Margins and the Tao of Interaction

2010 Boston Stata Conference

Phil Ender

UCLA Statistical Consulting Group

July 2010

Prolog

The **margins** command is new in Stata 11. But interactions have, of course, been around a lot longer.

When it comes to deconstructing and understanding interactions **margins** is your best friend.

In fact, the **margins** command is more flexible and versatile than anything found in S*S, S*SS, or even *.

Why?

Because, **margins** groks interactions.

General comments

In addition to the two predictor variables, each of the models will also include a continuous covariate that is not part of the interaction.

In general, continuous covariates, which are not part of the interaction, are easy to deal with in linear models. However, the same is not true in nonlinear models where the values for covariates can make a large difference.

About the output

Please note the output has been heavily edited for space considerations.

And yes, I know, there are way too many numbers on most of the screens.

Categorical by Categorical Interaction

Meet the model

```
. use http://www.ats.ucla.edu/stat/data/hsbanova, clear

. anova write c.read grp##female
        Number of obs =      200      R-squared     =  0.5008
        Root MSE      = 6.83602      Adj R-squared =  0.4799

  Source |  Partial SS    df       MS           F     Prob > F
  -------+----------------------------------------------------
    read |  3818.04142     1  3818.04142     81.70     0.0000
     grp |  776.490174     3  258.830058      5.54     0.0011
  female |  1328.81274     1  1328.81274     28.44     0.0000
grp#femal|  427.388047     3  142.462682      3.05     0.0299
         |
Residual |  8925.65863   191   46.731197
  -------+----------------------------------------------------
   Total |   17878.875   199   89.843593
```

Margins time - compute the 8 adjusted cell means

```
. estimates store m1
. margins grp#female, asbalanced post

Expression   : Linear prediction, predict()
---------------------------------------------------------------
     |       Delta-method
     |  Margin  Std. Err.    z    P>|z|     [95% Conf. Interval]
-----+---------------------------------------------------------
grp#female
1 0  |   44.99   1.468    30.65   0.000     42.11     47.87
1 1  |   53.64   1.348    39.78   0.000     50.99     56.28
2 0  |   48.80   1.492    32.70   0.000     45.87     51.72
2 1  |   55.78   1.423    39.19   0.000     52.99     58.57
3 0  |   50.91   1.297    39.25   0.000     48.36     53.45
3 1  |   55.71   1.222    45.60   0.000     53.32     58.11
4 0  |   55.27   1.610    34.32   0.000     52.12     58.43
4 1  |   55.78   1.361    40.99   0.000     53.11     58.44
```

Collect values for graphing

The Kronecker product can be very useful in generating sequences of numbers.

```
. matrix m = e(b)'

. matrix g = (1\2\3\4)#(1\1)

. matrix f = (1\1\1\1)#(0\1)

. matrix m = g,f,m

. svmat m
```

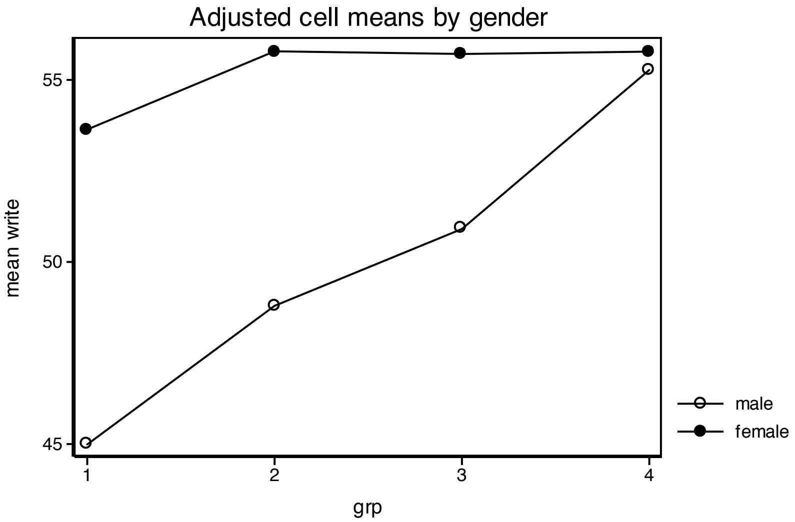
Here's what matrix m looks like

```
. matrix list m

m[8,3]
               c1:         c1:
               c1          c1          y1
r1:r1           1           0   44.988635
r1:r2           1           1   53.637754
r2:r3           2           0   48.795544
r2:r4           2           1   55.784561
r3:r5           3           0   50.905864
r3:r6           3           1   55.711471
r4:r7           4           0   55.273395
r4:r8           4           1    55.77617
```

