graph twoway lfit — Twoway linear prediction plots

Description
twoway lfit calculates the prediction for yvar from a linear regression of yvar on xvar and plots the resulting line.

Quick start
A linear fit prediction plot for y on x
twoway lfit y x

A scatterplot with line of best fit
twoway scatter y x || lfit y x

A separate graph area for each level of catvar
twoway scatter y x || lfit y x, by(catvar)

Distinct lines and points for catvar = 0 and catvar = 1 in the same graph area
twoway scatter y x if catvar==0 || scatter y x if catvar==1 || lfit y x if catvar==0 || lfit y x if catvar==1

Add the title “My Title” to a scatterplot with line of best fit
twoway scatter y x || lfit y x, title("My Title")

Add the title “X Variable” to the x axis
twoway scatter y x || lfit y x, title("My Title") xtitle("X Variable")

Display the line of best fit as a dashed black line
twoway scatter y x || lfit y x, lcolor(black) lpattern(dash)

Menu
Graphics > Twoway graph (scatter, line, etc.)
Syntax

twoway lfit yvar xvar [if] [in] [weight] [, options]

<table>
<thead>
<tr>
<th>options</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>range(#)</td>
<td>range over which predictions calculated</td>
</tr>
<tr>
<td>n(#)</td>
<td>number of prediction points</td>
</tr>
<tr>
<td>atobs</td>
<td>calculate predictions at xvar</td>
</tr>
<tr>
<td>estopts(regress_options)</td>
<td>options for regress</td>
</tr>
<tr>
<td>predopts(predict_options)</td>
<td>options for predict</td>
</tr>
<tr>
<td>cline_options</td>
<td>change look of predicted line</td>
</tr>
<tr>
<td>axis_choice_options</td>
<td>associate plot with alternative axis</td>
</tr>
<tr>
<td>twoway_options</td>
<td>titles, legends, axes, added lines and text, by, regions, name, aspect ratio, etc.</td>
</tr>
</tbody>
</table>

All options are rightmost; see [G-4] Concept: repeated options.

yvar and xvar may contain time-series operators; see [U] 11.4.4 Time-series varlists.

aweights, fweights, and pweights are allowed. Weights, if specified, affect estimation but not how the weighted results are plotted. See [U] 11.1.6 weight.

Options

range(#) specifies the x range over which predictions are to be calculated. The default is range(. .), meaning the minimum and maximum values of xvar. range(0 10) would make the range 0 to 10, range(. 10) would make the range the minimum to 10, and range(0 .) would make the range 0 to the maximum.

n(#) specifies the number of points at which predictions over range() are to be calculated. The default is n(3).

atobs is an alternative to n(). It specifies that the predictions be calculated at the xvar values. atobs is the default if predopts() is specified and any statistic other than the xb is requested.

estopts(regress_options) specifies options to be passed along to regress to estimate the linear regression from which the line will be predicted; see [R] regress. If this option is specified, estopts(nocons) is also often specified.

predopts(predict_options) specifies options to be passed along to predict to obtain the predictions after estimation by regress; see [R] regress postestimation.

cline_options specify how the prediction line is rendered; see [G-3] cline_options.

axis_choice_options associate the plot with a particular y or x axis on the graph; see [G-3] axis_choice_options.

twoway_options are a set of common options supported by all twoway graphs. These options allow you to title graphs, name graphs, control axes and legends, add lines and text, set aspect ratios, create graphs over by() groups, and change some advanced settings. See [G-3] twoway_options.
Remarks and examples

Remarks are presented under the following headings:

Typical use
Cautions
Use with by()

Typical use

twoway lfit is nearly always used in conjunction with other twoway plottypes, such as

. use https://www.stata-press.com/data/r16/auto
(1978 Automobile Data)
. scatter mpg weight || lfit mpg weight

Results are visually the same as typing

. regress mpg weight
. predict fitted
. scatter mpg weight || line fitted weight
Cautions

Do not use `twoway lfit` when specifying the `axis_scale_options` `yscale(log)` or `xscale(log)` to create log scales. Typing

```
.scatter mpg weight, xscale(log) || lfit mpg weight
```

The line is not straight because the regression estimated for the prediction was for `mpg` on `weight`, not `mpg` on `log(weight)`. (The default for `n()` is 3 so that, if you make this mistake, you will spot it.)

Use with `by()`

`lfit` may be used with `by()` (as can all the `twoway` plot commands):

```
.scatter mpg weight || lfit mpg weight ||, by(foreign, total row(1))
```

Also see

[G-2] graph twoway fpfit — Twoway fractional-polynomial prediction plots

[G-2] graph twoway line — Twoway line plots

[G-2] graph twoway mband — Twoway median-band plots

[G-2] graph twoway mspline — Twoway median-spline plots

[G-2] graph twoway qfit — Twoway quadratic prediction plots

[G-2] graph twoway lfitci — Twoway linear prediction plots with CIs

[R] regress — Linear regression