Title

> **xtreg postestimation** — Postestimation tools for xtreg[+]

[+]This command includes features that are part of StataNow.

Postestimation commands

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

Command	Description
xttest0	Breusch and Pagan LM test for random effects

The following standard postestimation commands are also available:

Command	Description
contrast	contrasts and ANOVA-style joint tests of estimates
*estat ic	Akaike's, consistent Akaike's, corrected Akaike's, and Schwarz's Bayesian information criteria (AIC, CAIC, AICc, and BIC)
estat summarize	summary statistics for the estimation sample
estat vce	variance–covariance matrix of the estimators (VCE)
estimates	cataloging estimation results
etable	table of estimation results
†forecast	dynamic forecasts and simulations
hausman	Hausman's specification test
lincom	point estimates, standard errors, testing, and inference for linear combinations of coefficients
*lrtest	likelihood-ratio test
margins	marginal means, predictive margins, marginal effects, and average marginal effects
marginsplot	graph the results from margins (profile plots, interaction plots, etc.)
nlcom	point estimates, standard errors, testing, and inference for nonlinear combinations of coefficients
predict	linear predictions, residuals, error components
predictnl	point estimates, standard errors, testing, and inference for generalized predictions
pwcompare	pairwise comparisons of estimates
test	Wald tests of simple and composite linear hypotheses
testnl	Wald tests of nonlinear hypotheses

*`estat ic` and `lrtest` are not appropriate after `xtreg` with the `pa` or `re` option.

†`forecast` is not appropriate with `mi` estimation results.

predict

Description for predict

predict creates a new variable containing predictions such as fitted values, standard errors, predicted values, linear predictions, and the equation-level score.

Menu for predict

Statistics > Postestimation

Syntax for predict

Fixed-effects model

   predict [*type*] *newvar* [*if*] [*in*] [, *FE_statistic*]

Between-effects and GLS and ML random-effects model

   predict [*type*] *newvar* [*if*] [*in*] [, *RE_statistic*]

Population-averaged model

   predict [*type*] *newvar* [*if*] [*in*] [, *PA_statistic* <u>nooff</u>set]

FE_statistic	Description
Main	
xb	$\alpha + \mathbf{x}_{it}\boldsymbol{\beta}$, fitted values; the default
stdp	standard error of the fitted values
ue	$u_i + e_{it}$, the combined residual
* xbu	$\alpha + \mathbf{x}_{it}\boldsymbol{\beta} + u_i$, prediction including panel effect
* u	u_i, the fixed-error component
+* xbd	$\alpha + \mathbf{x}_{it}\boldsymbol{\beta} + d_{\text{absorbvars}}$, prediction including effects of absorbed variables
+* d	$d_{\text{absorbvars}}$, effects of absorbed variables
* e	e_{it}, the overall error component

[+]These features are part of [StataNow].

Unstarred statistics are available both in and out of sample; type predict ... if e(sample) ... if wanted only for the estimation sample. Starred statistics are calculated only for the estimation sample, even when if e(sample) is not specified.

RE_statistic	Description
Main	
xb	$\alpha + \mathbf{x}_{it}\boldsymbol{\beta}$, fitted values; the default
stdp	standard error of the fitted values
ue	$u_i + e_{it}$, the combined residual
* xbu	$\alpha + \mathbf{x}_{it}\boldsymbol{\beta} + u_i$, prediction including panel effect
* u	u_i, the random-error component
* e	e_{it}, the overall error component

Unstarred statistics are available both in and out of sample; type predict ... if e(sample) ... if wanted only for the estimation sample. Starred statistics are calculated only for the estimation sample, even when if e(sample) is not specified.

PA_statistic	Description
Main	
mu	predicted value of *depvar*; considers the offset()
rate	predicted value of *depvar*
xb	linear prediction
stdp	standard error of the linear prediction
<u>score</u>	first derivative of the log likelihood with respect to $\mathbf{x}_{it}\boldsymbol{\beta}$

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

Options for predict

> **Main**

xb calculates the linear prediction, that is, $\alpha + \mathbf{x}_{it}\boldsymbol{\beta}$. This is the default for all except the population-averaged model.

stdp calculates the standard error of the linear prediction. For the fixed-effects model, this excludes the variance due to uncertainty about the estimate of u_i.

mu and rate both calculate the predicted value of *depvar*. mu takes into account the offset(), and rate ignores those adjustments. mu and rate are equivalent if you did not specify offset(). mu is the default for the population-averaged model.

