Lab 2a: Interpreting Simple Interactions, using Excel

```
. use http://www-personal.umich.edu/~franzese/WS_Data_level1_merge_level2.dta

. reg eu_support leftright GENGOV06 lrGS

      Source |       SS       df       MS              Number of obs =    8536
-------------+------------------------------           F(  3,  8532) =    8.42
       Model |  183.363072     3  61.1210239           Prob > F      =  0.0000
    Residual |  61910.1959  8532  7.25623486           R-squared     =  0.0030
-------------+------------------------------           Adj R-squared =  0.0026
       Total |  62093.5589  8535  7.27516801           Root MSE      =  2.6937

------------------------------------------------------------------------------
  eu_support |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   leftright |   .3102197   .0692995     4.48   0.000     .1743759    .4460636
    GENGOV06 |   .0310353   .0085924     3.61   0.000     .0141921    .0478785
        lrGS |  -.0074389    .001577    -4.72   0.000    -.0105302   -.0043475
       _cons |   5.326216   .3764811    14.15   0.000     4.588222     6.06421
------------------------------------------------------------------------------
```

Copy output from Stata to Excel, copy as text from ". reg" to end---, paste to Excel in A1, using Text Import Wizard. Coefficient on leftright should be in cell D13.

What's the effect of leftright? …of GS?

$$\frac{\partial EU\sup}{\partial \text{lftrt}} = .31 - .0074 GovSize \qquad \frac{\partial EU\sup}{\partial \text{GSize}} = .031 - .0074 lftrt$$

```
Recall that:
```

- Only for modifying effect does standard regression output tell us directly.

- What are the standard errors of these effects?
  - Only for modifying effect does standard regression output tell us directly.
- Recall: *Main Effects* refer to beyond-range values; they not direct evidence on whether effect (generally) positive

- What are the standard errors of these effects? Need this:

```
. vce

Covariance matrix of coefficients of regress model

        e(V) |  leftright    GENGOV06        lrGS       _cons
-------------+------------------------------------------------
   leftright |  .00480243
    GENGOV06 |  .00052614   .00007383
        lrGS | -.00010751   -.0000122   2.487e-06
       _cons | -.02346701  -.00318089   .00052616   .14173804
```

Copy output from Stata to Excel, copy as text from ". vce" to end last line, paste to Excel in A19, using Text Import Wizard. Est'd variance of leftright should be in cell D25.

Finally, you'll want summary stats for your variables:

```
. tabstat leftright GENGOV06 lrGS, statistics(mean max min median iqr skewness sd)

   stats |  leftri~t  GENGOV06      lrGS
---------+------------------------------
    mean |  4.977976  43.27785  215.8879
     max |        10      52.7       527
     min |         0      20.9         0
     p50 |         5      45.3     219.5
     iqr |         2      10.2     108.4
skewness | -.0185371 -1.222382  .2163898
      sd |  2.342092  7.794739  110.6589
----------------------------------------
```

Copy output from Stata to Excel, copy as text from ". tab" to end -- line, paste to Excel in A30, using Text Import Wizard. Mean of leftright should be in cell D34.

From here, you can work in Excel.

Column 1: Conditioning Variable (GENGOV06 in dEUsup/dLR)

- **1st cell (in A45) enter =$E$36 (location of min GS)**
    - **You can put a label GovSize in A44.**
- **2nd cell (in A46) enter =$A45+($E$35-$E$36)/50**
    - **i.e., previous value +(max-min)/#steps (large # for smoothness)**
- **Copy down**

Column 2: Conditional Effect (label dEUsup/dLR)

- 1st cell (in B45) enter =$D$13+$D$15*$A45
  - This is coeff(LR) + coeff(LR*GS)*GS.
  - left $ on col A references optional but makes next steps easier
- Copy down

Column 3: Standard Error (optional; can skip to bounds)

- 1st cell (in C45) enter =($D$25+$F$27*$A45^2+2*$D$27*$A45)^0.5
  - This is $[V(c(LR))+V(c(LR*GS))*GS^2+2*Cov(LR,LRGS)GS]^{.5}$
- Copy down

Column 4: Lower Bound

- Put probability level for desired c.i. in cell above label (D43). Label goes in D44.
- 1st cell (in D45) enter =$B45-TINV($D$43,$E$6)*$C45
  - This is conditional effect-critical t*s.e.(effect)
- Copy down

Column 5: Upper Bound

- Copy lower 1st cell (D45) to upper bound 1st cell (E45)
- Edit the "-" sign to a "+" sign, and copy down.