Graph it

```
. graph twoway  ///
  (connect m3 m1 if m2==0)(connect m3 m1 if m2==1), ///
  title(Adjusted cell means by gender)              ///
  ytitle(mean write) xtitle(grp)                    ///
  legend(order(1 "male" 2 "female")) scheme(lean1)
```

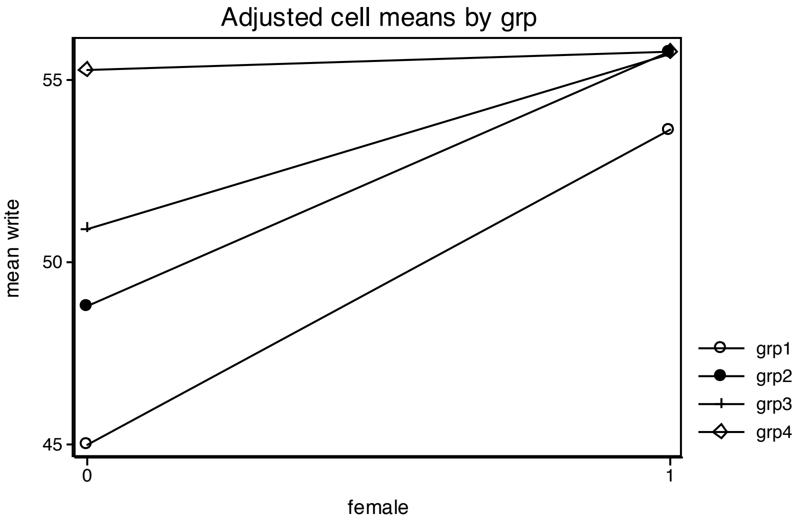
The view by gender



Graph it again

```
. graph twoway  ///
  (connect m3 m2 if m1==1)(connect m3 m2 if m1==2)   ///
  (connect m3 m2 if m1==3)(connect m3 m2 if m1==4),  ///
  title(Adjusted cell means by grp) xlabel(0 1)      ///
  ytitle(mean write) xtitle(female)                  ///
  legend(order(1 "grp1" 2 "grp2" 3 "grp3" 4 "grp4")) ///
  scheme(lean1)
```

The view by grp



Tests of simple main effects: female at grp (screen 1)

```
. test 1.grp#0.female = 1.grp#1.female  /* @ grp=1 */

( 1)  1bn.grp#0bn.female - 1bn.grp#1.female = 0
           chi2(  1) =   19.87
         Prob > chi2 =    0.0000

. test 2.grp#0.female = 2.grp#1.female  /* @ grp=2 */

( 1)  2.grp#0bn.female - 2.grp#1.female = 0
           chi2(  1) =   11.43
         Prob > chi2 =    0.0007
```

Tests of simple main effects: female at grp (screen 2)

```
. test 3.grp#0.female = 3.grp#1.female  /* @ grp=3 */

( 1)  3.grp#0bn.female - 3.grp#1.female = 0
           chi2(  1) =    7.37
         Prob > chi2 =    0.0066

. test 4.grp#0.female = 4.grp#1.female  /* @ grp=4 */

( 1)  4.grp#0bn.female - 4.grp#1.female = 0
           chi2(  1) =    0.06
         Prob > chi2 =    0.8079
```

Tests of simple main effects: grp at female (screen 1)

```
. test (1.grp#0.female = 2.grp#0.female) ///
       (1.grp#0.female = 3.grp#0.female) ///
       (1.grp#0.female = 4.grp#0.female)     /* @ female=0 */

( 1)  1bn.grp#0bn.female - 2.grp#0bn.female = 0
( 2)  1bn.grp#0bn.female - 3.grp#0bn.female = 0
( 3)  1bn.grp#0bn.female - 4.grp#0bn.female = 0

           chi2(  3) =   22.19
         Prob > chi2 =    0.0001
```

Tests of simple main effects: grp at female (screen 2)

```
. test (1.grp#1.female = 2.grp#1.female) ///
       (1.grp#1.female = 3.grp#1.female) ///
       (1.grp#1.female = 4.grp#1.female)      /* @ female=1 */

 ( 1)  1bn.grp#1.female - 2.grp#1.female = 0
 ( 2)  1bn.grp#1.female - 3.grp#1.female = 0
 ( 3)  1bn.grp#1.female - 4.grp#1.female = 0

           chi2(  3) =    1.86
         Prob > chi2 =    0.6028
```

Alternate method

The method just shown computed the simple main effects using the individual adjusted cell means.

