Suggested Answers, Problem Set 5
ECON 30331

Bill Evans
Spring 2010

1. The program that generates these results is called q1_ps5.do. Below are a few tables that summarize the results for this problem. Please note that the variables v2 and v3 that are used to construct the new variables are produced from draws to a random number generator. Each time you run a program, the computer will generate a different sequence of random numbers so your results will differ slightly. Given the sample sizes, the mean and standard deviation of v2 should close to 0 and 1 respectively, while the same values for v3 should be 0 and 2.

Means, Standard Deviations, and Variances of Key Variables

Variable	Mean	Std. deviation	Variance
years_educ	13.16	2.80	7.84
v2	-0.0011	0.997	0.994
v3	-0.0006	1.996	3.985
educ2	13.16	2.97	8.82
educ3	13.16	3.43	11.76
ln(weekly_earn)	6.067	0.513	0.263
y2	6.066	1.118	1.250
y2	6.068	2.057	4.231

OLS Estimates

Problem	Dependent Variable	Independent Variable	Parameter on independent	Std error on independent
1a	Ln(weekly_earn)	years_educ	0.0741	0.0012
1c	Ln(weekly_earn)	educ2	0.0653	0.0011
1d	Ln(weekly_earn)	educ3	0.0483	0.0010
1e	y2	years_educ	0.0745	0.0028
1e	y3	years_educ	0.0763	0.0052

1a) When education does not have measurement error, the coefficient on that variable is 0.0741 indicating that each additional year of education increases earnings by 7.4 percent.

1b) The random variable z2 has a mean of roughly zero and a variance of approximately 1. Therefore, the new constructed variable educ2 has a mean of 13.16 which is exactly the mean of years_educ, but now the variance of years_educ has increased by approximately 1, from 7.84 to 8.82.

1c) When ln(weekly earn) is regressed on educ2, notice that the coefficient on the education variable falls to 0.0653. Notice also that the ratio of this estimate to the one without measurement error is simply 0.0653/0.0741=0.881. Is this to be expected? Yes. Recall two facts. First, in large samples, when there is random measurement error in x, the coefficient on x falls by the size of the reliability ratio. Notice that the variance of educ2 is simply the variance of years_educ plus the variance of v2. Therefore, the reliability ratio is $\theta = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_v^2}$ =7.84/8.82 = 0.888. the reliability ratio suggests that the coefficient on educ2 should be about 12 percent lower and it is roughly 12 percent lower.

1d) Educ3=years_educ+v3 and notice that the variance for educ3 is about 4 larger than the variance of years_educ. Therefore, the reliability ratio in this context is $\theta = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_v^2}$ = 7.84/11.76=0.666. This is the devil's number but it also suggests that using educ3 instead of years_educ should reduce the coefficient on education by 33 percent. Notice that the ratio of the new to the original estimate is 0.0483/0.0741=0.651 or about 35 percent lower.

1e) On problems sets 2 and 3, we demonstrated that with random measurement error in the dependent variable, the estimate for β_1 is still unbiased (problem set 2) but the standard error should rise considerably. In this problem, notice that the two variables with measurement error (y2 and y3) have essentially the same mean as ln_weekly_earn but the variance increases by 1 and 4 respectively over the initial value. Therefore, when we replace ln_weekly_earn with y2 and y3, we see little change in the coefficient estimate for β_1 but a large change in the estimate for the standard error of β_1.

2. a) Notice also that if $y_t = \rho y_{t-1} + \varepsilon_t$ and $y_{t-1} = \rho y_{t-2} + e_{t-1}$ then

$$y_t = \rho y_{t-1} + \varepsilon_t = \rho(\rho y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t = \rho^2 y_{t-2} + \rho\varepsilon_{t-1} + \varepsilon_t$$

And since $y_{t-2} = \rho y_{t-3} + \varepsilon_{t-2}$ then you can also show that

$$y_t = \rho^3 y_{t-3} + \rho^2\varepsilon_{t-2} + \rho\varepsilon_{t-1} + \varepsilon_t$$

Doing this multiple times, we then generate that identify that

$$y_t = \rho^{n-1} y_1 + \varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + \rho^3\varepsilon_{t-3} + ... = \rho^{n-1} y_1 + \sum_{j=1}^{n-2} \rho^{j-1}\varepsilon_{t-j+1}$$

b) Taking expectations of the equation above, note that

$$E[y_t] = E[\rho^{t-1} y_1] + E[\varepsilon_t] + E[\rho\varepsilon_{t-1}] + E[\rho^2\varepsilon_{t-2}] + E[\rho^3\varepsilon_{t-3}] + ...$$

$$E[y_t] = \rho^{t-1} y_1$$

Because $E[\rho^{j-1}\varepsilon_{t-j+1}] = \rho^{j-1}E[\varepsilon_{t-j+1}] = 0$ and $\rho^{t-1} y_1$ is a constant

