Graphics tricks for models

Bill Rising

StataCorp LP

2011 Australian/New Zealand Stata Users Group Meeting
Fremantle, WA
September 17, 2011

Goals

- Show two new Stata graphics tools
  - Plotting marginal effects and predictive margins using `marginsplot`
  - Making contour plots via `twoway contour` and `twoway contourline`

Getting Started

- This will be an interactive demonstration
  - This will help show some of the new interface as well as some other new tools
- Start by opening up the `nhanes2` dataset

```
. webuse nhanes2
```

  - Good for continuation from yesterday
- These are survey data

```
. svyset
```

  - We will need to use the `svy:` prefix for estimation

Pros and Cons of `margins`

- The `margins` command is great!
  - Can easily compute averaged predicted values
  - Can easily compute average marginal effects
  - Both of these make tables of values
- Curse the `margins` command!
  - The tabular output can be difficult to read

A Simple Regression with a Quadratic

- Here is a simple model for BMI as a function of age and sex

```
. svy: regress bmi c.age##c.age i.sex
```

- Can see that the parabola is concave down, and that women have smaller BMIs then men

Taking a Look at Predictive Margins

- We can take a look at the predictive margins at several ages within each sex

```
. margins, at(age==(25(10)75)) over(sex)
```

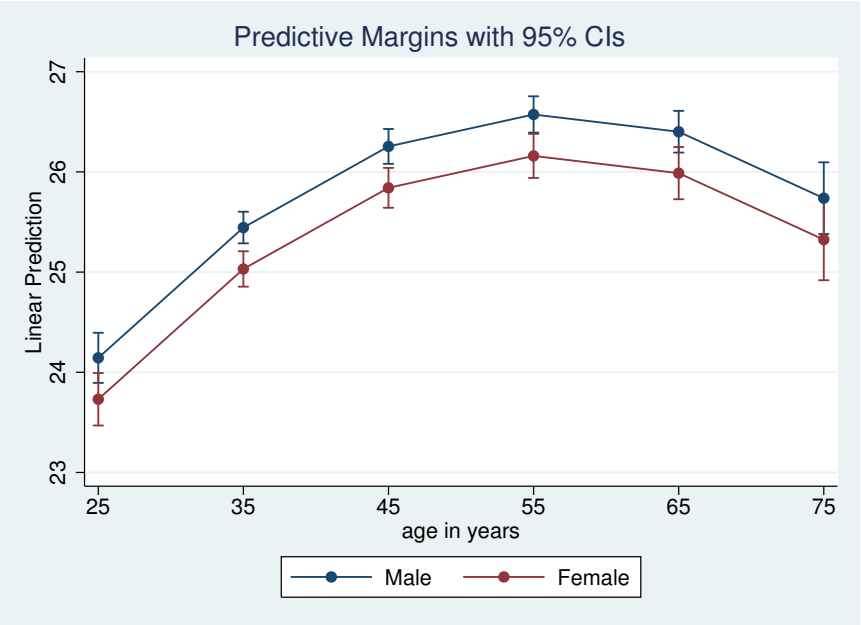
  - **Note:** Because of using `svy`, we should really be specifying `vce(unconditional)` for all these examples; this is being left off to keep the commands short

- It looks like the differences between sexes are constant (as expected)
- As one would expect from the concave-down parabola the BMI's increase and then decrease

Visualizing the Predictive Margins

- We can get a much better look at the margins by using the `marginsplot` command
- Here it is in its simplest form

```
. marginsplot
```