An alternative approach uses the **dydx()** option to compute the simple main effects directly from the **margins** output.

Simple main effects for female at grp

```
. estimates restore m1
. margins grp, dydx(female) asbalanced post

Expression   : Linear prediction, predict()
------------------------------------------------------------------
    |            Delta-method
    |      dy/dx   Std. Err.     z     P>|z|     [95% Conf. Int.]
----+-------------------------------------------------------------
grp |
 1  |   8.649119   1.940487    4.46   0.000    4.846   12.452
 2  |   6.989016   2.067385    3.38   0.001    2.937   11.041
 3  |   4.805607   1.770274    2.71   0.007    1.336    8.275
 4  |   .5027748    2.06747    0.24   0.808   -3.549    4.555
------------------------------------------------------------------
```

Simple main effects for grp at female

```
. estimates restore m1
. margins female, dydx(grp) asbalanced post

Expression   : Linear prediction, predict()
--------------------------------------------------------------------
   |            Delta-method
   |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Int.]
---+----------------------------------------------------------------
2.grp
0  |   3.806909   2.099492     1.81   0.070    -.3080197    7.921838
1  |   2.146807   1.917853     1.12   0.263    -1.612115    5.905729
3.grp
0  |   5.917229   1.978675     2.99   0.003     2.039097    9.795361
1  |   2.073717   1.848835     1.12   0.262    -1.549932    5.697366
4.grp
0  |   10.28476   2.236825     4.60   0.000     5.900663    14.66886
1  |   2.138416   1.951454     1.10   0.273    -1.686363    5.963195
--------------------------------------------------------------------
```

Results: Simple main effects for grp at female

```
. test ([2.grp]0.female=0)([3.grp]0.female=0)([4.grp]0.female=0)

 ( 1)  [2.grp]0bn.female = 0
 ( 2)  [3.grp]0bn.female = 0
 ( 3)  [4.grp]0bn.female = 0

           chi2(  3) =   22.19
         Prob > chi2 =    0.0001

. test ([2.grp]1.female=0)([3.grp]1.female=0)([4.grp]1.female=0)

 ( 1)  [2.grp]1.female = 0
 ( 2)  [3.grp]1.female = 0
 ( 3)  [4.grp]1.female = 0

           chi2(  3) =    1.86
         Prob > chi2 =    0.6028
```

Categorical by Continuous Interaction

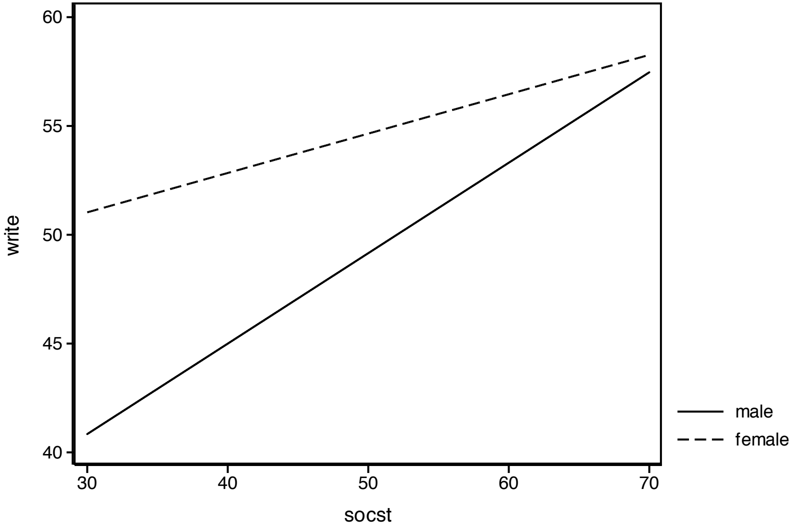
Regression model w/ categorical by continuous interaction

```
. regress write read female##c.socst, noheader

. estimates store m1

------------------------------------------------------------------
   write |   Coef. Std. Err.  t     P>|t|  [95% Conf. Interval]
---------+--------------------------------------------------------
    read |  .3747    .0584    6.41   0.000   .2595     .4899
1.female |  17.23    4.658    3.70   0.000   8.046     26.42
   socst |  .4156    .0693    6.00   0.000   .2790     .5522
         |
  female#|
 c.socst | -.2347    .0870   -2.70   0.008  -.4063    -.0631
         |
   _cons |  8.802    3.527    2.50   0.013   1.846     15.76
```

