
```
      name:  Grant Aarons Assignment 4
       log:  C:\Users\gaarons\Git\Notes\Stata\2016F\Metrics\logs\stata_4.smcl
  log type:  smcl
 opened on:  13 Feb 2017, 15:37:54

 1 . * Use s or t to get smcl or text log file
 2 . /*
   > Grant Aarons
   > gaarons@london.edu
   > Econometrics 1, London Business School
   > */
 3 . global programdir C:\Users\gaarons\Git\Notes\Stata\2016F\Metrics\programs

 4 . global datadir C:\Users\gaarons\Git\Notes\Stata\2016F\Metrics\data

 5 . global outputdir C:\Users\gaarons\Git\Notes\Stata\2016F\Metrics\output

 6 . /*
   > Program: stata_41.do
   > Description: Introduction to econometrics in stata
   > */
 7 .
 8 . ********************************************************************
 9 . ************************* START QUESTION 3 *************************
10 . ********************************************************************
11 . do $programdir/stata_41.do

12 . /*
   > Grant Aarons
   > gaarons@london.edu
   > Econometrics 1, London Business School
   > Assignment 4
   > */
13 . * Load the data
14 . import excel "$datadir/problems.xls", sheet("problems12") firstrow clear

15 . * pause Dataset has been loaded
16 .
17 . * Generate all neccessary or helpful variables
18 . *gen datenum = year+(quarter-1)*(0.25)
19 . gen t = _n

20 . gen newt = tq(1948q1) + t - 1

21 . tsset newt, quarterly
        time variable:  newt, 1948q1 to 2002q4
                delta:  1 quarter

22 .
23 . gen cpta = (realconsumptionofnondurables +  realconsumptionofservices)/population

24 . gen ypta = (realdisposableincome)/population

25 .
26 . * Prettier ways of doing this but whatever
27 . gen lcpta_1 = l.cpta
   (1 missing value generated)

28 . gen lcpta_2 = l.lcpta_1
   (2 missing values generated)

29 . gen lcpta_3 = l.lcpta_2
   (3 missing values generated)

30 . gen lcpta_4 = l.lcpta_3
   (4 missing values generated)
```

```
31 . gen lcpta_5 = l.lcpta_4
   (5 missing values generated)

32 . gen lcpta_6 = l.lcpta_5
   (6 missing values generated)

33 . gen lypta_1 = l.ypta
   (1 missing value generated)

34 .
35 . * first difference income per capita
36 . gen fd_ypta = ypta-lypta_1
   (1 missing value generated)

37 . *cpta add lag 1
38 . gen fd_cpta = cpta-lcpta_1
   (1 missing value generated)

39 .
40 . * Logical thing to do, take logs and then do all the generated variables again
41 . gen ln_cpta = ln(cpta)

42 . gen ln_ypta = ln(ypta)

43 .
44 . * Prettier ways of doing this but whatever
45 . gen ln_lcpta_1 = l.ln_cpta
   (1 missing value generated)

46 . gen ln_lcpta_2 = l.ln_lcpta_1
   (2 missing values generated)

47 . gen ln_lcpta_3 = l.ln_lcpta_2
   (3 missing values generated)

48 . gen ln_lcpta_4 = l.ln_lcpta_3
   (4 missing values generated)

49 . gen ln_lcpta_5 = l.ln_lcpta_4
   (5 missing values generated)

50 . gen ln_lcpta_6 = l.ln_lcpta_5
   (6 missing values generated)

51 . gen ln_lypta_1 = l.ln_ypta
   (1 missing value generated)

52 .
53 . * logarithm of first difference income per capita
54 . gen ln_fd_ypta = ln_ypta-ln_lypta_1
   (1 missing value generated)

55 . * logarithm of first difference consumption per capita
56 . gen ln_fd_cpta = ln_cpta-ln_lcpta_1
   (1 missing value generated)

57 .
58 . gen instrument1 = ln_lcpta_2 - ln_lcpta_3
   (3 missing values generated)

59 . gen instrument2 = ln_lcpta_3 - ln_lcpta_4
   (4 missing values generated)

60 . gen instrument3 = ln_lcpta_4 - ln_lcpta_5
   (5 missing values generated)

61 . gen instrument4 = ln_lcpta_5 - ln_lcpta_6
   (6 missing values generated)
```

```
62 .
63 . *** Part A
64 . * Do the simple regression of C_t on 4 period lags of C_t and a constant
65 . reg cpta lcpta_1 lcpta_2 lcpta_3 lcpta_4
```