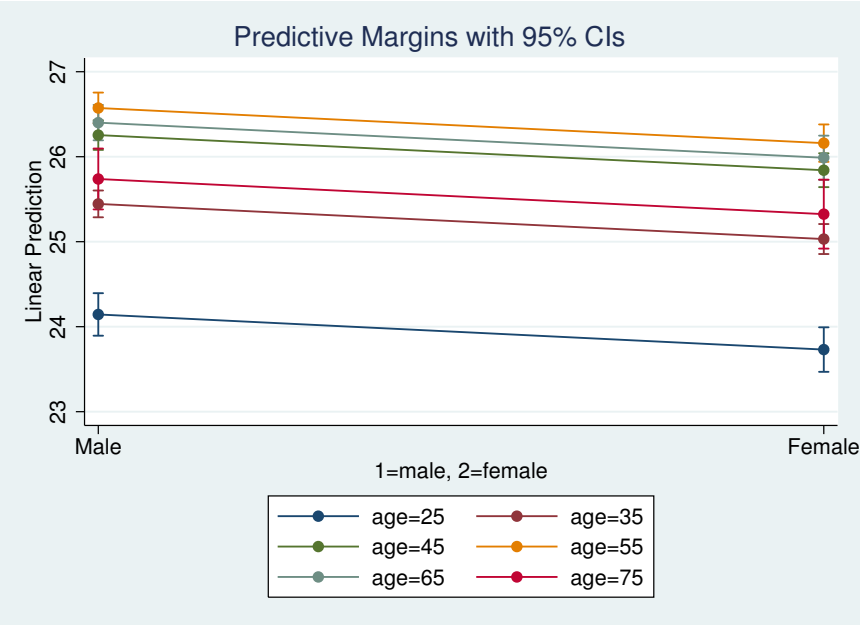
Default Behavior

- The default behavior is to
  - Draw a connected-line plot
  - Draw pointwise confidence intervals at each point
- Stata is bright enough to use the `at()` variable for the x-axis and to overlay the two curves

Switching Axes

- We can change this default behavior, of course
- Here are the same data in a less-useful form (though it shows the constant offset for the females)

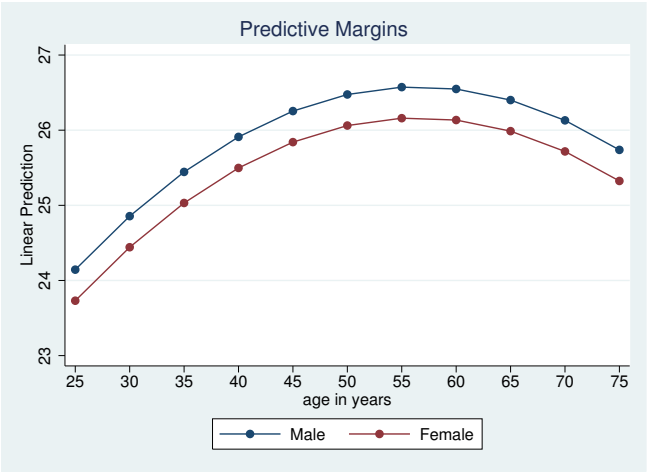
```
. marginsplot, x(sex)
```



Making a Finer Grid

- We can also make a finer grid and omit the points
  - We'll also turn off the CIs here

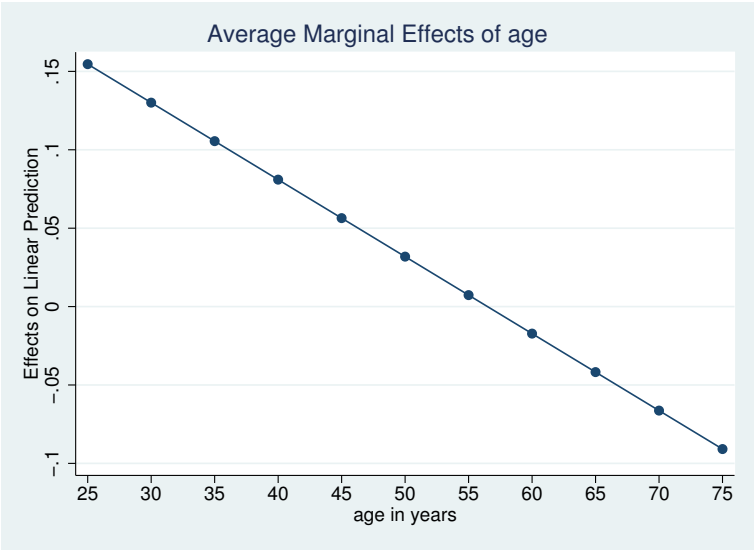
```
. quietly margins, at(age==(25(5)75)) over(sex)
. marginsplot, noci
```