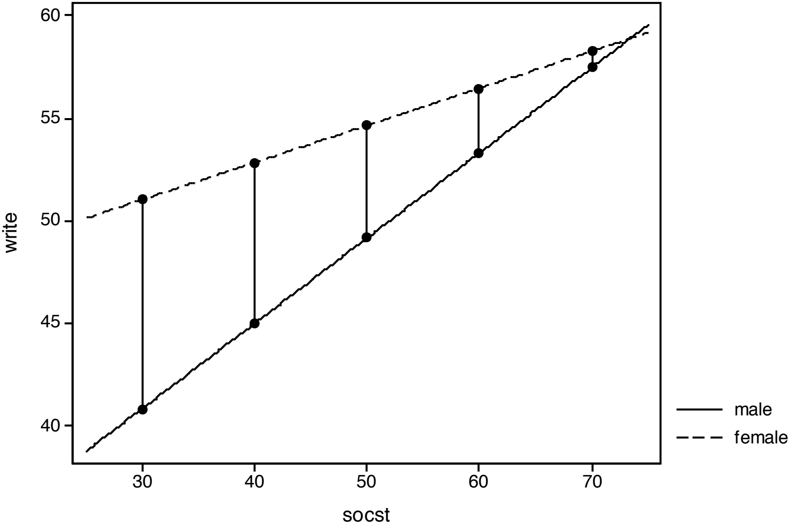
Getting slopes and intercepts

```
. margins female, dydx(socst)    /* slopes */

Average marginal effects                    Number of obs   = 200

socst        |      dy/dx   Std. Err.      z    P>|z|
  female     |
          1  |   .4156419   .0692631    6.00   0.000
          2  |    .180911   .0721559    2.51   0.012

. margins female, at(socst=0)   /* intercepts */

Predictive margins                          Number of obs   = 200
             |     Margin   Std. Err.      z    P>|z|
          1  |   28.37166   3.636821    7.80   0.000
          2  |   45.60334   3.884672   11.74   0.000
```

Graph of simple slopes by gender



Suppose we want gender difference at 5 values of socst



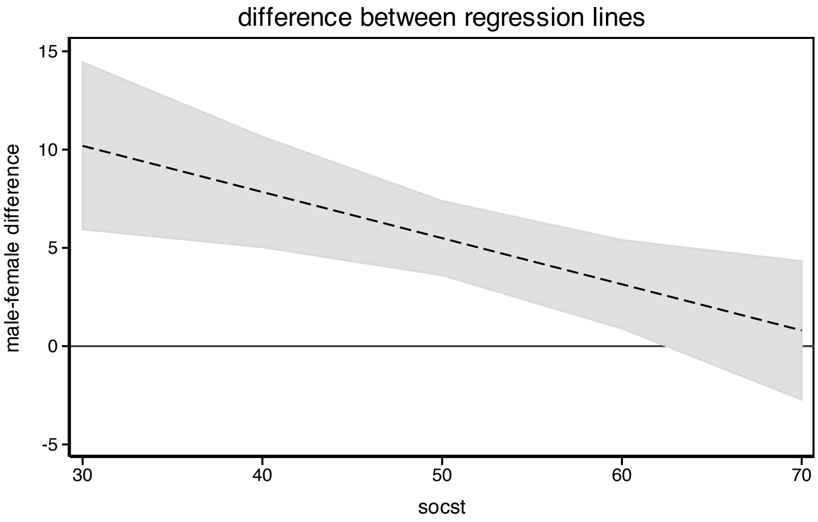
Margins - adjusted means

```
. margins female, at(socst=(30(10)70)) post noatlegend

Expression   : Linear prediction, predict()
------------------------------------------------------------------
     |       Delta-method
     |  Margin  Std. Err.    z     P>|z|    [95% Conf. Interval]
-----+------------------------------------------------------------
_at#female
1 0  |   40.84   1.644    24.84   0.000     37.62     44.06
1 1  |   51.03   1.784    28.61   0.000     47.54     54.53
2 0  |   44.99   1.056    42.60   0.000     42.93     47.07
2 1  |   52.84    1.14    46.46   0.000     50.61     55.07
(output omitted)
5 0  |   57.47   1.452    39.59   0.000     54.62     60.31
5 1  |   58.27   1.371    42.50   0.000     55.58     60.95
-------------------------------------------------------------------
```

