Economics 6511: Advanced Applied Econometrics

Lecture 5: Midterm Revision

CSU, East Bay

February 7, 2018

Question 1

Consider the model of fertility from class that relates the general fertility rate (gfr), the number of children born to every 1,000 women of childbearing age, to the real dollar value of the personal tax exemption (pe). The data cover the years 1913 through 1984.

$$gfr_t = \beta_0 + \beta_1 pe_t + \beta_2 pe_{t-1} + \beta_3 pe_{t-2} + u_t$$

Regression output from the above model and other specifications are included below.

1. Interpret the coefficient on pe_t.
2. What is your estimate of the long-run propensity (LRP) of the personal tax exemption? Explain how you could obtain a 95% confidence interval for this estimate.
3. The first order autocorrelations of gfr and pe are 0.977 and 0.964. What is the significance of this?
4. Write the Stata code for obtaining the first difference of the variable gfr.

Question 1

5. Consider the following model in first differences:

$$\Delta gfr_t = \beta_0 + \beta_1 \Delta pe_t + \beta_2 \Delta pe_{t-1} + \beta_3 \Delta pe_{t-2} + \Delta u_t$$

   In the Stata output, $cgfr = \Delta gfr$ and $cpe = \Delta pe$. How do you interpret the estimate of β_3? What is your estimate of the LRP?

6. Suppose you add two variables to the model in part (e): a dummy variable for World War 2 (*ww2*) and a dummy variable indicating whether the birth control pill is available (*pill*). Interpret the coefficients on these two variables.

7. Test the joint significance of *ww2* and *pill* at the 5 percent level.

8. Suppose you add a linear time trend (t) in addition to *ww2* and *pill* to the equation in (e). What happens to the coefficient on *pill* as compared with that in part (f)?

9. Using the model from part (h), what is your estimate of the LRP? Compare this to your result from part (a). Would you say that the link between fertility and the value of the personal tax exemption is a particularly robust finding?

Question 1

```
. reg gfr pe pe_1 pe_2
```

Source	SS	df	MS
Model	159.461148	3	53.153716
Residual	25832.9717	66	391.408663
Total	25992.4329	69	376.701926

Number of obs	70
F(3, 66)	0.14
Prob > F	0.9383
R-squared	0.0061
Adj R-squared	-0.0390
Root MSE	19.784

gfr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
pe	-.0158445	.140256	-0.11	0.910	-.2958747	.2641856
pe_1	-.0213365	.2152292	-0.10	0.921	-.4510555	.4083826
pe_2	.0539005	.1381132	0.39	0.698	-.2218513	.3296524
_cons	93.15791	4.499654	20.70	0.000	84.17406	102.1418

Question 1

```
. reg cgfr cpe cpe_1 cpe_2
```

Source	SS	df	MS
Model	293.259859	3	97.7532864
Residual	968.199959	65	14.895384
Total	1261.45982	68	18.5508797

Number of obs	= 69
F(3, 65)	= 6.56
Prob > F	= 0.0006
R-squared	= 0.2325
Adj R-squared	= 0.1971
Root MSE	= 3.8595

cgfr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
cpe	-.0362021	.0267737	-1.35	0.181	-.089673	.0172687
cpe_1	-.0139706	.0275539	-0.51	0.614	-.0689997	.0410584
cpe_2	.1099896	.0268797	4.09	0.000	.0563071	.1636721
_cons	-.9636787	.4677599	-2.06	0.043	-1.89786	-.0294976

Question 1

```
. reg cgfr cpe cpe_1 cpe_2 t
```

Source	SS	df	MS
Model	294.856737	4	73.7141842
Residual	966.603082	64	15.1031732
Total	1261.45982	68	18.5508797

Number of obs	69
F(4, 64)	4.88
Prob > F	0.0017
R-squared	0.2337
Adj R-squared	0.1859
Root MSE	3.8863

cgfr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
cpe	-.0348352	.0272856	-1.28	0.206	-.0893445	.019674
cpe_1	-.0131442	.0278616	-0.47	0.639	-.0688042	.0425158
cpe_2	.11109	.0272773	4.07	0.000	.0565974	.1655826
t	.0078781	.0242282	0.33	0.746	-.0405233	.0562796
_cons	-1.267445	1.046219	-1.21	0.230	-3.357507	.822617

Question 1

```
. reg cgfr cpe cpe_1 cpe_2 ww2 pill
```

Source	SS	df	MS
Model	372.894284	5	74.5788568
Residual	888.565534	63	14.1042148
Total	1261.45982	68	18.5508797

