R code to Stata code for Examples in FEDAI Textbook

Contents

FEDAI Page 158

Box 5.4 on FEDAI Page 158

```
itt_fit <- lm(VOTED ~ ASSIGNED)
         coeftest(itt_fit,vcovHC(itt_fit))


t test of coefficients:

            Estimate Std. Error t value  Pr(>|t|)
(Intercept) 0.375376   0.006446 58.2344 < 2.2e-16 ***
ASSIGNED    0.038464   0.014479  2.6565  0.007914 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Box 5.4 Stata Version

```
reg VOTED ASSIGNED, vce(hc3)


Linear regression                               Number of obs   =      7,090
                                                F(1, 7088)      =       7.06
                                                Prob > F        =     0.0079
                                                R-squared       =     0.0010
                                                Root MSE        =     .48599


------------------------------------------------------------------------------
             |               Robust HC3
      VOTED |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   ASSIGNED |   .0384644   .0144795     2.66   0.008     .0100802    .0668485
      _cons |   .3753764    .006446    58.23   0.000     .3627404    .3880124
------------------------------------------------------------------------------
```

FEDAI Page 159

Box 5.5 on FEDAI Page 159

```
itt_d_fit <- lm(TREATED ~ ASSIGNED)
        coeftest(itt_d_fit,vcovHC(itt_d_fit))


t test of coefficients:

            Estimate Std. Error t value  Pr(>|t|)
(Intercept) 2.5284e-14 6.3517e-16  39.806 < 2.2e-16 ***
ASSIGNED    2.7336e-01 1.1733e-02  23.299 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Box 5.5 Stata Version

```
reg TREATED ASSIGNED, vce(hc3)


Linear regression                               Number of obs   =       7,090
                                                F(1, 7088)      =      542.84
                                                Prob > F        =      0.0000
                                                R-squared       =      0.2305
                                                Root MSE        =      .20123


------------------------------------------------------------------------------
             |               Robust HC3
    TREATED  |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    ASSIGNED |   .2733564   .0117325    23.30   0.000     .2503571    .2963557
       _cons |   -1.26e-14   4.41e-10    -0.00   1.000    -8.64e-10    8.64e-10
------------------------------------------------------------------------------
```

FEDAI Page 160

Box 5.6 on FEDAI Page 160

```
cace_fit <- ivreg(VOTED ~ TREATED,~ASSIGNED)
        coeftest(cace_fit,vcovHC(cace_fit))



t test of coefficients:

            Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.375376   0.006446 58.2344   <2e-16 ***
TREATED     0.140711   0.052434  2.6836   0.0073 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Box 5.6: Stata Version

```
ivregress 2sls VOTED (TREATED =ASSIGNED), vce(robust)


Instrumental variables (2SLS) regression          Number of obs   =      7,090
                                                   Wald chi2(1)    =       7.21
                                                   Prob > chi2     =     0.0072
                                                   R-squared       =     0.0063
                                                   Root MSE        =     .48464


------------------------------------------------------------------------------
             |               Robust
      VOTED |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    TREATED |   .1407115   .0524015     2.69   0.007     .0380065    .2434165
      _cons |   .3753764   .0064448    58.24   0.000     .3627448    .3880081
------------------------------------------------------------------------------
Instrumented:   TREATED
Instruments:    ASSIGNED
```

FEDAI Page 186

Box 6.3 on FEDAI Page 186

```
lm(formula = TREATED ~ ASSIGNED)

Residuals:
    Min      1Q  Median      3Q     Max
-0.3663 -0.3663 -0.1616  0.6337  0.8384

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.16162    0.01931   8.367  < 2e-16 ***
ASSIGNED     0.20472    0.02718   7.532 1.12e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4297 on 998 degrees of freedom
Multiple R-squared:  0.05379,Adjusted R-squared:  0.05284
F-statistic: 56.73 on 1 and 998 DF,  p-value: 1.116e-13
```

Box 6.3 Stata Version