ue calculates the prediction of $u_i + e_{it}$.

xbu calculates the prediction of $\alpha + \mathbf{x}_{it}\boldsymbol{\beta} + u_i$, the prediction including the fixed or random component.

u calculates the prediction of u_i, the estimated fixed or random effect.

xbd is part of StataNow. It calculates the prediction of $\alpha + \mathbf{x}_{it}\boldsymbol{\beta} + d_{\text{absorbvars}}$, the prediction including the absorbed variables' effects.

d is part of StataNow. It calculates the prediction of $d_{\text{absorbvars}}$, the absorbed variables' effects.

e calculates the prediction of e_{it}.

score calculates the equation-level score, $u_{it} = \partial \ln L(\mathbf{x}_{it}\boldsymbol{\beta}) / \partial(\mathbf{x}_{it}\boldsymbol{\beta})$.

nooffset is relevant only if you specified offset(*varname*) for xtreg, pa. It modifies the calculations made by predict so that they ignore the offset variable; the linear prediction is treated as $\mathbf{x}_{it}\boldsymbol{\beta}$ rather than $\mathbf{x}_{it}\boldsymbol{\beta} + \text{offset}_{it}$.

margins

Description for margins

margins estimates margins of response for fitted values, probabilities, and linear predictions.

Menu for margins

Statistics > Postestimation

Syntax for margins

> margins [*marginlist*] [, *options*]
>
> margins [*marginlist*] , <u>predict</u>(*statistic* ...) [<u>predict</u>(*statistic* ...) ...] [*options*]

Fixed-effects model

FE_statistic	Description
xb	$\alpha + \mathbf{x}_{it}\boldsymbol{\beta}$, fitted values; the default
stdp	not allowed with margins
ue	not allowed with margins
xbu	not allowed with margins
u	not allowed with margins
xbd	not allowed with margins
d	not allowed with margins
e	not allowed with margins

Between-effects and GLS and ML random-effects model

RE_statistic	Description
xb	$\alpha + \mathbf{x}_{it}\boldsymbol{\beta}$, fitted values; the default
stdp	not allowed with margins
ue	not allowed with margins
xbu	not allowed with margins
u	not allowed with margins
e	not allowed with margins

Population-averaged model

PA_statistic	Description
mu	probability of *depvar*; considers the offset()
rate	probability of *depvar*
xb	linear prediction
stdp	not allowed with margins
<u>score</u>	not allowed with margins

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

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

xttest0

Description for xttest0

xttest0, for use after xtreg, re, presents the Breusch and Pagan (1980) Lagrange multiplier test for random effects, a test that $\text{Var}(\nu_i) = 0$.

Menu for xttest0

Statistics > Longitudinal/panel data > Linear models > Lagrange multiplier test for random effects

Syntax for xttest0

xttest0

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

Remarks and examples   stata.com

▷ Example 1

Continuing with our xtreg, re estimation example (example 4) in xtreg, we can see that xttest0 will report a test of $\nu_i = 0$. In case we have any doubts, we could type

```
. use https://www.stata-press.com/data/r18/nlswork
(National Longitudinal Survey of Young Women, 14-24 years old in 1968)
. xtreg ln_w grade age c.age#c.age ttl_exp c.ttl_exp#c.ttl_exp
> tenure c.tenure#c.tenure 2.race not_smsa south, re theta
 (output omitted)
. xttest0

Breusch and Pagan Lagrangian multiplier test for random effects

        ln_wage[idcode,t] = Xb + u[idcode] + e[idcode,t]

        Estimated results:
                              Var      SD = sqrt(Var)

            ln_wage |     .2283326        .4778416
                  e |     .0845002        .2906892
                  u |     .0665151        .2579053

Test: Var(u) = 0

                       chibar2(01) = 14779.98
                    Prob > chibar2 =   0.0000
```

◁

▷ Example 2

More importantly, after `xtreg, re` estimation, `hausman` will perform the Hausman specification test. If our model is correctly specified, and if ν_i is uncorrelated with $\mathbf{x}_{it}$, the (subset of) coefficients that are estimated by the fixed-effects estimator and the same coefficients that are estimated here should not statistically differ:

```
. xtreg ln_w grade age c.age#c.age ttl_exp c.ttl_exp#c.ttl_exp
> tenure c.tenure#c.tenure 2.race not_smsa south, re
(output omitted )

. estimates store random_effects

. xtreg ln_w grade age c.age#c.age ttl_exp c.ttl_exp#c.ttl_exp
> tenure c.tenure#c.tenure 2.race not_smsa south, fe
(output omitted )

. hausman . random_effects
```

