Title

stata.com

fvset — Declare factor-variable settings

Description

`fvset base`, `fvset design`, and `fvset clear` manage factor-variable settings, which identify the base level and specify how to accumulate statistics over levels. `fvset base` declares the base level for each specified variable; the default for factor variables without a declared base level is the lowest value. `fvset design` specifies how the `margins` command is to accumulate over the levels of a factor variable. `fvset clear` removes factor-variable settings for each variable in *varlist*. `fvset clear _all` removes all factor-variable settings from all variables.

`fvset report` reports the current factor-variable settings for each variable in *varlist*. `fvset` without arguments is a synonym for `fvset report`.

Quick start

Set the base category of categorical variable `a1` to 3

```
fvset base 3 a1
```

Set the base category of `a2`, `a3`, and `a4` to each variable's largest observed value

```
fvset base last a2 a3 a4
```

Set the base category of `a5` to the most frequent category

```
fvset base frequent a5
```

Set `a6` to have no base category

```
fvset base none a6
```

Restore the default base category (first) for `a5`

```
fvset base default a5
```

Specify that `margins` should treat `a2` as though it is balanced

```
fvset design asbalanced a2
```

Clear factor-variable settings for `a2` to `a4`

```
fvset clear a2-a4
```

List factor-variable settings for all factor variables

```
fvset report
```

Syntax

Declare base settings

```
fvset base base_spec varlist
```

Declare design settings

```
fvset design design_spec varlist
```

Clear the current settings

```
fvset clear varlist
```

Report the current settings

```
fvset report [varlist] [, base(base_spec) design(design_spec)]
```

base_spec	Description
`default`	default base
`first`	lowest level value; the default
`last`	highest level value
`frequent`	most frequent level value
`none`	no base
#	nonnegative integer value

design_spec	Description
`default`	default design
`asbalanced`	accumulate using $1/k$, k = number of levels
`asobserved`	accumulate using observed relative frequencies; the default

Options

`base(`*base_spec*`)` restricts `fvset report` to report only the factor-variable settings for variables with the specified *base_spec*.

`design(`*design_spec*`)` restricts `fvset report` to report only the factor-variable settings for variables with the specified *design_spec*.

Remarks and examples

stata.com

▷ Example 1

Using `auto2.dta`, we include factor variable `i.rep78` in a regression:

```
. use http://www.stata-press.com/data/r14/auto2
(1978 Automobile Data)
. regress mpg i.rep78, baselevels
```

Source	SS	df	MS	Number of obs	=	69
				F(4, 64)	=	4.91
Model	549.415777	4	137.353944	Prob > F	=	0.0016
Residual	1790.78712	64	27.9810488	R-squared	=	0.2348
				Adj R-squared	=	0.1869
Total	2340.2029	68	34.4147485	Root MSE	=	5.2897

mpg	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
rep78						
Poor	0	(base)				
Fair	-1.875	4.181884	-0.45	0.655	-10.22927	6.479274
Average	-1.566667	3.863059	-0.41	0.686	-9.284014	6.150681
Good	.6666667	3.942718	0.17	0.866	-7.209818	8.543152
Excellent	6.363636	4.066234	1.56	0.123	-1.759599	14.48687
_cons	21	3.740391	5.61	0.000	13.52771	28.47229

We specified the `baselevels` option so that the base level would be included in the output. By default, the first level is the base level. We can change the base level to 2:

```
. fvset base 2 rep78
. regress mpg i.rep78, baselevels
```

Source	SS	df	MS	Number of obs	=	69
				F(4, 64)	=	4.91
Model	549.415777	4	137.353944	Prob > F	=	0.0016
Residual	1790.78712	64	27.9810488	R-squared	=	0.2348
				Adj R-squared	=	0.1869
Total	2340.2029	68	34.4147485	Root MSE	=	5.2897

mpg	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
rep78						
Poor	1.875	4.181884	0.45	0.655	-6.479274	10.22927
Fair	0	(base)				
Average	.3083333	2.104836	0.15	0.884	-3.896559	4.513226
Good	2.541667	2.247695	1.13	0.262	-1.948621	7.031954
Excellent	8.238636	2.457918	3.35	0.001	3.32838	13.14889
_cons	19.125	1.870195	10.23	0.000	15.38886	22.86114

Let's set rep78 to have no base level and fit a cell-means regression:

```
. fvset base none rep78
. regress mpg i.rep78, noconstant
```