```
reg TREATED ASSIGNED


      Source |       SS           df       MS      Number of obs   =     1,000
-------------+----------------------------------   F(1, 998)       =     56.73
       Model |  10.4765702         1  10.4765702   Prob > F        =    0.0000
    Residual |   184.29843       998  .184667765   R-squared       =    0.0538
-------------+----------------------------------   Adj R-squared   =    0.0528
       Total |     194.775       999   .19496997   Root MSE        =   .42973


------------------------------------------------------------------------------
     TREATED |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    ASSIGNED |   .2047205   .0271799     7.53   0.000     .1513842    .2580567
       _cons |   .1616162   .0193149     8.37   0.000     .1237136    .1995187
------------------------------------------------------------------------------
```

Box 6.4 on FEDAI Page 187

```
lm(formula = Y ~ ASSIGNED)

Residuals:
    Min      1Q  Median      3Q     Max
-0.4753 -0.4753 -0.4182  0.5248  0.5818

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.41818    0.02233  18.725   <2e-16 ***
ASSIGNED     0.05707    0.03143   1.816   0.0697 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4969 on 998 degrees of freedom
Multiple R-squared:  0.003293,Adjusted R-squared:  0.002294
F-statistic: 3.297 on 1 and 998 DF,  p-value: 0.06969
```

Box 6.4 Stata Version

```
reg Y ASSIGNED


      Source |       SS           df       MS      Number of obs   =     1,000
-------------+----------------------------------   F(1, 998)       =      3.30
       Model |  .814042304         1  .814042304   Prob > F        =    0.0697
    Residual |  246.376958       998  .246870699   R-squared       =    0.0033
-------------+----------------------------------   Adj R-squared   =    0.0023
       Total |    247.191       999  .247438438   Root MSE        =    .49686


------------------------------------------------------------------------------
           Y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    ASSIGNED |   .0570657   .0314258     1.82   0.070    -.0046025     .118734
       _cons |   .4181818   .0223322    18.73   0.000     .3743583    .4620053
------------------------------------------------------------------------------
```

6

FEDAI Page 188

Box 6.5 on FEDAI Page 188

```
out.tsls <- tsls(Y ~ TREATED, ~ ASSIGNED)
summary(out.tsls)


 2SLS Estimates

Model Formula: Y ~ TREATED

Instruments: ~ASSIGNED

             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.3731314  0.0434628 8.58508  < 2e-16 ***
TREATED     0.2787494  0.1529948 1.82195 0.068761 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4952064 on 998 degrees of freedom
```

Box 6.5 Stata Version

```
ivregress 2sls Y (TREATED =ASSIGNED)


Instrumental variables (2SLS) regression       Number of obs   =      1,000
                                                Wald chi2(1)    =       3.33
                                                Prob > chi2     =     0.0682
                                                R-squared       =     0.0099
                                                Root MSE        =     .49471


------------------------------------------------------------------------------
         Y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   TREATED |   .2787494   .1528417     1.82   0.068    -.0208149    .5783137
     _cons |   .3731314   .0434193     8.59   0.000     .2880311    .4582317
------------------------------------------------------------------------------
Instrumented:  TREATED
Instruments:   ASSIGNED
```

FEDAI Page 199

Box 6.6 **on FEDAI Page 199**

```
coef(summary(lm(Y ~ D+Z)))
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.23418367	0.02061000	11.362621	4.898133e-29
D	0.43469388	0.02416360	17.989618	3.538618e-67
Z	-0.05867347	0.02157878	-2.719035	6.603925e-03