Source	SS	df	MS	Number of obs	=	216
				F(4, 211)	>	99999.00
Model	.004822372	4	.001205593	Prob > F	=	0.0000
Residual	1.2145e-06	211	5.7561e-09	R-squared	=	0.9997
				Adj R-squared	=	0.9997
Total	.004823586	215	.000022435	Root MSE	=	7.6e-05

cpta	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lcpta_1	1.192335	.0673783	17.70	0.000	1.059514	1.325156
lcpta_2	-.1281352	.1057009	-1.21	0.227	-.3365003	.08023
lcpta_3	.1443307	.105489	1.37	0.173	-.0636167	.3522782
lcpta_4	-.2074643	.06753	-3.07	0.002	-.3405841	-.0743444
_cons	.0000244	.0000192	1.27	0.207	-.0000136	.0000623

```
66 .
67 . * Test the hypothesis that the beta_2 = beta_3 = beta_4 = 0
68 . test lcpta_2=lcpta_3=lcpta_4=0

 ( 1)  lcpta_2 - lcpta_3 = 0
 ( 2)  lcpta_2 - lcpta_4 = 0
 ( 3)  lcpta_2 = 0

       F(  3,   211) =    8.41
            Prob > F =    0.0000

69 .
70 . *pause Part A completed
71 .
72 . *** Part B
73 . * Set up the LHS variable to match the C_t - C_{t-1} in the homework write up
74 . * Do the simple regression of C_t - C_{t-1} on first difference per capita income
75 . * and a constant
76 . reg fd_cpta fd_ypta
```

Source	SS	df	MS	Number of obs	=	219
				F(1, 217)	=	55.15
Model	2.8226e-07	1	2.8226e-07	Prob > F	=	0.0000
Residual	1.1105e-06	217	5.1177e-09	R-squared	=	0.2027
				Adj R-squared	=	0.1990
Total	1.3928e-06	218	6.3890e-09	Root MSE	=	7.2e-05

fd_cpta	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
fd_ypta	.176198	.0237255	7.43	0.000	.1294361	.22296
_cons	.00006	5.40e-06	11.12	0.000	.0000494	.0000707

```
77 .
78 . *pause Part B completed
79 .
80 . *** Part C
81 . * Take logs of the relevant variables before running any regressions
82 . * Run 2SLS with robust White standard errors as the question requests
83 . * Run the first stage regression, where we try to separate the endogeneity of
84 . * first difference output from residuals from regression of consumption on fd output
85 . reg ln_fd_ypta instrument1 instrument2 instrument3 instrument4
```

Source	SS	df	MS	Number of obs	=	214
				F(4, 209)	=	2.51
Model	.000976969	4	.000244242	Prob > F	=	0.0430
Residual	.020344562	209	.000097342	R-squared	=	0.0458
				Adj R-squared	=	0.0276
Total	.021321531	213	.000100101	Root MSE	=	.00987

ln_fd_ypta	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
instrument1	**-.0185833**	**.1358824**	**-0.14**	**0.891**	**-.2864591**	**.2492925**
instrument2	**.3372076**	**.1367773**	**2.47**	**0.014**	**.0675677**	**.6068475**
instrument3	**.170375**	**.1363403**	**1.25**	**0.213**	**-.0984035**	**.4391535**
instrument4	**-.2218674**	**.1351864**	**-1.64**	**0.102**	**-.4883711**	**.0446363**
_cons	**.0039572**	**.0012099**	**3.27**	**0.001**	**.0015721**	**.0063423**

```
86 . predict x_hat
(option xb assumed; fitted values)
(6 missing values generated)

87 . reg ln_fd_cpta x_hat, robust

Linear regression                               Number of obs     =        214
                                                F(1, 212)         =       4.59
                                                Prob > F          =     0.0332
                                                R-squared         =     0.0323
                                                Root MSE          =     .00513
```

ln_fd_cpta	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
x_hat	**.4369222**	**.2038442**	**2.14**	**0.033**	**.0351009**	**.8387434**
_cons	**.002613**	**.0012046**	**2.17**	**0.031**	**.0002384**	**.0049875**

```
88 .
89 . * Should get the same results as the one-shot Stata command:
90 . ivregress 2sls ln_fd_cpta (ln_fd_ypta=instrument1 instrument2 instrument3 instrument4), robust

Instrumental variables (2SLS) regression          Number of obs   =        214
                                                  Wald chi2(1)    =       4.65
                                                  Prob > chi2     =     0.0310
                                                  R-squared       =     0.0685
                                                  Root MSE        =     .00501
```

ln_fd_cpta	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ln_fd_ypta	**.4369222**	**.2025541**	**2.16**	**0.031**	**.0399234**	**.833921**
_cons	**.002613**	**.0011573**	**2.26**	**0.024**	**.0003448**	**.0048811**

```
Instrumented:  ln_fd_ypta
Instruments:   instrument1 instrument2 instrument3 instrument4

91 .
92 . * The two results have exact same coefficient values, but the top two step that
93 . * I set up for 2 SLS has a larger SE because I have not corrected for the compounded SE
94 . * from the first stage.  Should use 1 step command, because corrects for this which
95 . * I was not told to do, and do not want to do here.
96 .
97 . * In MATLAB, I will try to replicate the larger SE version that hasnt had correction
98 . * for regressor/regression compounding of the SE.
99 . *pause Part C completed
100 .
end of do-file

101 .
102 . log close _all
      name:  Grant Aarons Assignment 4
       log:  C:\Users\gaarons\Git\Notes\Stata\2016F\Metrics\logs\stata_4.smcl
  log type:  smcl
 closed on:  13 Feb 2017, 15:37:54
```