Source	SS	df	MS	Number of obs	=	69
				F(5, 64)	=	227.47
Model	31824.2129	5	6364.84258	Prob > F	=	0.0000
Residual	1790.78712	64	27.9810488	R-squared	=	0.9467
				Adj R-squared	=	0.9426
Total	33615	69	487.173913	Root MSE	=	5.2897

mpg	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
rep78						
Poor	21	3.740391	5.61	0.000	13.52771	28.47229
Fair	19.125	1.870195	10.23	0.000	15.38886	22.86114
Average	19.43333	.9657648	20.12	0.000	17.504	21.36267
Good	21.66667	1.246797	17.38	0.000	19.1759	24.15743
Excellent	27.36364	1.594908	17.16	0.000	24.17744	30.54983

◁

▷ Example 2

By default, margins assumes that factor variables are to be treated asobserved and accumulates a margin by using the observed relative frequencies of the factor levels or the sum of the weights if weights have been specified.

```
. regress mpg i.foreign
```

Source	SS	df	MS	Number of obs	=	74
				F(1, 72)	=	13.18
Model	378.153515	1	378.153515	Prob > F	=	0.0005
Residual	2065.30594	72	28.6848048	R-squared	=	0.1548
				Adj R-squared	=	0.1430
Total	2443.45946	73	33.4720474	Root MSE	=	5.3558

mpg	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
foreign						
Foreign	4.945804	1.362162	3.63	0.001	2.230384	7.661225
_cons	19.82692	.7427186	26.70	0.000	18.34634	21.30751

```
. margins
Predictive margins                              Number of obs     =         74
Model VCE    : OLS
Expression   : Linear prediction, predict()
```

	Margin	Delta-method Std. Err.	t	P>|t|	[95% Conf.	Interval]
_cons	21.2973	.6226014	34.21	0.000	20.05616	22.53843

Let's set `foreign` to always accumulate using equal relative frequencies:

```
. fvset design asbalanced foreign
. regress mpg i.foreign
```

Source	SS	df	MS		
				Number of obs	= 74
				F(1, 72)	= 13.18
Model	378.153515	1	378.153515	Prob > F	= 0.0005
Residual	2065.30594	72	28.6848048	R-squared	= 0.1548
				Adj R-squared	= 0.1430
Total	2443.45946	73	33.4720474	Root MSE	= 5.3558

mpg	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
foreign						
Foreign	4.945804	1.362162	3.63	0.001	2.230384	7.661225
_cons	19.82692	.7427186	26.70	0.000	18.34634	21.30751

```
. margins
Adjusted predictions                            Number of obs     =         74
Model VCE    : OLS
Expression   : Linear prediction, predict()
at           : foreign          (asbalanced)
```

	Margin	Delta-method Std. Err.	t	P>|t|	[95% Conf.	Interval]
_cons	22.29983	.6810811	32.74	0.000	20.94211	23.65754

Suppose that we issued the `fvset design` command earlier in our session and that we cannot remember which variables we set as `asbalanced`. We can retrieve this information by using the `fvset report` command:

```
. fvset report, design(asbalanced)
  Variable    Base    Design
  ------------------------------
  foreign             asbalanced
  ------------------------------
```

◁

❑ Technical note

`margins` is aware of a factor variable's design setting only through the estimation results it is working with. The design setting is stored by the estimation command; thus changing the design setting between the estimation command and `margins` will have no effect. For example, the output from the following two calls to `margins` yields the same results:

```
. fvset clear foreign
. regress mpg i.foreign
```

Source	SS	df	MS		
Model	378.153515	1	378.153515	Number of obs	= 74
Residual	2065.30594	72	28.6848048	F(1, 72)	= 13.18
				Prob > F	= 0.0005
Total	2443.45946	73	33.4720474	R-squared	= 0.1548
				Adj R-squared	= 0.1430
				Root MSE	= 5.3558

mpg	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
foreign						
Foreign	4.945804	1.362162	3.63	0.001	2.230384	7.661225
_cons	19.82692	.7427186	26.70	0.000	18.34634	21.30751

```
. margins
Predictive margins                              Number of obs     =         74
Model VCE    : OLS
Expression   : Linear prediction, predict()
```

	Delta-method Margin	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_cons	21.2973	.6226014	34.21	0.000	20.05616	22.53843

```
. fvset design asbalanced foreign
. margins
Predictive margins                              Number of obs     =         74
Model VCE    : OLS
Expression   : Linear prediction, predict()
```

	Delta-method Margin	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_cons	21.2973	.6226014	34.21	0.000	20.05616	22.53843

❑

Stored results

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

Macros

`r(varlist)`	*varlist*
`r(baselist)`	base setting for each variable in *varlist*
`r(designlist)`	design setting for each variable in *varlist*