

areg postestimation — Postestimation tools for areg

Postestimation commands
References

predict
Also see

margins

Remarks and examples

Postestimation commands

The following postestimation commands are available after `areg`:

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

*`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, residuals, and the equation-level score.

Menu for predict

Statistics > Postestimation

Syntax for predict

```
predict [type] newvar [if] [in] [, statistic]
```

where $y_j = \mathbf{x}_j \mathbf{b} + d_{\text{absorbvar}} + e_j$ and *statistic* is

<i>statistic</i>	Description
------------------	-------------

Main

<code>xb</code>	$\mathbf{x}_j \mathbf{b}$, fitted values; the default
<code>stdp</code>	standard error of the prediction
<code><u>d</u>residuals</code>	$d_{\text{absorbvar}} + e_j = y_j - \mathbf{x}_j \mathbf{b}$
* <code>xbd</code>	$\mathbf{x}_j \mathbf{b} + d_{\text{absorbvar}}$
* <code>d</code>	$d_{\text{absorbvar}}$
* <code><u>r</u>esiduals</code>	residual
* <code><u>s</u>core</code>	score; equivalent to <code>residuals</code>

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.

Options for predict

Main

`xb`, the default, calculates the prediction of $\mathbf{x}_j \mathbf{b}$, the fitted values, by using the average effect of the absorbed variable. Also see `xbd` below.

`stdp` calculates the standard error of $\mathbf{x}_j \mathbf{b}$.

`dresiduals` calculates $y_j - \mathbf{x}_j \mathbf{b}$, which are the residuals plus the effect of the absorbed variable.

`xbd` calculates $\mathbf{x}_j \mathbf{b} + d_{\text{absorbvar}}$, which are the fitted values including the individual effects of the absorbed variable.

`d` calculates $d_{\text{absorbvar}}$, the individual coefficients for the absorbed variable.

`residuals` calculates the residuals, that is, $y_j - (\mathbf{x}_j \mathbf{b} + d_{\text{absorbvar}})$.

`score` is a synonym for `residuals`.

margins

Description for margins

`margins` estimates margins of response for fitted values.

Menu for margins

Statistics > Postestimation

Syntax for margins

```

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

```

<i>statistic</i>	Description
<code>xb</code>	$x_j\mathbf{b}$, fitted values; the default
<code>stdp</code>	not allowed with <code>margins</code>
<code>dresiduals</code>	not allowed with <code>margins</code>
<code>xbd</code>	not allowed with <code>margins</code>
<code>d</code>	not allowed with <code>margins</code>
<code>residuals</code>	not allowed with <code>margins</code>
<code>score</code>	not allowed with <code>margins</code>

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

For the full syntax, see [R] [margins](#).

Remarks and examples

[stata.com](https://www.stata.com)

► Example 1

Continuing with [example 1](#) of [R] [areg](#), we refit the model with robust standard errors and then obtain linear predictions and standard errors for those linear predictions.

```

. use https://www.stata-press.com/data/r17/auto2
(1978 automobile data)
. areg mpg weight gear_ratio, absorb(rep78) vce(robust)
(output omitted)
. predict xb_ar
(option xb assumed; fitted values)
. predict stdp_ar, stdp

```

We can obtain the same linear predictions by fitting the model with `xtreg`, `fe`, but we would first need to specify the panel structure by using `xtset`.

```

. xtset rep78
Panel variable: rep78 (unbalanced)
. xtreg mpg weight gear_ratio, fe vce(robust)
(output omitted)
. predict xb_xt
(option xb assumed; fitted values)
. predict stdp_xt, stdp
. summarize xb_ar xb_xt stdp*

```

Variable	Obs	Mean	Std. dev.	Min	Max
xb_ar	74	21.36805	4.286788	11.58643	28.07367
xb_xt	74	21.36805	4.286788	11.58643	28.07367
stdp_ar	74	.7105649	.1933936	.4270821	1.245179
stdp_xt	74	.8155919	.4826332	.0826999	1.709786

The predicted `xb` values above are the same for `areg` and `xtreg, fe`, but the standard errors for those linear predictions are different. The assumptions for these two estimators lead to different formulations for their standard errors. The robust variance estimates with `areg` are equivalent to the robust variance estimates using `regress`, including the panel dummies. The consistent robust variance estimates with `xtreg` are equivalent to those obtained by specifying `vce(cluster panelvar)` with that estimation command. For a theoretical discussion, see [Wooldridge \(2020\)](#), [Stock and Watson \(2008\)](#), and [Arellano \(2003\)](#); also see the [technical note](#) after example 3 of `[XT] xtreg`.



▶ Example 2

We would like to use `linktest` to check whether the dependent variable for our model is correctly specified:

```

. use https://www.stata-press.com/data/r17/auto2, clear
(1978 automobile data)
. areg mpg weight gear_ratio, absorb(rep78)
(output omitted)
. linktest, absorb(rep78)

```

Linear regression, absorbing indicators	Number of obs	=	69
Absorbed variable: rep78	No. of categories	=	5
	F(2, 62)	=	46.50
	Prob > F	=	0.0000
	R-squared	=	0.6939
	Adj R-squared	=	0.6643
	Root MSE	=	3.3990

mpg	Coefficient	Std. err.	t	P> t	[95% conf. interval]
_hat	-.9305602	.9537856	-0.98	0.333	-2.83715 .9760302
_hatsq	.0462785	.0227219	2.04	0.046	.0008582 .0916989
_cons	19.24899	9.725618	1.98	0.052	-.1922457 38.69022

F test of absorbed indicators: F(4, 62) = 1.278 Prob > F = 0.288

The squared residuals are significant in the regression for `mpg` on the linear and squared residuals; therefore, the test indicates that our dependent variable does not seem to be well specified. Let's transform the dependent variable into energy consumption, gallons per mile, fit the alternative model, and check the link test again.

```
. generate gpm = 1/mpg
. areg gpm weight gear_ratio, absorb(rep78)
  (output omitted)
. linktest, absorb(rep78)
```

```
Linear regression, absorbing indicators          Number of obs   =    69
Absorbed variable: rep78                       No. of categories =    5
                                                F(2, 62)        =   72.60
                                                Prob > F         =  0.0000
                                                R-squared        =  0.7436
                                                Adj R-squared    =  0.7187
                                                Root MSE        =  0.0068
```

gpm	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
_hat	.2842582	.7109124	0.40	0.691	-1.136835	1.705352
_hatsq	6.956965	6.862439	1.01	0.315	-6.760855	20.67478
_cons	.0175457	.0178251	0.98	0.329	-.0180862	.0531777

```
F test of absorbed indicators: F(4, 62) = 0.065          Prob > F = 0.992
```

The link test supports the use of the transformed dependent variable.

◀

References

- Arellano, M. 2003. *Panel Data Econometrics*. Oxford: Oxford University Press.
- Stock, J. H., and M. W. Watson. 2008. Heteroskedasticity-robust standard errors for fixed effects panel data regression. *Econometrica* 76: 155–174. <https://doi.org/10.1111/j.0012-9682.2008.00821.x>.
- Wooldridge, J. M. 2020. *Introductory Econometrics: A Modern Approach*. 7th ed. Boston: Cengage.

Also see

- [R] [areg](#) — Linear regression with a large dummy-variable set
- [U] [20 Estimation and postestimation commands](#)