Number of obs	= 69
F(5, 63)	= 5.29
Prob > F	= 0.0004
R-squared	= 0.2956
Adj R-squared	= 0.2397
Root MSE	= 3.7556

cgfr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
cpe	-.0751636	.0323566	-2.32	0.023	-.1398232	-.010504
cpe_1	-.0513865	.0331632	-1.55	0.126	-.1176579	.0148848
cpe_2	.0882556	.0279766	3.15	0.002	.0323488	.1441624
ww2	4.839225	2.831973	1.71	0.092	-.8200213	10.49847
pill	-1.676145	1.004766	-1.67	0.100	-3.684009	.3317186
_cons	-.6502546	.5817652	-1.12	0.268	-1.81282	.5123105

Question 1

```
. reg cgfr cpe cpe_1 cpe_2 ww2 pill t
```

Source	SS	df	MS
Model	453.089851	6	75.5149751
Residual	808.369968	62	13.0382253
Total	1261.45982	68	18.5508797

Number of obs	69
F(6, 62)	5.79
Prob > F	0.0001
R-squared	0.3592
Adj R-squared	0.2972
Root MSE	3.6108

cgfr	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
cpe	-.0618201	.0315717	-1.96	0.055	-.124931	.0012908
cpe_1	-.039124	.0322664	-1.21	0.230	-.1036236	.0253757
cpe_2	.0951653	.0270425	3.52	0.001	.0411081	.1492225
ww2	3.250812	2.797162	1.16	0.250	-2.340636	8.84226
pill	-4.888069	1.615706	-3.03	0.004	-8.117819	-1.658319
t	.0944006	.0380635	2.48	0.016	.0183127	.1704884
_cons	-3.141163	1.149618	-2.73	0.008	-5.439216	-.8431097

Question 2

Consider a model of earnings:

$$\log(wage_i) = \beta_0 + \beta_1 educ_i + \beta_2 union_i + \beta_3 black_i + \beta_4 hisp_i + \beta_5 poorhlth_i + u_i$$

where $educ_i$ is in years, *union* is a dummy equal to 1 if i is part of a union, *hisp* is a dummy equal to 1 if person i is hispanic, and *poorhlth* is a dummy equal to 1 if i is in poor health. The data are a cross-section of males from 1980. Estimates of this model and others are provided below.

1. What is the interpretation of $\hat{\beta}_1$?
2. Does the sign of the coefficient on *poorhlth* make sense? Explain.
3. Suppose you collected data on the *same* individuals for 1981 to 1987. Consider an unobserved effects model that allows you test whether the returns to education have changed over time:

$$\begin{aligned}\log(wage_{it}) =& \beta_0 + \delta_1 d81_t + \ldots + \delta_7 d87_t + \beta_1 educ_i \\ &+ \gamma_1 d81_t \cdot educ_i + \ldots + \gamma_7 d87_t \cdot educ_i + \beta_2 union_{it} \\ &+ \beta_3 black_i + \beta_4 hisp_i + \beta_5 poorhlth_{it} + a_i + u_{it}\end{aligned}$$

where $d81_t$ is a dummy variable for the year 1981. What factors are likely to be contained in a_i? Choose one factor and explain how it might bias the coefficient on one of the variables included in the model.

Question 2

5. How might you control for the unobserved effect, a_i?
6. Suppose you estimate the first-differenced equation. Which parameters can you not estimate using first differencing and why?
7. Estimates of the first-differenced equation are provided below (where "c" before a variable denotes the change in the variable). Formally test the null hypothesis (at the 5 percent level) that the return to education has not changed over time.
8. Why is the coefficient on $\Delta union$ in the first-differenced model (probably) much smaller than the coefficient on *union* from the model in part (a)?
9. The same model is estimated using the appendange "cluster(nr)" where *nr* identifies the respondent. What does the "cluster" command allow for? What happens to the statistical significance of $\Delta union$ with clustering?
10. Explain in words how you would test for serial correlation in the error term using estimates from the first-differenced model.

Question 2

```
. reg lwage educ union black hisp poorhlth if year==1980
```

Source	SS	df	MS
Model	12.9889096	5	2.59778192
Residual	156.090184	539	.289592177
Total	169.079093	544	.310807157

Number of obs	545
F(5, 539)	8.97
Prob > F	0.0000
R-squared	0.0768
Adj R-squared	0.0683
Root MSE	.53814

lwage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
educ	.0599163	.0135322	4.43	0.000	.0333339	.0864987
union	.2647581	.0535832	4.94	0.000	.1595006	.3700155
black	-.0800924	.0736824	-1.09	0.278	-.2248322	.0646474
hisp	.0028126	.0659805	0.04	0.966	-.1267977	.132423
poorhlth	.1697316	.1578938	1.07	0.283	-.1404311	.4798943
_cons	.6269718	.1651291	3.80	0.000	.3025963	.9513473

Question 2

```
. reg clwage cd81-cd87 cd81_educ-cd87_educ cunion cpoorhlth, nocons
```

Source	SS	df	MS
Model	20.9921267	16	1.31200792
Residual	747.620874	3799	.196794123
Total	768.613001	3815	.201471298