Margins - differences in adjusted means

```
. estimates restore m1
. margins, dydx(female) at(socst=(30(10)70)) noatlegend

Average marginal effects                      Number of obs    = 200
Expression: Linear prediction - dy/dx w.r.t. : 1.female
------------------------------------------------------------
   |         Delta-method
   |    dy/dx  Std. Err.    z      P>|z|    [95% Conf. Interval]
---+--------------------------------------------------------
1.female _at
1  |    10.19   2.166     4.70   0.000     5.945     14.43
2  |    7.842   1.433     5.47   0.000     5.035     10.65
3  |    5.495   .9633     5.70   0.000     3.607     7.383
4  |     3.15   1.148     2.74   0.006     .8977     5.398
5  |    .8005   1.795     0.45   0.656    -2.718     4.319
Note: dy/dx for factor levels is the discrete change from the ba
```

Graph of differences in adjusted means



About the graph

We will show a detailed example of creating a graph like this in the last section for the Bonus Interaction.

Continuous by Continuous Interaction

Regression model w/ continuous by continuous interaction

```
. regress read write c.math##c.socst, noheader

------------------------------------------------------------------
    read |    Coef.   Std. Err.    t     P>|t| [95% Conf. Interval]
---------+--------------------------------------------------------
   write |     .193    .0714     2.71   0.007     .0524     .3340
    math |   -.2285    .2903    -0.79   0.432    -.8011     .3441
   socst |   -.3206    .2700    -1.19   0.237    -.8532      .212
         |
  c.math#|
 c.socst |    .0120    .0052     2.32   0.021     .0018      .022
         |
   _cons |    37.17    14.32     2.60   0.010     8.931     65.41
```

Simple slopes the Aiken and West way

Recenter data at 3 points for one predictor:

- 1 standard deviation below the mean
- at the mean
- 1 standard deviation above the mean

then rerun regressions

In Stata 10.1, I would have used as series on **lincom** commands.

Using the **margins** command you do not need to recenter data and you can compute simple slopes for as many values as you wish.

Compute simple slopes using margins

```
. margins, dydx(math) at(socst=(30(5)75)) noatlegend

. matrix s = r(b)   /* capture slopes */

Average marginal effects       Number of obs    =  200
Expression : Linear prediction - dy/dx w.r.t. : math
    |  dy/dx   Std. Err.   z      P>|z|  [95% Conf. Interval]
----+---------------------------------------------------------
math _at
 1  |    .1308   .1448     0.90   0.366    -.1529    .4145
 2  |    .1907   .1227     1.55   0.120    -.0497    .4311
 3  |    .2506   .1023     2.45   0.014     .0500    .451
 (output omitted)
 8  |    .5500   .0865     6.36   0.000     .380     .7196
 9  |    .6099   .1043     5.85   0.000     .406     .814
10  |    .6698   .1248     5.37   0.000     .4251    .9145
```

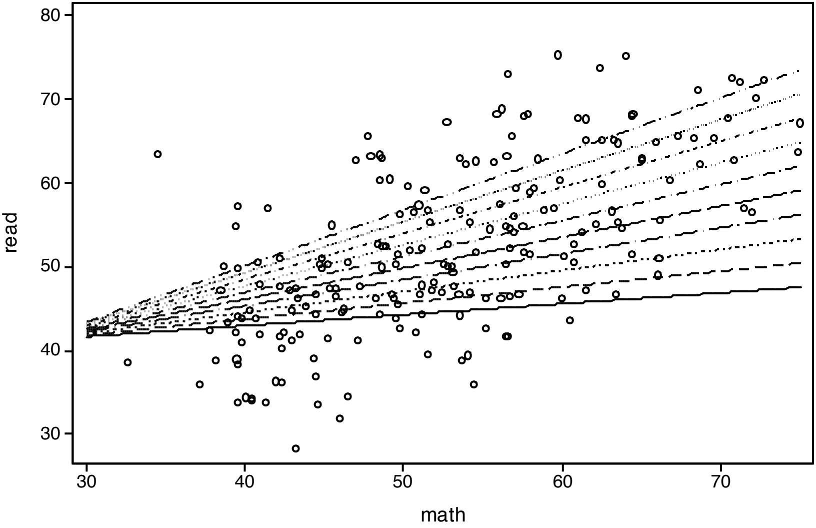
Compute intercepts using margins

```
. margins, at(math=0 socst=(30(5)75)) noatlegend

. matrix i = r(b)    /* capture intercepts */

Predictive margins     Number of obs    =   200
Expression     : Linear prediction, predict()
     |  Margin    Std. Err.  z       P>|z|    [95% Conf. Interval]
----+---------------------------------------------------------
_at |
 1  |   37.7     7.180      5.26   0.000     23.67     51.82
 2  |   36.14    6.057      5.97   0.000     24.27     48.02
 3  |   34.54    5.046      6.85   0.000     24.65     44.43
(output omitted)
 8  |   26.53    4.903      5.41   0.000     16.91     36.14
 9  |   24.92    5.890      4.23   0.000     13.38     36.47
10  |   23.32    6.999      3.33   0.001       9.6     37.04
```

