

Advanced Reporting of Regression Results

Introduction

Most papers these days don't involve reporting out regression results in the way that you see in intro textbooks. Instead, lots of times we need to run results that test various assumptions or only apply to certain parts of the data. The programming skills we've been developing all year long really come into their own in these situations.

Tables of Descriptives

One of the most common types of descriptive tables that needs to be run in modern papers is one that establishes baseline equivalence. This is the idea that the control variables should look similar at different levels of the treatment variable. The easiest version of this to run is when you have a binary treatment variable.

In the table in the do file, I regress the treatment variable on a series of control variables, and do so across different levels of test scores. This table tests whether the treatment (expects to go to college) has similar baseline characteristics across different levels of test scores.

```
. // Balance test: how different is the key covariate (treatment variable) by levels
> of the control variables
.
. foreach test_level of numlist 0(1)3
.   2.
.
. //Counter variable
. local counter=1
.   3.
. foreach race_var of local race
.   4.     quietly reg `race_var' `x' if test_group==`test_level'
.   5.
.   scalar my_diff = round(_b[`x'], `mysig')
.   6.
.   scalar my_t = round(_b[`x']/_se[`x'], `mysig')
.   7.
.   mat M= [my_diff_t]
.   8.
.   if `counter'==1
.   9.     mat M_col=M
.  10.
.  11.     else mat M_col=(M_col)
.  12.     local counter=`counter'+1
.  13. //end loop over race variables
.  14.
.   if `test_level'==0
.  15.     mat results_tab=M_col
.  16.
.  17.     else mat results_tab=(results_tab,M_col)
```

```

18.
. // End loop over test scores

.
. matrix rownames results_tab="Native American" "t value" "Asian" "t value" "Afric
> an American" "t value" "Hispanic" "t value" "Multiracial" "t value"

.
. matrix colnames results_tab ="Lowest Quartile" "2nd Quartile" "3rd Quartile" "4th
> Quartile"

.
. // Table
.
. estout matrix(results_tab) using "baseline_tab.`ttype`", style(tex) replace
(output written to baseline_tab.tex)

.

```

Quick? Exercise

Add a column for the full sample and a row for female to the above table.

Regression Results

Many studies these days are “single coefficient” studies. The idea is to test whether a treatment has an impact on an outcome, but to do so across multiple different specifications and/or samples. If the impact of the treatment is consistent, that’s good news for the analysis.

In the example below, I run a regression of whether or not the individual expects to go to college on whether they actually went to college, but do so with and without controls and at different levels of test scores.

```

. // Regression results
.
. foreach test_level of numlist 0(1)3
2.
. quietly reg `y' `x' if test_group==`test_level'
3.
. scalar my_coeff = round(_b[`x'], `mysig')
4.
. scalar my_se =round(_se[`x'],`mysig')
5.
. scalar my_n=round(e(N))
6.
. mat M1= [my_coeff_se_n]
7.
. quietly reg `y' `x' `controls' if test_group==`test_level'
8.
. scalar my_coeff = round(_b[`x'], `mysig')
9.
. scalar my_se =round(_se[`x'],`mysig')
10.
. scalar my_n=round(e(N))
11.
. mat M2= [my_coeff_se_n]
12.
. mat M=(M1,M2)
13.
. if `test_level'==0
14.     mat reg_results=M

```

```

15.
16.     else mat reg_results=(reg_results,M)
17.
. // end loop over test levels
.

```

Quick? Exercise

Create another table that separates results by SES levels.

Graphics

Similar to the above, we many times want to create graphics that are for different parts of the sample or involve different assumptions. The problem many times is getting the graphics on the same scale and not repeating the legend. The totally awesome *grc1leg2* package can do this for us. In the example, I create a scatterplot of test scores by SES, with separate colors for those who do and don't plan to go to college, and separate it out by level of test score.

```

. // Complex Graphics
.
. local test_level=0

. foreach test_level of numlist 0(1)3
2. local quartile=`test_level'+1
3.
. graph twoway (scatter byncls2m byses1 if test_group==`test_level' & expect_college
> ==0,msize(vtiny) color(red) mcolor(%10)) ///
> (lfit byncls2m byses1 if test_group==`test_level' & expect_college==0
> ,lwidth(thin) lcolor(red)) ///
> (scatter byncls2m byses1 if test_group==`test_level' & expect_college
> ==1,msize(vtiny) color(blue) mcolor(%10)) ///
> (lfit byncls2m byses1 if test_group==`test_level' & expect_college==1,l
> width(thin) color(blue)), ///
> legend(order(2 "Doesn't expect to go to college" 4 "Expects to go to college")) yt
> itle("Test Scores") xtitle("SES") title("Quartile=`quartile'")
4.
. graph save "scatter_`quartile'.gph", replace
5.
(file scatter_1.gph saved)
(file scatter_2.gph saved)
(file scatter_3.gph saved)
(file scatter_4.gph saved)

.
. // Combine all levels
.
. grc1leg2 scatter_1.gph scatter_2.gph scatter_3.gph scatter_4.gph, legendfrom("scat
> ter_1.gph") rows(2) name(scatter,replace) xcommon ycommon

.
. graph export scatter.eps, replace
(file scatter.eps written in EPS format)
.

```

Quick? Exercise

Create a graphic that shows attendance in college as a function of reading test scores, for males and females, across four different SES levels.