GRAPH:

- **Select B44 (label cond effect) - E95 (last up bound)**
- **Delete the s.e. line.**
- **Change x-axis labels to Column A.**
- **Pretty it up at will.**

- **Now you can copy that block over somewhere, edit slightly to plot the analogously conditional effect of Government Size on EU support.**

- **Can template the prettied graph too.**

- **Can use these templates for next models you estimate, editing location of coefficients and v-cov elements, etc. May make sense, therefore, to put these analyses in a second (& further) sheet(s).**

Alternative, more Stata-intensive, methods avoiding excel…

- In Stata, plot dY/dX w/ c.i. from smpl min-max:
  - `egen zmin = min(z)` ; `egen zmax = max(z)` finds those sample min & max for variable z. (z=leftright in our case, i.e., GS)
  - `gen z0 = (_n-1)/(v-1)*(zmax-zmin)+zmin in 1/v` creates var counting v equal-size steps from sample min to max.
  - `gen dyhatdx=_b[x]+_b[xz]*z0` creates var of dY/dX ests (x=education)
  - Stata code tedious to get to s.e.'s & c.i. plots (bit better in matrix form)
    - First have to work in matrices for bit, then back to vars:
      - `matrix V = get(VCE)` (makes matrix of v-cov mat)
      - `matrix C= V[3,1]` (grabs 3,1 element as covar)
      - `gen column1 = 1 in 1/v` (makes a variable equal to all ones)
      - `mkmat column1, matrix(col1)` (makes vector called col1 of that var)
      - `matrix cov_x_xz = C*col1` (makes a constant vector of covar)
      - `svmat cov_x_xz, name(cov_x_xz)` (makes that vector a variable)
    - Finally, you can generate variances & std errors, which you could tabulate:
      - `gen vardyhatdx=(_se[x])^2+(z0*z0)*(_se[xz]^2)+2*z0*cov_x_xz`
      - `gen sedyhatdx=sqrt(vardyhatdx)` (makes variable equal to s.e. of effect)
      - `tabdisp z0, cellvar(dyhatdx sedyhatdx)` (makes table effects & s.e.'s)
    - Or you can generate the confidence interval bounds & plot:
      - `gen LBdyhatdx=dyhatdx-invttail(e(df_r),.05)*sedyhatdx`
      - `gen UBdyhatdx=dyhatdx+invttail(e(df_r),.05)*sedyhatdx`
      - `twoway connected dyhatdx LBdyhatdx UBdyhatdx z0`

Calculating Predicted-Values & Standard Errors

Procedures	Command syntax

Create the variables which set the values of the variables *x*, *z*, and *xz* (& other variables, if any) for which $\hat{y}$ will be calculated.	`egen xmin = min(x)`
	`egen xmax = max(x)`
	`gen xh = ((_n-1)/(v-1))*(xmax-xmin) in 1/v`
	`egen zh=mean(z)`
	`gen xhzh=xh*zh`
	`gen col1h=1`

Assemble the variables into a matrix, **Mh**	`mkmat xh zh xhzh col1h in 1/v, matrix(Mh)`
Create **betas**, a column vector with *k x 1* dimensions.	`matrix betas=e(b)'`
Calculate the predicted values.	`matrix yhat=Mh*betas`
Convert the vector into a variable.	`svmat yhat, name(yhat)`
Create a matrix from the estimated covariance matrix of the coefficient estimates.	`matrix V = get(VCE)`
Calculate the variance of the predicted values.	`matrix VYH=Mh*V*Mh'`
Extract the diagonal elements into a row vector.	`matrix DVYH= vecdiag(VYH)`
Transpose elements into a column vector.	`matrix VARYHAT=DVYH'`
Convert the vector into a variable.	`svmat VARYHAT, name(varyhat)`
Calculate the estimated standard error of each predicted probability.	`gen seyhat1 = sqrt(varyhat1)`
Present a table of predicted values with corresponding standard errors of the predicted values.	`tabdisp xh, cellvar(yhat1 seyhat1)`

Generate lower and upper confidence interval bounds. Graph the predicted probabilities and the upper and lower confidence intervals.	`gen LByhat1=yhat1-invttail(e(df_m),.05)*seyhat1`
	`gen UByhat1=yhat1+invttail(e(df_m),.05)*seyhat1`
	
	`twoway connected yhat1 LByhat1 UByhat1 xh`

Google

'kam franzese michigan press' for that data, and stata & excel resources or go directly to `http://www.press.umich.edu/KamFranzese/Interactions.html`

Google

'Matt Golder interaction' or go directly to `http://homepages.nyu.edu/~mrg217/interaction.html`

Stata

```
help mfx
help predictnl
```

HOMEWORK: Restimate as follows. Explain why we might want to “fe” and “robust” here. Regenerate the dEUsup/dLeftRight conditional-effect w/ c.i. graph.

```
. xtreg eu_support leftright GENGOV06 lrGS, fe vce(robust)
```