

lassoinfo — Display information about lasso estimation results

[Description](#)
[Option](#)

[Quick start](#)
[Remarks and examples](#)

[Menu](#)
[Stored results](#)

[Syntax](#)
[Also see](#)

Description

`lassoinfo` displays basic information about the lasso or lassos fit by all commands that fit lassos.

Quick start

After any command that fits lassos

```
lassoinfo
```

`dsregress` was run and the results stored under the name `mygreatmodel` using `estimates store`;
 show information about all the lassos in `mygreatmodel`

```
lassoinfo mygreatmodel
```

Same as above, but three models were stored

```
lassoinfo mygreatmodel mygoodmodel myfairmodel
```

After an `xpo` command, show information about every single lasso fit

```
lassoinfo, each
```

Menu

Statistics > Postestimation

Syntax

For all lasso estimation results

```
lassoinfo [namelist]
```

For xpo estimation results

```
lassoinfo [namelist] [, each]
```

namelist is a name of a stored estimation result, a list of names, `_all`, or `*`. `_all` and `*` mean the same thing. See [\[R\] estimates store](#).

`collect` is allowed; see [\[U\] 11.1.10 Prefix commands](#).

Option

each applies to xpo models only. It specifies that information be shown for each lasso for each cross-fit fold to be displayed. If `resample` was specified, then information is shown for each lasso for each cross-fit fold in each resample. By default, summary statistics are shown for the lassos.

Remarks and examples

stata.com

`lassoinfo` is intended for use after `ds`, `po`, `xpo` commands and after `telasso` to see basic information about the lassos they fit. It is a good idea to *always* run `lassoinfo` after these commands to see how many variables were selected in each lasso.

Running `lassoinfo` is a first step toward doing a sensitivity analysis. The lassos listed by `lassoinfo` can be examined using `coefpath`, `cvplot`, `lassocoef`, `lassoknots`, and `lassoselect`.

➤ Example 1: lasso

`lassoinfo` works after `lasso`, `sqrtilasso`, and `elasticnet`, but it does not display much useful information for these commands.

Here is an example using `lasso` from [LASSO] [lasso examples](#). We load the data and make the `v1` variable lists active.

```
. use https://www.stata-press.com/data/r18/fakesurvey_v1
(Fictitious survey data with v1)

. v1 rebuild
Rebuilding v1 macros ...

(output omitted)
```

We fit the lasso.

```
. lasso linear q104 $idemographics $ifactors $v1continuous, rseed(1234)
10-fold cross-validation with 100 lambdas ...
Grid value 1:      lambda = .9090511   no. of nonzero coef. =    0
Folds: 1...5....10   CVF = 18.33331

(output omitted)

Grid value 28:      lambda = .0737359   no. of nonzero coef. =   80
Folds: 1...5....10   CVF = 11.92887
... cross-validation complete ... minimum found

Lasso linear model                No. of obs      =      914
                                No. of covariates =      277
Selection: Cross-validation       No. of CV folds =       10
```

ID	Description	lambda	No. of nonzero coef.	Out-of- sample R-squared	CV mean prediction error
1	first lambda	.9090511	0	-0.0010	18.33331
23	lambda before	.1174085	58	0.3543	11.82553
* 24	selected lambda	.1069782	64	0.3547	11.81814
25	lambda after	.0974746	66	0.3545	11.8222
28	last lambda	.0737359	80	0.3487	11.92887

* lambda selected by cross-validation.

lassoinfo tells us nothing new.

```
. lassoinfo
```

```
Estimate: active
Command: lasso
```

Dependent variable	Model	Selection method	Selection criterion	lambda	No. of selected variables
q104	linear	cv	CV min.	.1069782	64

Replaying the command gives more information.

```
. lasso
```

```
Lasso linear model          No. of obs      =      914
                             No. of covariates =      277
Selection: Cross-validation  No. of CV folds =      10
```

ID	Description	lambda	No. of nonzero coef.	Out-of-sample R-squared	CV mean prediction error
1	first lambda	.9090511	0	-0.0010	18.33331
23	lambda before	.1174085	58	0.3543	11.82553
* 24	selected lambda	.1069782	64	0.3547	11.81814
25	lambda after	.0974746	66	0.3545	11.8222
28	last lambda	.0737359	80	0.3487	11.92887

* lambda selected by cross-validation.



► Example 2: dsregress

lassoinfo gives important information after the ds, po, and xpo commands.

We load the data used in [\[LASSO\] lasso examples](#). See that entry for details about the data.

```
. use https://www.stata-press.com/data/r18/fakesurvey_v1, clear
(Fictitious survey data with v1)
. v1 rebuild
Rebuilding v1 macros ...
(output omitted)
```

We are going to fit a `dsregress` model with `q104` as our dependent variable and variables of interest `q41` and `q22`. These variables of interest are currently in the variable lists `factors` and `vlcontinuous`, which we will use to specify the control variables. So we need to move them out of these variable lists.

```
. v1 modify factors = factors - (q41)
note: 1 variable removed from $factors.
. v1 move (q22) vlother
note: 1 variable specified and 1 variable moved.
(output omitted)
. v1 rebuild
Rebuilding v1 macros ...
(output omitted)
```

After we moved the variables out of the variable lists, we typed `v1 rebuild` to update the variable list `ifactors` created from `factors`. See [D] [v1](#) for details.