Note that when 10 observations, the expected value of the 10th observation still has some information about the 1st observation. In particular, $E[y_{10}] = \rho^9 y_1$. How much information still persists is a function of the value of ρ. If ρ=0.25, then 0.25^9=3.81e-6 and virtually none Y1 persists until period 10. However if ρ=0.5, ρ^9 =0.002 and if ρ=0.9, ρ^9 =0.39.

c) If ρ=1, then note that $E[y_t] = \rho^{t-1} y_1 = y_1$ and regardless of the period in the future, the best prediction of Y will always be y_1.

3. a) A regression of the form $\hat{\varepsilon}_t = \rho\hat{\varepsilon}_{t-1} + v_t$ will generate the estimate $\hat{\rho} = \frac{\sum_{t=2}^{T}\hat{\varepsilon}_{t-1}\hat{\varepsilon}_t}{\sum_{t=2}^{T}\hat{\varepsilon}_{t-1}^2}$.

b) There is no constant in the model because we assume $E[\varepsilon_t] = 0$ and we know that within our sample $\bar{\hat{\varepsilon}} = 0$

4. Start with the Durbin-Watson Statistic and complete the square in the numerator:

$$\hat{d} = \frac{\sum_{t=2}^{n}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2} = \frac{\sum_{t=2}^{n}(\hat{\varepsilon}_t^2 - 2\hat{\varepsilon}_t\hat{\varepsilon}_{t-1} + \hat{\varepsilon}_{t-1}^2)}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2} = \frac{\sum_{t=2}^{n}\hat{\varepsilon}_t^2}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2} - 2\frac{\sum_{t=2}^{n}\hat{\varepsilon}_t\hat{\varepsilon}_{t-1}}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2} + \frac{\sum_{t=2}^{n}\hat{\varepsilon}_{t-1}^2}{\sum_{t=1}^{n}\hat{\varepsilon}_t^2} = a - 2b + c$$

Notice that ratios (a) and (c) are both approximately equal to 1 and ratio b is approximately equal to $\hat{\rho}$, so $\hat{d} \approx 2(1-\hat{\rho})$.

5. We hope to estimate the model $y_t = \beta_0 + x_t\beta_2 + \varepsilon_t$ and there is third order autocorrelation, $\varepsilon_t = \rho_1\varepsilon_{t-1} + \rho_2\varepsilon_{t-2} + \rho_3\varepsilon_{t-3} + v_t$

Step 1: estimate $y_t = \beta_0 + x_t\beta_2 + \varepsilon_t$ by OLS and generate an estimate of $\hat{\varepsilon}_t = y_t - \hat{\beta}_0 - x_t\hat{\beta}_1$

Step 2: Lag $\hat{\varepsilon}_t$ three times then run the regression $\hat{\varepsilon}_t = \rho_1\hat{\varepsilon}_{t-1} + \rho_2\hat{\varepsilon}_{t-2} + \rho_3\hat{\varepsilon}_{t-3} + v_t$ to obtain estimates for $\hat{\rho}_1, \hat{\rho}_2, and\ \hat{\rho}_3$

Step 3: rho-difference the data three times

$$y_t^* = y_t - \hat{\rho}_1 y_{t-1} - \hat{\rho}_2 y_{t-2} - \hat{\rho}_3 y_{t-3}$$
$$x_t^* = x_t - \hat{\rho}_1 x_{t-1} - \hat{\rho}_2 x_{t-2} - \hat{\rho}_3 x_{t-3}$$
$$v_t = \varepsilon_t - \hat{\rho}_1\varepsilon_{t-1} - \hat{\rho}_2\varepsilon_{t-2} - \hat{\rho}_3\varepsilon_{t-3}$$

Step 4: run the synthetic regression

$$y_t^* = \beta_0^* + x_t^*\beta_2 + v_t$$

Notice that by rho-differencing the data three times, we generate the error term v_t in step 4.