Graph it

```
. graph twoway  ///
  (function y = i[1, 1] + s[1, 1]*x, range(30 75)) ///
  (function y = i[1, 2] + s[1, 2]*x, range(30 75)) ///
  (function y = i[1, 3] + s[1, 3]*x, range(30 75)) ///
  (function y = i[1, 4] + s[1, 4]*x, range(30 75)) ///
  (function y = i[1, 5] + s[1, 5]*x, range(30 75)) ///
  (function y = i[1, 6] + s[1, 6]*x, range(30 75)) ///
  (function y = i[1, 7] + s[1, 7]*x, range(30 75)) ///
  (function y = i[1, 8] + s[1, 8]*x, range(30 75)) ///
  (function y = i[1, 9] + s[1, 9]*x, range(30 75)) ///
  (function y = i[1,10] + s[1,10]*x, range(30 75)) ///
  (scatter read math, msym(oh) jitter(3)),         ///
  xlabel(30(10)75) legend(off) ytitle(read)        ///
  xtitle(math) scheme(lean1)
```

Simple slopes for 10 values of socst from 30 to 75



Bonus Interaction

Categorical by continuous logistic interaction

Logistic regression model

```
. use http://www.ats.ucla.edu/stat/data/logitcatcon, clear

. logit y cv1 i.f##c.s, nolog noheader

-----------------------------------------------------------------
     y |  Coef   Std. Err.    z      P>|z|    [95% Conf. Interval]
------+----------------------------------------------------------
   cv1 | .1877    .0348      5.40   0.000      .1195     .256
   1.f | 9.984     3.05      3.27   0.001      4.001     15.97
     s | .1751    .0470      3.72   0.000      .0829     .2672
       |
f#c.s  |
    1  |-.1595    .0570     -2.80   0.005     -.2713    -.0477
       |
_cons  |-19.01    3.371     -5.64   0.000     -25.61    -12.39
```

Hold cv1 constant, let s vary (probability metric)

```
. margins f, at(s=(40 50 60) cv1=50) noatlegend
Adjusted predictions    Number of obs    = 200
Expression    : Pr(y), predict()
     |     Margin   Std. Err.       z    P>|z|     [95% Conf. Interval]
-----+------------------------------------------------------------------
1 0  |    .068052   .0448994     1.52   0.130    -.0199492    .1560531
1 1  |   .7282426   .0816421     8.92   0.000     .5682269    .8882582
2 0  |   .2960206   .0768246     3.85   0.000     .1454472     .446594
2 1  |   .7578957   .0512602    14.79   0.000     .6574276    .8583637
3 0  |   .7077261   .0953456     7.42   0.000     .5208521       .8946
3 1  |   .7852672   .0669634    11.73   0.000     .6540214     .916513

. margins, dydx(f) at(s=(40 50 60) cv1=50) noatlegend
Conditional marginal effects    Number of obs    = 200
Expression    : Pr(y), predict() - dy/dx w.r.t. : 1.f
    |      dy/dx   Std. Err.       z    P>|z|     [95% Conf. Interval]
----+------------------------------------------------------------------
 1  |   .6601906   .0983425     6.71   0.000     .4674428    .8529385
 2  |   .4618751   .0965359     4.78   0.000     .2726681     .651082
 3  |   .0775412   .1164177     0.67   0.505    -.1506333    .3057156
Note: dy/dx for factor levels is the discrete change from the base level.
```