| | ── Coefficients ── | | | |
| | (b) | (B) | (b-B) | sqrt(diag(V_b-V_B)) |
	.	random_eff~s	Difference	Std. err.
age	.0359987	.0368059	-.0008073	.0013177
c.age#c.age	-.000723	-.0007133	-9.68e-06	.0000184
ttl_exp	.0334668	.0290208	.0044459	.001711
c.ttl_exp# c.ttl_exp	.0002163	.0003049	-.0000886	.000053
tenure	.0357539	.0392519	-.003498	.0005797
c.tenure# c.tenure	-.0019701	-.0020035	.0000334	.0000373
not_smsa	-.0890108	-.1308252	.0418144	.0062745
south	-.0606309	-.0868922	.0262613	.0081345

```
                       b = Consistent under H0 and Ha; obtained from xtreg.
           B = Inconsistent under Ha, efficient under H0; obtained from xtreg.
   Test of H0: Difference in coefficients not systematic
       chi2(8) = (b-B)'[(V_b-V_B)^(-1)](b-B)
               = 149.43
   Prob > chi2 = 0.0000
```

We can reject the hypothesis that the coefficients are the same. Before turning to what this means, note that `hausman` listed the coefficients estimated by the two models. It did not, however, list `grade` and `2.race`. `hausman` did not make a mistake; in the Hausman test, we compare only the coefficients estimated by both techniques.

What does this mean? We have an unpleasant choice: we can admit that our model is misspecified—that we have not parameterized it correctly—or we can hold that our specification is correct, in which case the observed differences must be due to the zero correlation of ν_i and the $\mathbf{x}_{it}$ assumption.

◁

❑ Technical note

We can also mechanically explore the underpinnings of the test's dissatisfaction. In the comparison table from `hausman`, it is the coefficients on `not_smsa` and `south` that exhibit the largest differences. In equation (1′) of [XT] **xtreg**, we showed how to decompose a model into within and between effects. Let's do that with these two variables, assuming that changes in the average have one effect, whereas transitional changes have another:

```
. egen avgnsmsa = mean(not_smsa), by(id)

. generate devnsma = not_smsa -avgnsmsa
(8 missing values generated)

. egen avgsouth = mean(south), by(id)

. generate devsouth = south - avgsouth
(8 missing values generated)

. xtreg ln_w grade age c.age#c.age ttl_exp c.ttl_exp#c.ttl_exp tenure c.tenure#
> c.tenure 2.race avgnsm devnsm avgsou devsou
```

Random-effects GLS regression   Number of obs   =   28,091
Group variable: idcode   Number of groups =   4,697

R-squared: Obs per group:
    Within = 0.1723   min =   1
    Between = 0.4809   avg =   6.0
    Overall = 0.3737   max =   15

                                                 Wald chi2(12)   =   9319.56
corr(u_i, X) = 0 (assumed)   Prob > chi2   =   0.0000

ln_wage	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
grade	.0631716	.0017903	35.29	0.000	.0596627	.0666805
age	.0375196	.0031186	12.03	0.000	.0314072	.043632
c.age#c.age	−.0007248	.00005	−14.50	0.000	−.0008228	−.0006269
ttl_exp	.0286543	.0024207	11.84	0.000	.0239098	.0333989
c.ttl_exp#						
c.ttl_exp	.0003222	.0001162	2.77	0.006	.0000945	.0005499
tenure	.0394423	.001754	22.49	0.000	.0360044	.0428801
c.tenure#						
c.tenure	−.0020081	.0001192	−16.85	0.000	−.0022417	−.0017746
race						
Black	−.0545936	.0102101	−5.35	0.000	−.074605	−.0345821
avgnsmsa	−.1833237	.0109339	−16.77	0.000	−.2047537	−.1618937
devnsma	−.0887596	.0095071	−9.34	0.000	−.1073931	−.070126
avgsouth	−.1011235	.0098789	−10.24	0.000	−.1204858	−.0817611
devsouth	−.0598538	.0109054	−5.49	0.000	−.081228	−.0384797
_cons	.2682987	.0495778	5.41	0.000	.171128	.3654694
sigma_u	.2579182					
sigma_e	.29068923					
rho	.44047745	(fraction of variance due to u_i)				

We will leave the reinterpretation of this model to you, except that if we were really going to sell this model, we would have to explain why the between and within effects are different. Focusing on residence in a non-SMSA, we might tell a story about rural people being paid less and continuing to get paid less when they move to the SMSA. Given our panel data, we could create variables to measure this (an indicator for moved from non-SMSA to SMSA) and to measure the effects. In our assessment of this model, we should think about women in the cities moving to the country and their relative productivity in a bucolic setting.