6. The program and results are in the files okun.do and okun.log respectively.

a) Regressing unemp on its lag generates a lag coefficient of 0.734 on lagged unemployment and the 95% confidence interval is only 0.558,0.910) so we can reject the null the coefficient on the lag is 1 it does not look like unemployment is a stationary series. In contrast, the regression of lgdpr on its lag produces an estimate of 0.992 on the lagged lgdpr and the confidence interval contains 1 so we cannot reject the null the coefficient on

```
      Source |       SS       df       MS              Number of obs =      59
-------------+------------------------------           F(  1,    57) =   69.74
       Model |  69.9782083     1  69.9782083           Prob > F      =  0.0000
    Residual |  57.1957759    57  1.00343466           R-squared     =  0.5503
-------------+------------------------------           Adj R-squared =  0.5424
       Total |  127.173984    58   2.1926549           Root MSE      =  1.0017

------------------------------------------------------------------------------
       unemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      unemp1 |   .7343877   .0879404     8.35   0.000     .5582901    .9104854
       _cons |   1.507979   .5105941     2.95   0.005     .4855324    2.530427
------------------------------------------------------------------------------

. test unemp1=1

 ( 1)  unemp1 = 1

       F(  1,    57) =    9.12
            Prob > F =    0.0038

. reg lgdpr lgdpr1

      Source |       SS       df       MS              Number of obs =      59
-------------+------------------------------           F(  1,    57) =29288.15
       Model |  18.3867782     1  18.3867782           Prob > F      =  0.0000
    Residual |  .035783977    57  .000627789           R-squared     =  0.9981
-------------+------------------------------           Adj R-squared =  0.9980
       Total |  18.4225622    58  .317630383           Root MSE      =  .02506

------------------------------------------------------------------------------
       lgdpr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lgdpr1 |   .9920535   .0057968   171.14   0.000     .9804456    1.003661
       _cons |   .1000495   .0490175     2.04   0.046     .0018935    .1982054
------------------------------------------------------------------------------

. test lgdpr1=1

 ( 1)  lgdpr1 = 1

       F(  1,    57) =    1.88
            Prob > F =    0.1758
```

b) The Durbin Watson statistic is estimated to be 0.038 which is incredibly low. In this case k=2 and N=60 (roughly) so the lower and upper values of the DW test statistic are 1.51 and 1.662, respectively, and since $\hat{d}$ <1.514, we can easily reject the null of no autocorrelation.

```
. estat dwatson

Durbin-Watson d-statistic(  2,    60) =  .0383833
```

c) Okun's law looks pretty good after all these years. A regression of dlgdpr on dunemp peoduce a coefficient on dunemp of -1.98 and since the confidence interval is (-2.3,-1.6), we cannot reject the null that the coefficient equals -2 as originally proposed by Okun almost 46 years ago. We cannot reject the null the constant equals 3 and we cannot reject the joint null hypothesis H_o: β_1 =-2 and β_0=3. The

```
. reg dlgdpr dunemp

      Source |       SS       df       MS              Number of obs =      59
-------------+------------------------------           F(  1,    57) =  136.26
       Model |  260.614306     1  260.614306           Prob > F      =  0.0000
    Residual |  109.022975    57  1.91268377           R-squared     =  0.7051
-------------+------------------------------           Adj R-squared =  0.6999
       Total |  369.637281    58  6.37305657           Root MSE      =   1.383

------------------------------------------------------------------------------
      dlgdpr |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      dunemp |   -1.98189   .1697861   -11.67   0.000    -2.321881   -1.641899
       _cons |    3.33389    .180074    18.51   0.000     2.973298    3.694482
------------------------------------------------------------------------------

. test dunemp=-2

 ( 1)  dunemp = -2

       F(  1,    57) =    0.01
            Prob > F =    0.9154

. test _cons=3

 ( 1)  _cons = 3

       F(  1,    57) =    3.44
            Prob > F =    0.0689

. test (dunemp=-2)(_cons=3)

 ( 1)  dunemp = -2
 ( 2)  _cons = 3

       F(  2,    57) =    1.73
            Prob > F =    0.1868


. estat dwatson

Durbin-Watson d-statistic(  2,    59) =  2.003222
```

d) The Durbin-Watson statistic is 2 and with k=2 and n=59, the lower and upper critical values of the DW test statistic are 1.514 and 1.652, respectively, so we cannot reject the null that ρ=0 (no autocorrelation) in the data.

7.

a) N=95, k=6, $\hat{d}=1.85$, lower=1.535, upper=1.805, since $\hat{d}$ >upper, we cannot reject that the null ρ=0 (no autocorrelation)
b) N=33, k=15, $\hat{d}=0.51$, lower=0.488, upper=2.796, since lower< $\hat{d}$ <upper, the test is inconclusive
c) N=55, k=2, $\hat{d}=1.55$, lower=1.49 upper=1.641, since lower< $\hat{d}$ <upper, the test is inconclusive