Average Marginal Effects

- We can also look at average marginal effects

```
. quietly margins, at(age==(25(5)75)) dydx(age)
. marginsplot, noci
```



Looking at a Logit

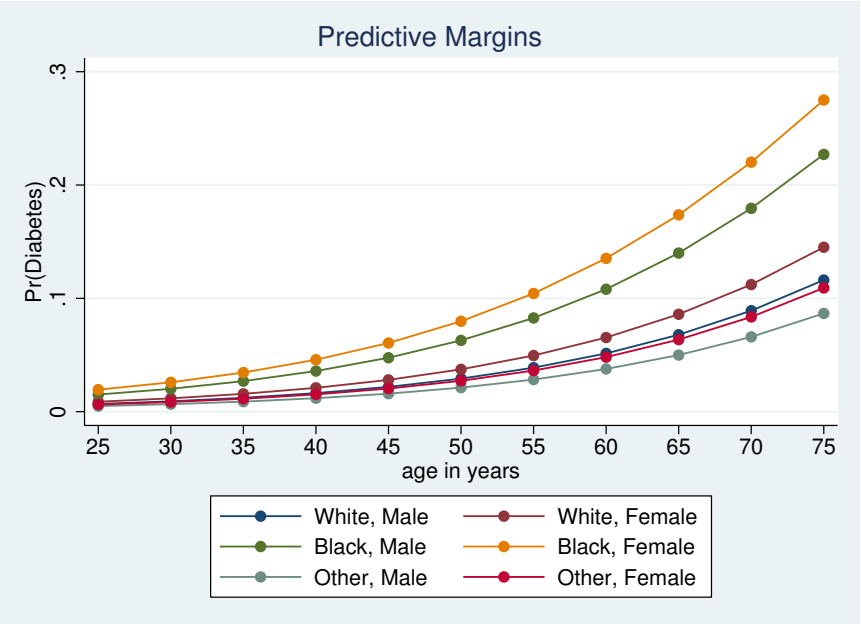
- Working with something which is non-linear in the natural (not the model) measure is a bit more interesting
- Here is a simple logistic regression

```
. svy: logistic diabetes age i.(race sex)
```

The Predictive Margins are More Interesting

- Now the predictive margins are more interesting

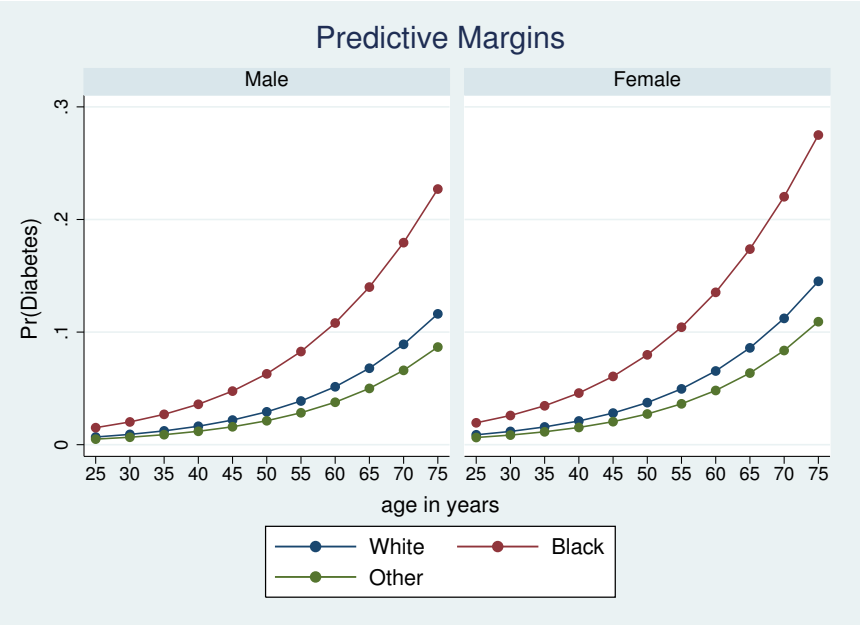
```
. quietly margins, at(age==(25(5)75)) over(race sex)
. marginsplot, noci
```



