented under the following headings:

Marginal effects
Obtaining predicted values

Marginal effects

Example 1

We can obtain marginal effects by using the `margins` command after `ivfprobit`. We will calculate average marginal effects by using the 401(k) participation model of example 1 in [R] `ivfprobit`.

```
. use https://www.stata-press.com/data/r18/401k
   (Firm-level data on 401k participation)
. ivfprobit prate c.ltotemp##c.ltotemp i.sole (mrate = c.age##c.age)
   (output omitted)
. margins, dydx(mrate) predict(cm)
```

```
Average marginal effects
Model VCE: Robust
Expression: Conditional mean of prate, predict(cm)
dy/dx wrt: mrate
```

```
<table>
<thead>
<tr>
<th></th>
<th>Delta-method</th>
<th></th>
<th></th>
<th></th>
<th>[95% conf. interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dy/dx</td>
<td>std. err.</td>
<td>z</td>
<td>P&gt;</td>
<td>z</td>
</tr>
<tr>
<td>mrate</td>
<td>.5288314</td>
<td>.0442438</td>
<td>11.95</td>
<td>0.000</td>
<td>.4421152 .6155477</td>
</tr>
</tbody>
</table>
```

The marginal effect from `ivfprobit` suggests that a $1 increase in the matching rate (per dollar) given by employers can increase participation by approximately 50%.

Obtaining predicted values

After fitting your model with `ivfprobit`, you can obtain the conditional mean, or 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`.

Stored results

`estat correlation` stores the following results in `r()`:

Matrices
- `r(corr)` : correlation matrix of the errors

`estat covariance` stores the following results in `r()`:

Matrices
- `r(cov)` : covariance matrix of the errors
Methods and formulas

Recall that the model fit by `ivfprobit` is given by two equations:

\[
E[y_{1i}|x_{1i}, x_{2i}, y_{2i}, u_i] = \Phi(y_{2i}\beta + x_{1i}\gamma + u_i)
\]

\[
y_{2i} = x_{i1}\Pi_1 + x_{i2}\Pi_2 + v_i
\]

The linear prediction for observation \(i\) is calculated as \(y_{2i}\hat{\beta} + x_{1i}\hat{\gamma}\). The predicted conditional mean is given by \(\Phi(\hat{m}_i)\), where \(\Phi(\cdot)\) is the standard normal distribution function and \(\hat{m}_i\) is the plugin estimator of

\[
m_i = \frac{y_{2i}\hat{\beta} + x_{1i}\hat{\gamma} + \rho v_i/\sigma}{\sqrt{1 - \rho^2}}
\]

where \(\sigma\) is the standard deviation of \(v_i\) and \(\rho\) is the correlation coefficient between \(u_i\) and \(v_i\). The ASF uses \(\hat{m}_i\) instead of \(y_{2i}\hat{\beta} + x_{1i}\hat{\gamma}\) to evaluate \(\Phi(\cdot)\) and account for endogeneity in the model. The fixed ASF is evaluated at \(y_{2i}\hat{\beta} + x_{1i}\hat{\gamma}\).

Also see

[R] `ivfprobit` — Fractional probit model with continuous endogenous covariates

[U] 20 Estimation and postestimation commands