Hold s constant, let cv1 vary (probability metric)

```
. margins f, at(s=50 cv1=(40 50 60)) noatlegend
Adjusted predictions      Number of obs    = 200
Expression    : Pr(y), predict()
     |   Margin   Std. Err.       z     P>|z|      [95% Conf. Interval]
-----+---------------------------------------------------------------
1 0  | .0604557   .0329478     1.83   0.067    -.0041208    .1250322
1 1  | .3238822   .0808248     4.01   0.000     .1654685    .4822959
2 0  | .2960206   .0768246     3.85   0.000     .1454472     .446594
2 1  | .7578957   .0512602    14.79   0.000     .6574276    .8583637
3 0  | .7331854   .0823937     8.90   0.000     .5716966    .8946741
3 1  | .9533959   .0227391    41.93   0.000     .9088281    .9979637

. margins, dydx(f) at(s=50 cv1=(40 50 60)) noatlegend
Conditional marginal effects     Number of obs    = 200
Expression    : Pr(y), predict() - dy/dx w.r.t. : 1.f
    |       dy/dx    Std. Err.       z     P>|z|      [95% Conf. Interval]
----+----------------------------------------------------------------
 1  |    .2634265    .0682395     3.86   0.000     .1296795    .3971735
 2  |    .4618751    .0965359     4.78   0.000     .2726681     .651082
 3  |    .2202105    .0743402     2.96   0.003     .0745063    .3659147
Note: dy/dx for factor levels is the discrete change from the base level.
```

Let both s and cv1 vary

```
. margins, dydx(f) at(s=(25(5)70) cv1=(40 50 60)) noatlegend post

Conditional marginal effects   Number of obs   =   200
Expression    : Pr(y), predict() - dy/dx w.r.t. : 1.f

    |    dy/dx    Std. Err.      z    P>|z|     [95% Conf. Interval]
----+------------------------------------------------------------
1.f _at
 1  | .2443475   .1321009     1.85   0.064    -.0145655    .5032605
 2  | .2578855   .1135271     2.27   0.023     .0353765    .4803946
 3  | .2704118   .0954463     2.83   0.005     .0833405    .4574832
 4  | .2797622   .0798258     3.50   0.000     .1233066    .4362179
(output omitted)
27  | .0884101   .0436473     2.03   0.043     .0028629    .1739572
28  | .0192749   .0303776     0.63   0.526    -.0402642    .0788139
29  |-.0116134   .0243513    -0.48   0.633     -.059341    .0361142
30  |-.0237264     .02315    -1.02   0.305    -.0690996    .0216469
Note: dy/dx for factor levels is the discrete change from the base leve
```

Capture the data for graphing

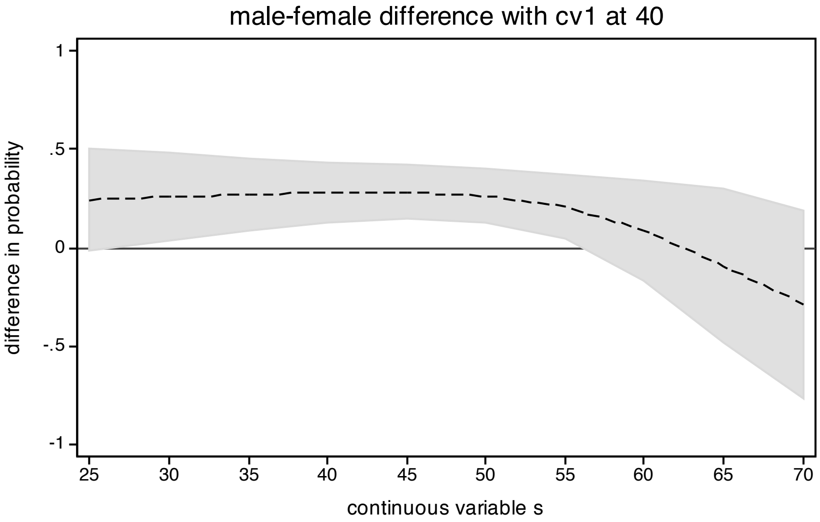
```
matrix t = J(30,3,.)
matrix cv = (40\50\60)#(1\1\1\1\1\1\1\1\1\1)
matrix iv = (1\1\1)#(25\30\35\40\45\50\55\60\65\70)

forvalues i=1/30 {
  quietly lincom _b[1.f:`i'._at]
  matrix t[`i',1] = r(estimate)
  matrix t[`i',2] = r(estimate) - 1.96*r(se)
  matrix t[`i',3] = r(estimate) + 1.96*r(se)
}

matrix t = t,iv,cv
svmat t
```