Making Side-by-Side graphs

- We can use `plotdimension()` option to make side-by-side graphs
- This graphs race by sex

```
. marginsplot, noci plot(race)
```



Now for Something Complicated

- We'll now fit a model with (too) many interactions

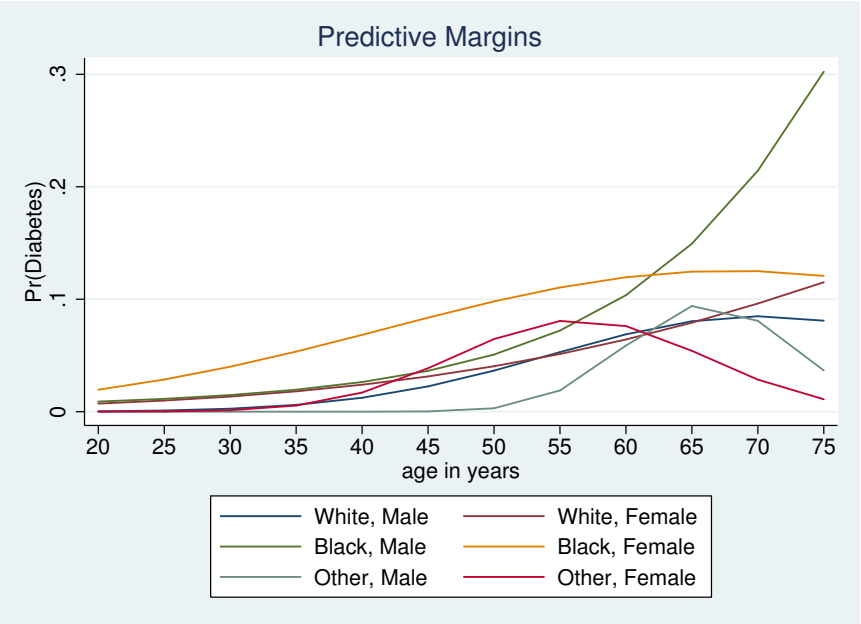
```
. svy: logit diabetes c.age##c.age##race##sex
```

- Looking at the interaction terms is nearly worthless
  - This is probably getting close to overfitting

Predictive Margins, One Last Time

- Here are the same set of margins applied to this different model

```
. quietly margins, at(age=(20(5)75)) over(race sex)
. marginsplot, noci recast(line)
```



Contour Plots

- Stata 12 now has both filled contour plots and contour line plots
- These are both `twoway` plots
  - `twoway contour` for filled contour plots
  - `twoway contourline` for contour line plots
- They are in the *twoway* dialog