Number of obs	3815
F(16, 3799)	6.67
Prob > F	0.0000
R-squared	0.0273
Adj R-squared	0.0232
Root MSE	.44361

clwage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
cd81	-.0238769	.1295078	-0.18	0.854	-.2777884	.2300346
cd82	-.0132354	.1831374	-0.07	0.942	-.3722925	.3458217
cd83	.0126088	.2242485	0.06	0.955	-.4270503	.4522678
cd84	.09034	.2589367	0.35	0.727	-.4173282	.5980083
cd85	.0416911	.2895167	0.14	0.886	-.5259319	.6093142
cd86	.0589789	.3171537	0.19	0.852	-.562829	.6807868
cd87	.0919261	.3425436	0.27	0.788	-.5796609	.7635132
cd81_educ	.0121732	.0108868	1.12	0.264	-.0091713	.0335177
cd82_educ	.0162321	.0153951	1.05	0.292	-.0139513	.0464155
cd83_educ	.0181018	.018851	0.96	0.337	-.0188574	.0550609
cd84_educ	.0175557	.0217671	0.81	0.420	-.0251205	.060232
cd85_educ	.0258999	.0243371	1.06	0.287	-.0218152	.073615
cd86_educ	.0296088	.0266602	1.11	0.267	-.0226609	.0818785
cd87_educ	.0323049	.0287952	1.12	0.262	-.0241508	.0887605
cunion	.0406563	.0197451	2.06	0.040	.0019443	.0793683
cpoorhlth	-.0543165	.0429605	-1.26	0.206	-.1385444	.0299114

Question 2

```
. reg clwage cd81-cd87 cd81_educ-cd87_educ cunion cpoorhlth, nocons cluster(nr)

Linear regression                               Number of obs =    3815
                                                F( 16,   544) =   27.04
                                                Prob > F      =  0.0000
                                                R-squared     =  0.0273
                                                Root MSE      =  .44361

                                    (Std. Err. adjusted for 545 clusters in nr)
------------------------------------------------------------------------------
             |               Robust
      clwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        cd81 |  -.0238769   .1438893    -0.17   0.868    -.3065236    .2587697
        cd82 |  -.0132354   .1394727    -0.09   0.924    -.2872064    .2607356
        cd83 |   .0126088   .1539559     0.08   0.935    -.2898121    .3150296
        cd84 |    .09034    .1590508     0.57   0.570    -.2220889    .4027689
        cd85 |   .0416911   .1567712     0.27   0.790    -.2662598    .3496421
        cd86 |   .0589789    .170808     0.35   0.730     -.276545    .3945029
        cd87 |   .0919261    .157112     0.59   0.559    -.2166944    .4005467
   cd81_educ |   .0121732   .0121676     1.00   0.318     -.011728    .0360745
   cd82_educ |   .0162321    .011819     1.37   0.170    -.0069843    .0394485
   cd83_educ |   .0181018   .0130632     1.39   0.166    -.0075587    .0437622
   cd84_educ |   .0175557   .0138506     1.27   0.206    -.0096515    .0447629
   cd85_educ |   .0258999   .0135214     1.92   0.056    -.0006607    .0524605
   cd86_educ |   .0296088   .0147493     2.01   0.045     .0006361    .0585814
   cd87_educ |   .0323049   .0135425     2.39   0.017     .0057029    .0589068
      cunion |   .0406563   .0220915     1.84   0.066    -.0027388    .0840514
    cpoorhlth |  -.0543165    .055394    -0.98   0.327    -.1631287    .0544958
------------------------------------------------------------------------------
```

Question 2

```
. reg clwage cd81-cd87 cunion cpoorhlth, nocons
```

Source	SS	df	MS
Model	20.5619774	9	2.28466416
Residual	748.051023	3806	.196545198
Total	768.613001	3815	.201471298

Number of obs	3815
F(9, 3806)	11.62
Prob > F	0.0000
R-squared	0.0268
Adj R-squared	0.0245
Root MSE	.44333

clwage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
cd81	.119368	.0189906	6.29	0.000	.0821354	.1566007
cd82	.1777662	.0268571	6.62	0.000	.1251105	.2304218
cd83	.2256218	.032894	6.86	0.000	.1611304	.2901133
cd84	.2969162	.0379808	7.82	0.000	.2224515	.3713808
cd85	.3464835	.04247	8.16	0.000	.2632174	.4297496
cd86	.4074214	.0465253	8.76	0.000	.3162044	.4986383
cd87	.4720564	.0502457	9.39	0.000	.3735453	.5705675
cunion	.0413709	.0197167	2.10	0.036	.0027145	.0800273
cpoorhlth	-.0534563	.0429039	-1.25	0.213	-.1375731	.0306606

Next week

- Valentine's Day. . . and a midterm!