In any case, the Hausman test now is

```
. estimates store new_random_effects

. xtreg ln_w grade age c.age#c.age ttl_exp c.ttl_exp#c.ttl_exp
> tenure c.tenure#c.tenure 2.race avgnsm devnsm avgsou devsou, fe
  (output omitted )

. hausman . new_random_effects
```

| | ─── Coefficients ─── | | | |
| | (b) | (B) | (b-B) | sqrt(diag(V_b-V_B)) |
	.	new_random~s	Difference	Std. err.
age	.0359987	.0375196	−.0015209	.0013198
c.age#c.age	−.000723	−.0007248	1.84e−06	.0000184
ttl_exp	.0334668	.0286543	.0048124	.0017127
c.ttl_exp#				
c.ttl_exp	.0002163	.0003222	−.0001059	.0000531
tenure	.0357539	.0394423	−.0036884	.0005839
c.tenure#				
c.tenure	−.0019701	−.0020081	.000038	.0000377
devnsma	−.0890108	−.0887596	−.0002512	.000683
devsouth	−.0606309	−.0598538	−.0007771	.0007618

```
                  b = Consistent under H0 and Ha; obtained from xtreg.
         B = Inconsistent under Ha, efficient under H0; obtained from xtreg.
Test of H0: Difference in coefficients not systematic
    chi2(8) = (b-B)'[(V_b-V_B)^(-1)](b-B)
            =  92.52
Prob > chi2 = 0.0000
```

We have mechanically succeeded in greatly reducing the χ^2, but not by enough. The major differences now are in the age, experience, and tenure effects. We already knew this problem existed because of the ever-increasing effect of experience. More careful parameterization work rather than simply including squares needs to be done.

❏

Stored results

xttest0 stores the following in r():

Scalars
   r(lm)   Lagrange multiplier statistic
   r(df)   degrees of freedom
   r(p) p-value

Methods and formulas

Methods and formulas are presented under the following headings:

Predictions for fixed-effects model with absorbed variables (StataNow)
xttest0

Predictions for fixed-effects model with absorbed variables (StataNow)

The following uses the notation introduced in *Methods and formulas* in [XT] **xtreg**.

Suppose we fit the model

$$y_{it} = \alpha + \mathbf{x}_{it}\boldsymbol{\beta} + \eta_{it} + \epsilon_{it}$$

where $\eta_{it} = \mathbf{d}_{1(it)}\boldsymbol{\gamma}_1 + \mathbf{d}_{2(it)}\boldsymbol{\gamma}_2 + \cdots + \mathbf{d}_{h(it)}\boldsymbol{\gamma}_h$ and $\mathbf{d}_{k(it)}$ is an indicator vector for absorbed variable k in panel i at time t. By convention, we define the first absorb variable ($k = 1$) as the panel-ID variable and write $\eta_{it} = \nu_i + \sum_{k=2}^{h} \mathbf{d}_{k(it)}\boldsymbol{\gamma}_k = \nu_i + \Delta_{it}$. Let $\boldsymbol{\eta}$ be the $1 \times N$ vector with the values of η_{it} in the sample. Define vectors Δ and $\boldsymbol{\nu}$ similarly.

We are interested in making predictions for the panel effects $\boldsymbol{\nu}$ (option u), the absorbed variables' effects Δ (option d), and the residuals ϵ (option e). These postestimation predictions use alternating projection methods to avoid the computational burden of estimating the $\boldsymbol{\gamma}$'s. Two alternating projection method algorithms are available: Halperin and Cimmino. For a description of these two algorithms, see *Methods and formulas* in [XT] **xtreg**.