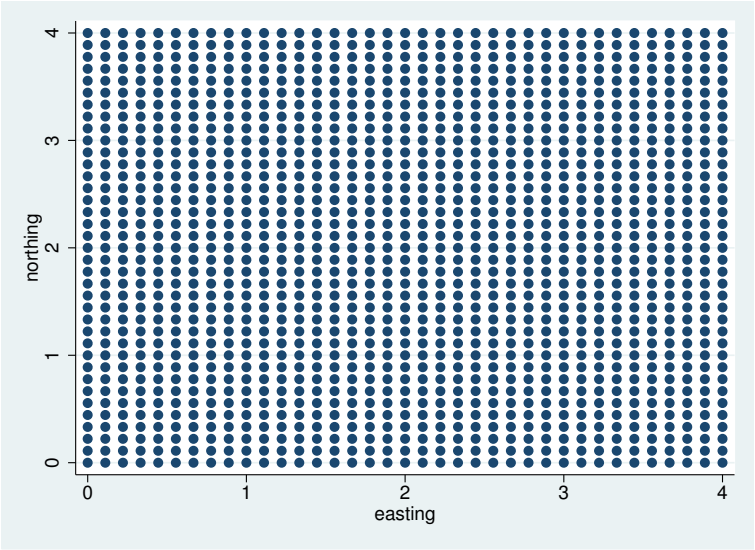
A Simple Example Dataset

- Here is a dataset meant to be artificial mountains

```
. use mtns
```

- These are heights of mountains measured on this grid

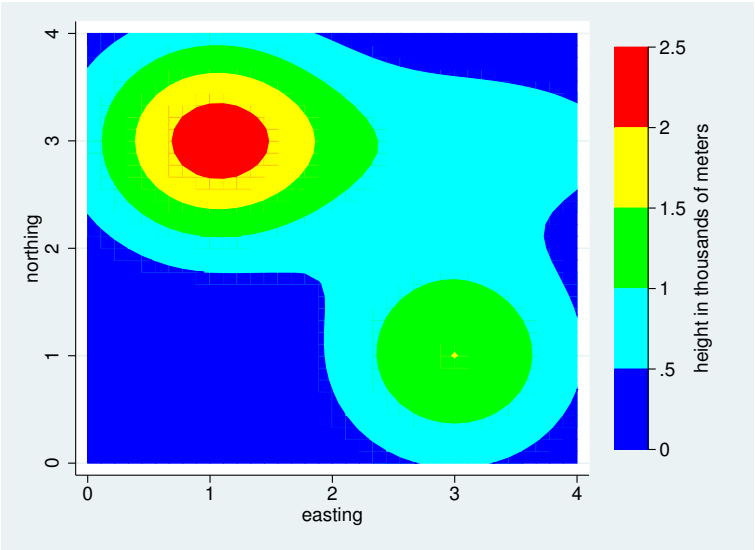
```
. scatter y x
```



Making a Filled Contour

- Here is the default filled contour

```
. twoway contour z y x
```



- Note: If this displays with line artifacts, turn off anti-aliasing in your pdf viewer—this is a known limitation of pdfs

Making the Contour Prettier

- Adding many levels smooths the color gradations

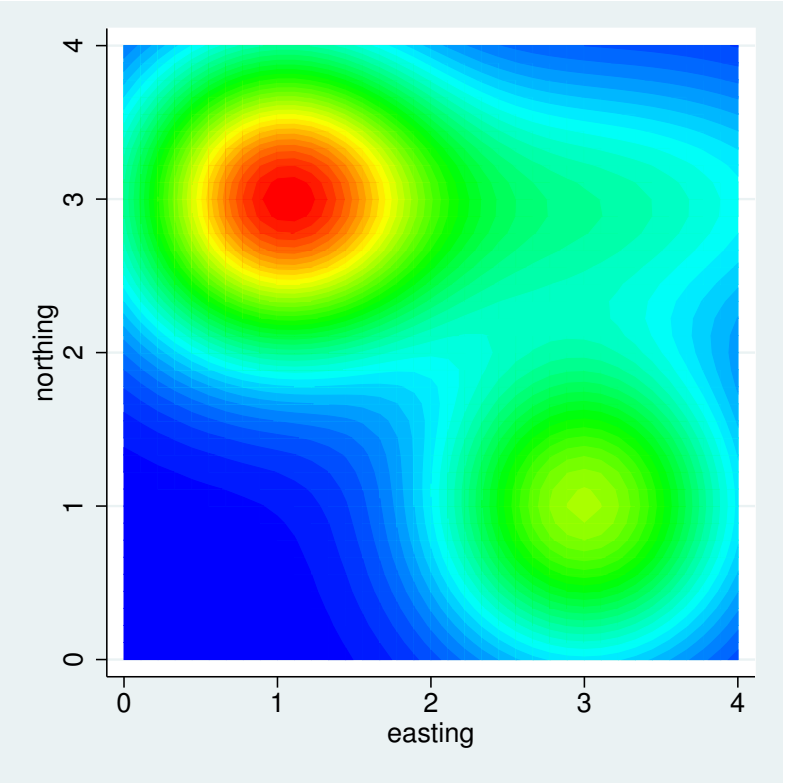
```
. twoway contour z y x, levels(40)
```