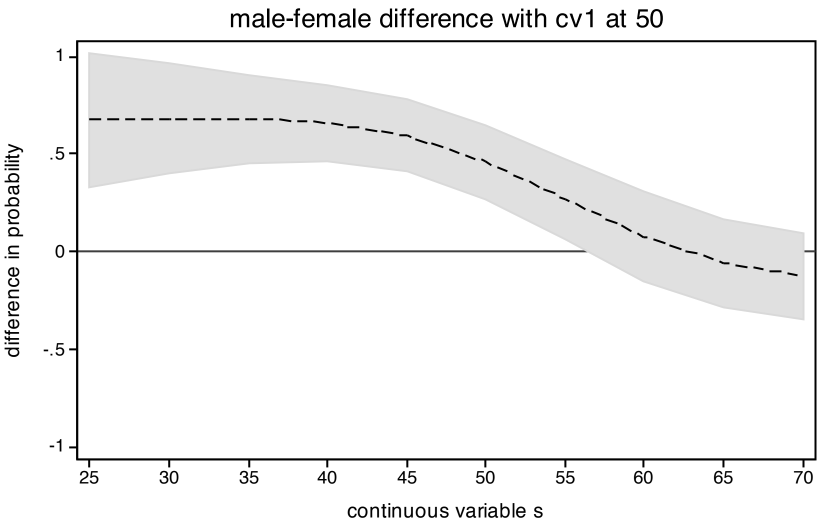
Here is what matrix t looks like

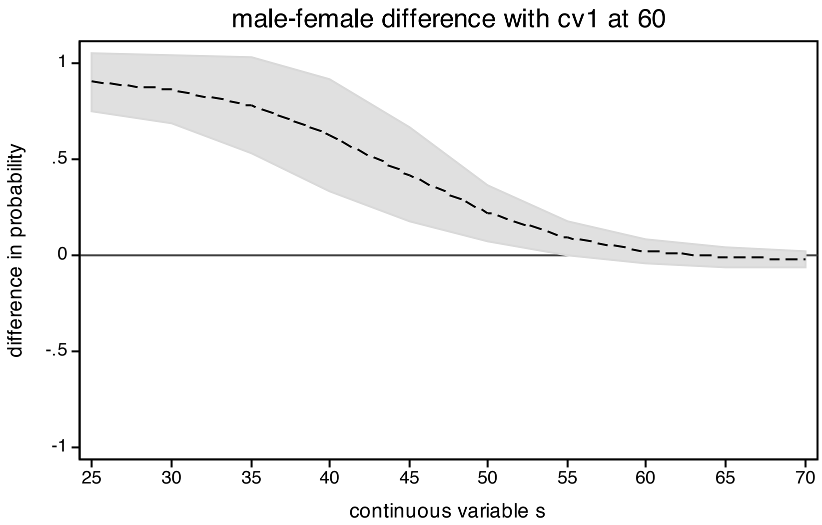
```
. matrix list t

t[30,5]
              c1          c2          c3      c1      c1
 r1     .2443475  -.01457025   .50326526      25      40
 r2    .25788554    .0353724   .48039867      30      40
 r3    .27041184   .08333707    .4574866      35      40
 r4    .27976224   .12330368   .43622079      40      40
 r5    .28098578   .14450356   .41746801      45      40
(output omitted)
r26    .22021051   .07450364   .36591739      50      60
r27    .08841006   .00286136   .17395877      55      60
r28    .01927488  -.04026528   .07881503      60      60
r29   -.01161343   -.0593419   .03611503      65      60
r30   -.02372637  -.06910046   .02164773      70      60
```

Make 3 graphs

```
forvalues i = 40(10)60 {
 graph twoway ///
  (rarea t2 t3 t4 if t5==`i', color(gs13) lcolor(gs13)) ///
  (line t1 t4 if t5==`i'), yline(0) legend(off)        ///
  xtitle(continuous variable s)                         ///
  ytitle(difference in probability)                     ///
  title(male-female difference with cv1 at `i')        ///
  scheme(lean1) xlabel(25(5)70) ylabel(-1(.5)1)        ///
  name(difference`i', replace)
}
```


male-female difference with cv1 at 40
difference in probability
1
.5
0
-.5
-1
25
30
35
40
45
50
55
60
65
70
continuous variable s


male-female difference with cv1 at 50
difference in probability
1
.5
0
-.5
-1
25
30
35
40
45
50
55
60
65
70
continuous variable s


male-female difference with cv1 at 60
difference in probability
1
.5
0
-.5
-1
25
30
35
40
45
50
55
60
65
70
continuous variable s

The End