Estimates for $\boldsymbol{\nu}$, Δ, and the residuals can be extracted from $\widehat{\boldsymbol{v}} = \mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}$, where $\widehat{\boldsymbol{\beta}}$ is the estimated regression coefficient vector. Here is the procedure. First, separate the absorbed variables into two sets: $\iota_1 = \{1\}$, the panel effect, and $\iota_2 = \{2, \ldots, h\}$, the remaining absorbed variables. Let project$(\widehat{\boldsymbol{v}}, \iota_i)$ denote the projection of vector $\widehat{\boldsymbol{v}}$ over the variables in set ι_i, using either the Halperin or Cimmino algorithm, and let p_error$(\widehat{\boldsymbol{v}}, \iota_i)$ denote the corresponding projection error. Second, initialize the values of the algorithm to $\widetilde{\boldsymbol{\epsilon}}^{(1)} = \widehat{\boldsymbol{v}}$, $\widetilde{\boldsymbol{\nu}}^{(1)} = \mathbf{0}$, and $\widetilde{\Delta}^{(1)} = \mathbf{0}$. Third and finally, iterate until convergence using the following formulas:

$$\widetilde{\boldsymbol{\nu}}^{(j+1)} = \widetilde{\boldsymbol{\nu}}^{(j)} + \text{project}(\widetilde{\boldsymbol{\epsilon}}^{(j)}, \iota_1)$$

$$\widetilde{\boldsymbol{\epsilon}}^{(j+1)} = \text{p\_error}\{\text{p\_error}(\widetilde{\boldsymbol{\epsilon}}^{(j)}, \iota_1), \iota_2\}$$

$$\widetilde{\Delta}^{(j+1)} = \widetilde{\Delta}^{(j)} + \text{project}\{\text{p\_error}(\widetilde{\boldsymbol{\epsilon}}^{(j)}, \iota_1), \iota_2\}$$

Convergence is declared at $j = \jmath$ when the elements of project$(\widetilde{\boldsymbol{\epsilon}}^{(\jmath)}, \iota_1)$ and project$(\widetilde{\boldsymbol{\epsilon}}^{(\jmath)}, \iota_2)$ are negligible. At convergence, $\widetilde{\boldsymbol{\epsilon}}^{(\jmath)}$ contains the estimated residuals, $\widetilde{\boldsymbol{\nu}}^{(\jmath)}$ the estimated panel effects, and $\widetilde{\Delta}^{(\jmath)}$ the estimated absorbed variables' effects.

xttest0

xttest0 reports the Lagrange multiplier test for random effects developed by Breusch and Pagan (1980) and as modified by Baltagi and Li (1990). The model

$$y_{it} = \alpha + \mathbf{x}_{it}\boldsymbol{\beta} + \nu_i$$

is fit via OLS, and then the quantity

$$\lambda_{\text{LM}} = \frac{(n\overline{T})^2}{2} \left\{ \frac{A_1^2}{(\sum_i T_i^2) - n\overline{T}} \right\}$$

is calculated, where

$$A_1 = 1 - \frac{\sum_{i=1}^{n}(\sum_{t=1}^{T_i} v_{it})^2}{\sum_i \sum_t v_{it}^2}$$

The Baltagi and Li modification allows for unbalanced data and reduces to the standard formula

$$
\lambda_{\mathrm{LM}} =
\begin{cases}
\frac{nT}{2(T-1)} \left\{ \frac{\sum_i (\sum_t v_{it})^2}{\sum_i \sum_t v_{it}^2} - 1 \right\}^2, & \widehat{\sigma}_u^2 \geq 0 \\
0, & \widehat{\sigma}_u^2 < 0
\end{cases}
$$

when $T_i = T$ (balanced data). Under the null hypothesis, λ_{LM} is distributed as a 50:50 mixture of a point mass at zero and $\chi^2(1)$.

References

Alejo, J., A. F. Galvao, G. Montes-Rojas, and W. Sosa-Escudero. 2015. Tests for normality in linear panel-data models. *Stata Journal* 15: 822–832.

Baltagi, B. H., and Q. Li. 1990. A Lagrange multiplier test for the error components model with incomplete panels. *Econometric Reviews* 9: 103–107. https://doi.org/10.1080/07474939008800180.

Breusch, T. S., and A. R. Pagan. 1980. The Lagrange multiplier test and its applications to model specification in econometrics. *Review of Economic Studies* 47: 239–253. https://doi.org/10.2307/2297111.

Hausman, J. A. 1978. Specification tests in econometrics. *Econometrica* 46: 1251–1271. https://doi.org/10.2307/1913827.

Sosa-Escudero, W., and A. K. Bera. 2008. Tests for unbalanced error-components models under local misspecification. *Stata Journal* 8: 68–78.

Verbeke, G., and G. Molenberghs. 2003. The use of score tests for inference on variance components. *Biometrics* 59: 254–262. https://doi.org/10.1111/1541-0420.00032.

Also see

[XT] **xtreg** — Fixed-, between-, and random-effects and population-averaged linear models[+]

[U] **20 Estimation and postestimation commands**