We fit our `dsregress` model using cross-validation to select λ^* 's in the lassos.

```
. dsregress q104 i.q41 q22,
> controls(($idemographics) $ifactors $v1continuous)
> selection(cv) rseed(1234)

Estimating lasso for q104 using cv
Estimating lasso for 1bn.q41 using cv
Estimating lasso for q22 using cv

Double-selection linear model      Number of obs      =      914
                                Number of controls      =      274
                                Number of selected controls =      123
                                Wald chi2(2)              =      10.96
                                Prob > chi2               =      0.0042
```

q104	Robust					
	Coefficient	std. err.	z	P> z	[95% conf. interval]	
q41						
Yes	.6003918	.2848483	2.11	0.035	.0420994	1.158684
q22	-.0681067	.0306219	-2.22	0.026	-.1281246	-.0080888

Note: Chi-squared test is a Wald test of the coefficients of the variables of interest jointly equal to zero. Lassos select controls for model estimation. Type `lassoinfo` to see number of selected variables in each lasso.

`lassoinfo` shows us how many variables were selected in each lasso.

```
. lassoinfo
      Estimate: active
      Command: dsregress
```

Variable	Model	Selection method	Selection criterion	lambda	No. of selected variables
q104	linear	cv	CV min.	.1116376	63
1bn.q41	linear	cv	CV min.	.0135958	68
q22	linear	cv	CV min.	.1624043	49

lassoinfo also gives useful information after fitting the model using the default selection(plugin).

```
. dsregress q104 i.q41 q22, controls(($idemographics) $ifactors $v1continuous)
Estimating lasso for q104 using plugin
Estimating lasso for 1bn.q41 using plugin
Estimating lasso for q22 using plugin
Double-selection linear model      Number of obs      =      914
                                   Number of controls      =      274
                                   Number of selected controls =      33
                                   Wald chi2(2)              =     18.72
                                   Prob > chi2               =     0.0001
```

q104	Robust		z	P> z	[95% conf. interval]	
	Coefficient	std. err.				
q41						
Yes	.8410538	.2691082	3.13	0.002	.3136114	1.368496
q22	-.0878443	.0310435	-2.83	0.005	-.1486884	-.0270001

Note: Chi-squared test is a Wald test of the coefficients of the variables of interest jointly equal to zero. Lassos select controls for model estimation. Type lassoinfo to see number of selected variables in each lasso.

```
. lassoinfo
Estimate: active
Command: dsregress
```

Variable	Model	Selection		No. of selected variables
		Method	lambda	
q104	linear	plugin	.1467287	18
1bn.q41	linear	plugin	.1467287	16
q22	linear	plugin	.1467287	15

See [\[LASSO\] lassoselect](#), where we continue this example and do a sensitivity analysis to examine the differences between the lassos fit using cross-validation and the lassos fit using the plugin estimator.



➤ Example 3: poivregress

We want to show you some differences that arise when you fit models containing endogenous variables using `poivregress` and `xpoivregress`.

We will not describe the data or the model here. See [\[LASSO\] Inference examples](#).

We load the data,

```
. use https://www.stata-press.com/data/r18/mroz2, clear
```

set v1 variable lists,

```
. vl create vars      = (kidslt6 kidsge6 age husage city exper)
note: $vars initialized with 6 variables.
. vl substitute vars2 = c.vars c.vars#c.vars
. vl create iv        = (huseduc motheduc fatheduc)
note: $iv initialized with 3 variables.
. vl substitute iv2    = c.iv c.iv#c.iv
```

and fit our model using `poivregress`.

```
. poivregress lwage (educ = $iv2), controls($vars2) selection(cv) rseed(12345)
Estimating lasso for lwage using cv
Estimating lasso for educ using cv
Estimating lasso for pred(educ) using cv
Partialing-out IV linear model      Number of obs      =      428
                                   Number of controls      =      27
                                   Number of instruments      =      9
                                   Number of selected controls =     16
                                   Number of selected instruments =      4
                                   Wald chi2(1)              =     11.10
                                   Prob > chi2               =      0.0009
```

lwage	Robust		z	P> z	[95% conf. interval]	
	Coefficient	std. err.				
educ	.0765154	.0229707	3.33	0.001	.0314936	.1215371

Endogenous: educ
Note: Chi-squared test is a Wald test of the coefficients of the variables of interest jointly equal to zero. Lassos select controls for model estimation. Type `lassoinfo` to see number of selected variables in each lasso.

```
. estimates store poivregresscv
```

We stored our estimation results using `estimates store`, and here we use `lassoinfo` with the name used to store them.

```
. lassoinfo poivregresscv
      Estimate: poivregresscv
      Command: poivregress
```

Variable	Model	Selection method	Selection criterion	lambda	No. of selected variables
lwage	linear	cv	CV min.	.0353704	3
educ	linear	cv	CV min.	.0530428	10
pred(educ)	linear	cv	CV min.	.013186	12

Note that we have two lassos for `educ` labeled by `lassoinfo` as `educ` and `pred(educ)`. `poivregress` and `xpoivregress` perform two lassos for each endogenous variable, one for the endogenous variable and one for its prediction. `lassoinfo` shows us how to refer to each of these lassos in other postestimation commands using the `for()` option. In this example, we would type `for(educ)` and `for(pred(educ))`, respectively.