- This makes the legend a bit absurd, though
- We should shut off the legend and change the aspect ratio

A Pretty Contour Plot

- This looks quite good

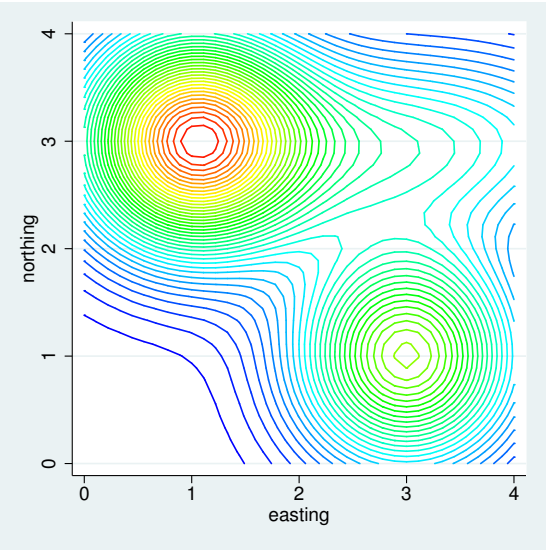
```
. twoway contour z y x, levels(40) clegend(off) xsize(5) ysize(
```



A Contour Line Plot

- Contour line plots outline elevations instead of filling them in
- They work best with color lines on

```
. twoway contourline z y x, levels(40) colorlines plegend(off)
```



What If There is No Grid?

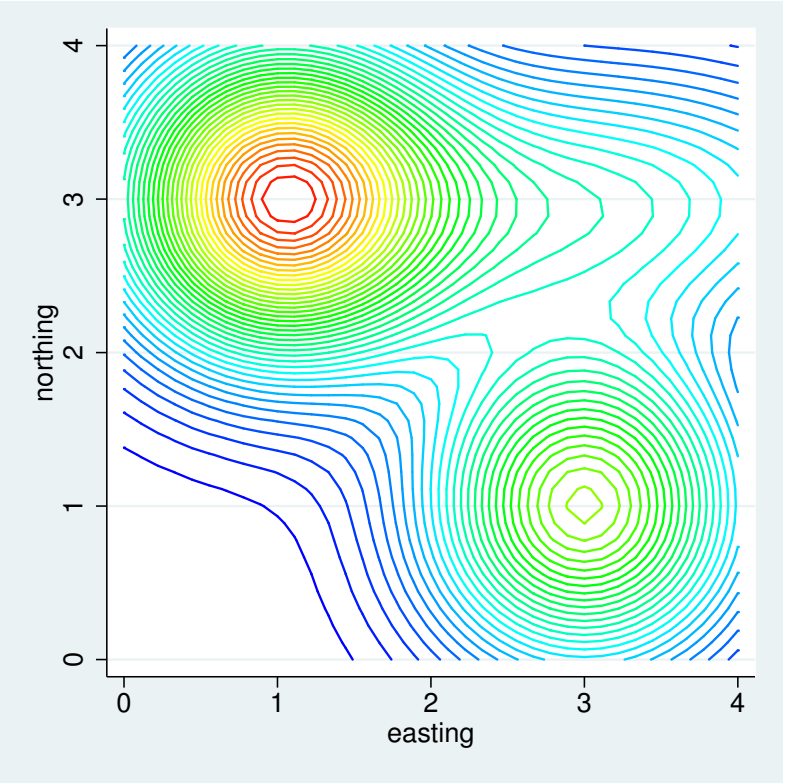
- The artificial mountain dataset was defined on a grid
- Both `twoway contour` and `twoway contourline` will use interpolation to fill the rest of the plot region if the data are not on a grid
- Let's look at this Excel file about a sandstone stratum under Ohio
- First we'll bring it into Stata

```
. import excel using sandstone, firstrow
```

