ECON 6511: Advanced Applied Econometrics

Homework 2 **January 24, 2018**

Submitted By:
Surabhi Asati
va7892

---

1. **(Based on Wooldridge, Chapter 12, Problem 2) Let $\{e_t : t = -1, 0, 1, \ldots\}$ be a sequence of independent, identically distributed random variables with mean zero and variance one. Define a stochastic process by $x_t = e_t - (1/2)\, e_{t-1} + (1/2)\, e_{t-2}$, $t = 1, 2, \ldots$.**

(a) Find $E(x_t)$ and $Var(x_t)$. Do either of these depend on t?

$E(x_t) = E(e_t) - (1/2)\, E(e_{t-1}) + (1/2)\, E(e_{t-2})$
$E(x_t) = 0 - (1/2) * 0 + (1/2) * 0$
$E(x_t) = 0$

As e_t is a sequence of independent stochastic variables, they are uncorrelated

$Var(x_t) = Var(e_t) + (1/4)\, Var(e_{t-1}) + (1/4)\, Var(e_{t-2})$

Also, $Var(e_t) = 1$ for all t. Therefore,

$Var(x_t) = 1 + (1/4) + (1/4) = 3/2$

(b) Show that Corr(x_t,x_{t+1}) = - ½ and Corr(x_t,x_{t+2}) =1/3. Note: For two random variables a and b, Corr(a, b) = $^{Cov(a,b)}$ / σa σb

Cov (x_t,x_{t+1}) = E(x_t,x_{t+1})

= E [(et – (1/2) et-1 + (1/2) et-2) (et+1 – (1/2) et + (1/2) et-1)]

= E(e_t * e_{t+1}) – (1/2) E (e_t^2) + (1/2) E(et*et-1) – (1/2) E(e_{t-1}*et+1) + (1/4(E(et-1*et) – (1/4) E (e_{t-1}^2) + (1/2) E (et-2*e_{t+1}) – (1/4) E(e_{t-2}*e_t) +(1/4) E(e_{t-2}*e_{t-1})

= – (1/2) E(e_t^2) – (1/4) E(e_t^2)

= – (1/2) * 1 – (1/4) * 1

= –3/4

Corr (a, b) = Cov(a,b) / (σa * σb)

Corr (x_t,x_{t+1}) = – (3/4) / (3/2)

= –1/2

E(e_t^2) = 1⁄2

Corr (x_t,x_{t+2}) = (1/2) / (3/2) = 1/3

(c) What is Corr (x_t, x_{t+h}) for h > 2?

Corr (x_t, x_{t+h}) for h > 2 = 0

(d) Is {x_t} covariance stationary? Is it weakly dependent?

x_t is weakly dependent because Correlation (x_t,x_{t+h}) < 1 i.e. observations more than two periods apart are uncorrelated

2. (Based on Wooldridge, Chapter 12, Problem 5) For the U.S. economy, let gprice denote the monthly growth in the overall price level and let gwage be the monthly growth in hourly wages. These are both obtained as differences of logarithms: gprice = Δlog(price) and gwage = Δlog(wage).

(a) Using the monthly data in WAGEPRC.dta, estimate a distributed lag model: gprice = $\alpha + \beta_1 gwage_t + \beta_2 gwage_{t-1} + \beta_3 gwage_{t-2} + \beta_4 gwage_{t-3} + \beta_5 gwage_{t-4} + \beta_6 gwage_{t-5} + \beta_7 gwage_{t-6} + \beta_8 gwage_{t-7} + \beta_9 gwage_{t-8} + \beta_{10} gwage_{t-9} + \beta_{11} gwage_{t-10} + \beta_{12} gwage_{t-11} + \beta_{13} gwage_{t-12} + u_t$ **and sketch the estimated lag distribution. At what lag is the effect of gwage on gprice largest? Which lag has the smallest coefficient?**

Using stata:

