**Finished and Revised FEDAI Stata Code Memo**

**by yi yin**

**3.8 revise:**

\*part a

before: scalar ate\_texas = el(r(b),1,3) - el(r(b),1,129

after: scalar ate\_texas = el(r(b),1,3) - el(r(b),1,1)

\*part b

se from naive regression is not suitable

\*part c

before : scalar total\_length = r(N)

after: scalar total\_length = \_N

\*part f

se from naive regression is not suitable

estimated ate is corrected for blocked assignment

**3.9 revise**

\* part b

before: ritest treatment \_b[treatment], strata(pair) reps(10000): regress preexperimentbets treatment

after: ritest treatment e(F), strata(pair) reps(10000) seed(1234): regress treatment preexperimentbets

**4.2 revise**

\*part a

before: ritest treat \_b[treat], reps(10000): regress pretest treat

after: tsrtest D r(Fs) using Fstat.dta, overwrite: Fstat

note: the p-value is different from R result due to rounding, the full schedule of permutation is exactly the same. to exam the F statistics of every possible permutation, we can use the file which save the statistics.

\*part b

generate full schedule of permutation, sorting the statistic to obtain 95% CI

note: the 95%CI is different from R result due to rounding, the full schedule of permutation is exactly the same. to exam the F statistics of every possible permutation, we can use the file which save the statistics.

\*part c

generate full schedule of permutation, sorting the statistic to obtain 95% CI

the 95%CI is the same as the R result.

**4.4 revise**

\*part f

finish

obtain CI by sorting the statistics from permutation

note: similar 95% CI but not exactly the same since the permutation is random

\*part g

finish

obtain CI by sorting the statistics from permutation

note: similar 95% CI but not exactly the same since the permutation is random

**4.5**

Finish

note: p-values, standard errors are slightly different since the probabilities under restricted randomization are calculated based on 10000 random permutations. But the calculation procedures are exactly the same since I also use R to read the permutation results from the Stata, which renders the same statistics.

**4.6 revise:**

\*part b

ensure blocked assignment

regression to check balance

\*part d

Finished

**4.10:**

Finished

note: p-values are slightly different due to random permutations

**5.11**

\*part c

Resolve “variable not defined” error:

sum d if z=="treatment"

scalar pr\_c\_treatment = r(mean)

sum d if z=="placebo"

scalar pr\_c\_placebo = r(mean)

**7.6**

finished

**8.5**

Resolve name conflict**:** y to Y

**8.9**

\*part b

revise, fit.3 to right coefficient by categorizing exposure

note: better to use csv file data, since dta file treat some variables as characters.

\*part c

Correct regression result due to data file miss-specificaiton.

**8.10**

Finish

**8.11**

\*part b

fix a typo: prob00zz to prob00z2

**9.6**

finished

**9.7**

\* part c

finish 3 regression model left before, generate journal style regression results for 4 models.

**9.9**

\* part c

data mis-matched on FEIDA website

should be Fieldhouse\_et\_al\_unpublished\_2010\_expanded.dta

NOT Fieldhouse\_et\_al\_unpublished\_2010.dta

**10.4**

Finish: use *tsrtest* module to calculate the full schedule of permutation.

\* part j

Note: Stata deal with perfect collinearity in a slightly different way which do not generate missing value in coefficient. Therefore, omitting missing values works in R but not Stata for this problem.

Have to manually remove the subset which have NA in R (but not missing in Stata), the script generates the small result in the solution file.

**10.4**

Finish

**11.6**

Finish part e

**11.8**

Finish part a,

the data mis-match on FEIDA website, should be GerberGreenBook\_Chapter11\_Table\_11\_3.csv

NOT Dupas\_WorkingPaper\_2010.dta

\* part b

the data mis-match on FEIDA website, should be Chapter\_11\_Dupas\_(2010)\_Dataset.csv (converted from Chapter 11\_Dupas (2010) Dataset.dta)

NOT Dupas\_WorkingPaper\_2010.dta

**11.9**

\* part b

Finish

the data mis-match on FEIDA website, should be Chapter\_11\_Dupas\_(2010)\_Dataset.csv (converted from Chapter 11\_Dupas (2010) Dataset.dta)

NOT Dupas\_WorkingPaper\_2010.dta

**12.11**

\* part h

fix “u\_b\_count not found error”

**13.1**

Finished

note: p-values, standard errors are slightly different since the probabilities under restricted randomization are calculated based on 10000 random permutations. But the calculation procedures are exactly the same since I also use R to read the permutation results from the Stata, which renders the same statistics.