◀

► Example 4: xporegress

The `xpo` commands fit many lassos. For each lasso fit by a `po` command, the corresponding `xpo` command fits `xfolds(#)` \times `resample(#)` lassos. `lassoinfo` can be used to get information about these lassos.

We will not describe the data or the model here. See [\[LASSO\] Inference examples](#).

We load the data,

```
. use https://www.stata-press.com/data/r18/breathe, clear
(Nitrogen dioxide and attention)
```

set `vl` variable lists,

```
. vl set
      (output omitted)
. vl move (siblings_old siblings_young) vlcontinuous
note: 2 variables specified and 2 variables moved.
      (output omitted)
. vl create mycontinuous = vlcontinuous - (react no2_class)
note: $mycontinuous initialized with 10 variables.
. vl substitute mycontrols = i.vlategorical mycontinuous
```

and fit our model using `xporegress` with the options `xfolds(3)` and `resample(2)`.

```
. xporegress react no2_class, controls($mycontrols) xfolds(3) resample(2)
> selection(cv) rseed(12345)

Resample 1 of 2 ...
Cross-fit fold 1 of 3 ...
Estimating lassos: 1.

Resample 1 of 2 ...
Cross-fit fold 2 of 3 ...
Estimating lassos: 1.

Resample 1 of 2 ...
Cross-fit fold 3 of 3 ...
Estimating lassos: 1.

Resample 2 of 2 ...
Cross-fit fold 1 of 3 ...
Estimating lassos: 1.

Resample 2 of 2 ...
Cross-fit fold 2 of 3 ...
Estimating lassos: 1.

Resample 2 of 2 ...
Cross-fit fold 3 of 3 ...
Estimating lassos: 1.

Cross-fit partialing-out linear model      Number of obs      =      1,036
                                           Number of controls =      32
                                           Number of selected controls =      27
                                           Number of folds in cross-fit =      3
                                           Number of resamples =      2
                                           Wald chi2(1)       =     20.99
                                           Prob > chi2        =     0.0000
```

react	Robust					[95% conf. interval]
	Coefficient	std. err.	z	P> z		
no2_class	2.332193	.5090902	4.58	0.000	1.334394	3.329991

Note: Chi-squared test is a Wald test of the coefficients of the variables of interest jointly equal to zero. Lassos select controls for model estimation. Type `lassoinfo` to see number of selected variables in each lasso.

For each cross-fit fold and each resample, `xporegress` fits lassos. So it fit six lassos for the dependent variable, `react`, and six for the variable of interest, `no2_class`. `lassoinfo` summarizes the numbers of variables selected across these six lassos for `react` and `no2_class`.

```
. lassoinfo
      Estimate: active
      Command: xporegress
```

Variable	Model	Selection method	No. of selected variables		
			min	median	max
no2_class	linear	cv	11	15	15
react	linear	cv	9	15	19

Specifying the option `each` gives us information on each lasso.

```
. lassoinfo, each
      Estimate: active
      Command: xporegress
```

Dependent variable	Model	Selection method	Resample number	xfold no.	Selection criterion	lambda	No. of sel. var.
no2_class	linear	cv	1	1	CV min.	.2663004	11
no2_class	linear	cv	1	2	CV min.	.2860957	15
no2_class	linear	cv	1	3	CV min.	.2887414	14
no2_class	linear	cv	2	1	CV min.	.2337636	15
no2_class	linear	cv	2	2	CV min.	.2824076	15
no2_class	linear	cv	2	3	CV min.	.2515777	15
react	linear	cv	1	1	CV min.	6.07542	9
react	linear	cv	1	2	CV min.	1.704323	19
react	linear	cv	1	3	CV min.	3.449884	15
react	linear	cv	2	1	CV min.	6.034922	9
react	linear	cv	2	2	CV min.	4.31785	16
react	linear	cv	2	3	CV min.	4.096779	15

See [\[LASSO\] lassocoeff](#) for an example where we list the variables selected by each lasso.



Stored results

`lassoinfo` stores the following in `r()`:

- Macros
- `r(names)`

names of estimation results displayed
- Matrices
- `r(table)`

matrix containing the numerical values displayed

Also see

- [\[LASSO\] lassoselect](#) — Select lambda after lasso
- [\[LASSO\] lasso postestimation](#) — Postestimation tools for lasso for prediction
- [\[LASSO\] lasso inference postestimation](#) — Postestimation tools for lasso inferential models