A Picture of the Lack of Grid

- Looking at it, there are gaps in the grid

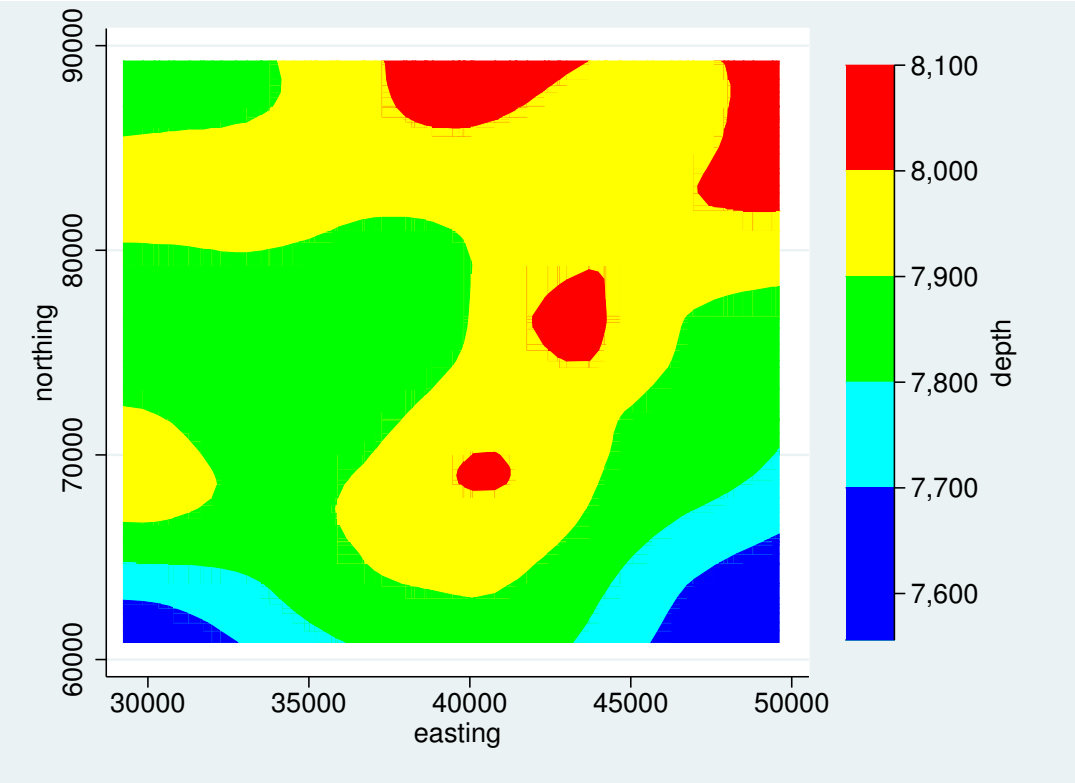
```
. scatter northing easting
```



We Can Still Make a Contour Plot

- We can still make a contour plot from this

```
. twoway contour depth northing easting
```



Taking A Look at Help

- We should take a look at the help for contour plots
- The help files have been spruced up, so that we can skip to the Remarks right away!

Just For Fun

- Believe it or not, it is possible to make a contour plot of predictive margins

```
. do margcon2
```

- Here is the picture