Box 6.6 **Stata Version**

```
reg y d z


      Source |       SS           df       MS      Number of obs   =     2,000
-------------+----------------------------------   F(2, 1997)      =    164.73
       Model |  70.6923469         2  35.3461735   Prob > F        =    0.0000
    Residual |  428.507653     1,997   .21457569   R-squared       =    0.1416
-------------+----------------------------------   Adj R-squared   =    0.1408
       Total |       499.2     1,999  .249724862   Root MSE        =    .46322


------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           d |   .4346939   .0241636    17.99   0.000     .3873054    .4820824
           z |  -.0586735   .0215788    -2.72   0.007    -.1009928   -.0163542
       _cons |   .2341837     .02061    11.36   0.000     .1937643     .274603
------------------------------------------------------------------------------
```

Box 7.1 on FEDAI Page 233

```r
# Generate a variable ("observed") indicating whether or not the unit is observed (r_i=1)
observed <- 1 - (read == 0)

# Use logistic regression to predict probabilities of being observed

probobs <- glm(observed~(vouch0*sex)+(vouch0*phone)+(vouch0*age),
               family=binomial(link="logit"))$fitted

# Compare distributions of predicted probabilities across experimental conditions
# Check to make sure that there are no zero predicted probabilities in either condition
summary(probobs[vouch0==0])

    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
0.005258 0.090593 0.295256 0.302246 0.413661 0.887591


summary(probobs[vouch0==1])

    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
0.006938 0.237711 0.449397 0.375784 0.503739 0.872052


# Generate weights: inverse of predicted probability of being observed
wt <- 1/probobs

# Restrict analysis to observed subjects.
sel_valid <- observed == 1
table(sel_valid)

sel_valid
FALSE   TRUE
 2319   1223


# Coefficients or unweighted regression (restricting analysis to observed subjects)
lm(read~vouch0,subset=sel_valid)$coefficients

(Intercept)  46.9208148492351
 vouch0       0.682737757271807


# Coefficients for IPW regression (restricting analysis to observed subjects)
lm(read~vouch0,weights=wt,subset=sel_valid)$coefficients

(Intercept)  46.4378181846688
 vouch0       0.723030349366462
```

Box 7.1 Stata Version

```stata
// Generate a variable ("observed") indicating
// whether or not the unit is observed (r_i=1)
gen observed = 1 - (read == 0)
recast int vouch0 sex age phone
gen vouch0sex = vouch0*sex
gen vouch0phone = vouch0*phone
gen vouch0age = vouch0*age

// Use logistic regression to predict probabilities of being observed
qui glm observed vouch0 sex phone age vouch0sex vouch0phone vouch0age, family(binomial)
predict probobs
```

```stata
// Compare distributions of predicted probabilities
//across experimental conditions

// Check to make sure that there are
// no zero predicted probabilities in either condition
tabstat probobs, by(vouch0) stat(min p25 med mean p75 max) nototal


// Generate weights: inverse of predicted probability of being observed
gen wt=1/probobs
```

```
Summary for variables: probobs
     by categories of: vouch0

  vouch0 |       min        p25        p50       mean        p75        max
---------+------------------------------------------------------------------
       0 |  .0052578   .0905925   .2952563   .3022464   .4136614    .887591
       1 |  .0069381   .2377109   .4493973   .3757839   .5037394   .8720517
----------------------------------------------------------------------------
```

```stata
// Restrict analysis to observed subjects.
gen sel_valid = observed == 1
tab sel_valid
```

```
 sel_valid |      Freq.     Percent        Cum.
-----------+-----------------------------------
         0 |      2,319       65.47       65.47
         1 |      1,223       34.53      100.00
-----------+-----------------------------------
     Total |      3,542      100.00
```

```stata
//Coefficients for unweighted regression
// (restricting analysis to observed subjects)

qui reg read vouch0 if sel_valid == 1
mat unweight = e(b)
mat li unweight
```

```
unweight[1,2]
       vouch0       _cons
y1   .68273776   46.920815
```

```stata
// Coefficients for IPW regression
// (restricting analysis to observed subjects)
qui reg read vouch0 [iw=wt] if sel_valid == 1
mat weighted = e(b)
mat li weighted
```

```
weighted[1,2]
       vouch0       _cons
y1   .72303035   46.437818
```