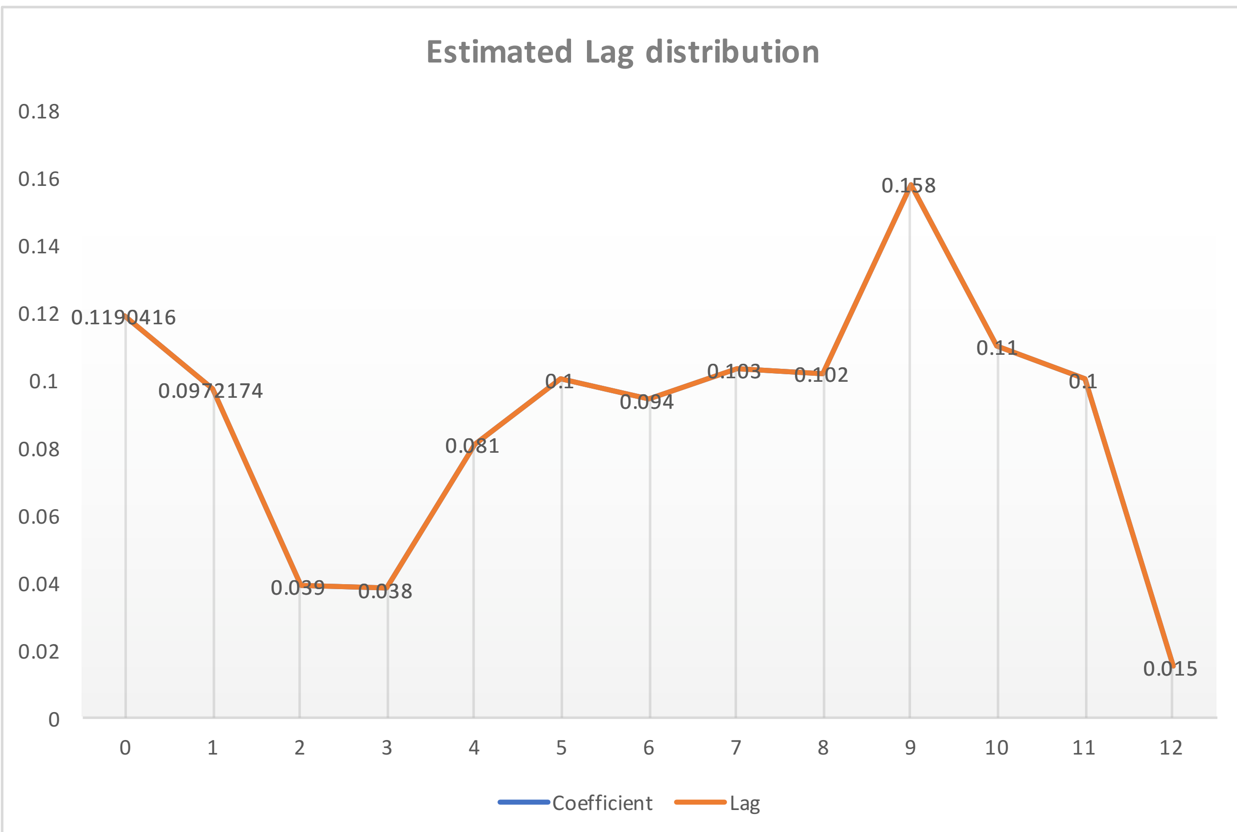
```
. use "/Users/surbhiasati/Desktop/Econometrics/Homework2/WAGEPRC.DTA"

. reg gprice gwage gwage_1 gwage_2 gwage_3 gwage_4 gwage_5 gwage_6 gwage_7 gwage_8 gwage_
> 9 gwage_10 gwage_11 gwage_12
```

Source	SS	df	MS			
				Number of obs	=	273
				F(13, 259)	=	9.25
Model	.000981458	13	.000075497	Prob > F	=	0.0000
Residual	.002113658	259	8.1608e-06	R-squared	=	0.3171
				Adj R-squared	=	0.2828
Total	.003095116	272	.000011379	Root MSE	=	.00286

gprice	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
gwage	.1190416	.0517725	2.30	0.022	.0170929	.2209903
gwage_1	.0972174	.0390409	2.49	0.013	.0203393	.1740954
gwage_2	.0399518	.0390717	1.02	0.307	-.0369869	.1168905
gwage_3	.0382652	.0391513	0.98	0.329	-.0388301	.1153605
gwage_4	.0813362	.0393483	2.07	0.040	.0038528	.1588195
gwage_5	.106852	.0391937	2.73	0.007	.0296731	.1840308
gwage_6	.0949731	.0392186	2.42	0.016	.0177451	.1722011
gwage_7	.1037922	.0393788	2.64	0.009	.0262488	.1813355
gwage_8	.1025629	.0394884	2.60	0.010	.0248037	.180322
gwage_9	.1585079	.0393341	4.03	0.000	.0810526	.2359632
gwage_10	.1104412	.0392229	2.82	0.005	.0332049	.1876776
gwage_11	.1033206	.0394388	2.62	0.009	.0256591	.180982
gwage_12	.0156575	.0518343	0.30	0.763	-.0864128	.1177278
_cons	-.0009296	.0005662	-1.64	0.102	-.0020445	.0001853

Model:

gprice = - .00 + .11 $gwage_t$ + .09 $gwage_{t-1}$ + .04 $gwage_{t-2}$ + .04 $gwage_{t-3}$ + .08 $gwage_{t-4}$ + .10 $gwage_{t-5}$ + .09 $gwage_{t-6}$ + .10 $gwage_{t-7}$ + .10 $gwage_{t-8}$ + .15 $gwage_{t-9}$ + .11 $gwage_{t-10}$ + .10 $gwage_{t-11}$ + .01 $gwage_{t-12}$



The largest effect is at the ninth lag 0.158. The smallest effect is at the twelfth lag 0.015.

(b) For which lags are the t statistics less than two?

The t statistics is less than two at the second ($t = 1.02$), third ($t = 0.98$) and twelfth ($t = 0.30$) lag.

(c) What is the estimated long-run propensity? Is it much different than one? Explain what the LRP tells us in this example.

The estimated long-run propensity is the sum of the lag coefficients from zero through twelve. In a distributed lag model, the eventual change in the dependent variable given a permanent, one-unit increase in the independent variable.

LRP = 1.172

As, LRP > 1, is a cumulative effect of wage on price. This difference indicates error.

(d) What regression would you run to obtain the standard error of the LRP directly?

Assuming LRP coefficients = q1, q2, q3, …., q12

LRP = q0, q1, q2, q3, …., q12

q0 = LRP – q1 -q2 -q3- …. -q12

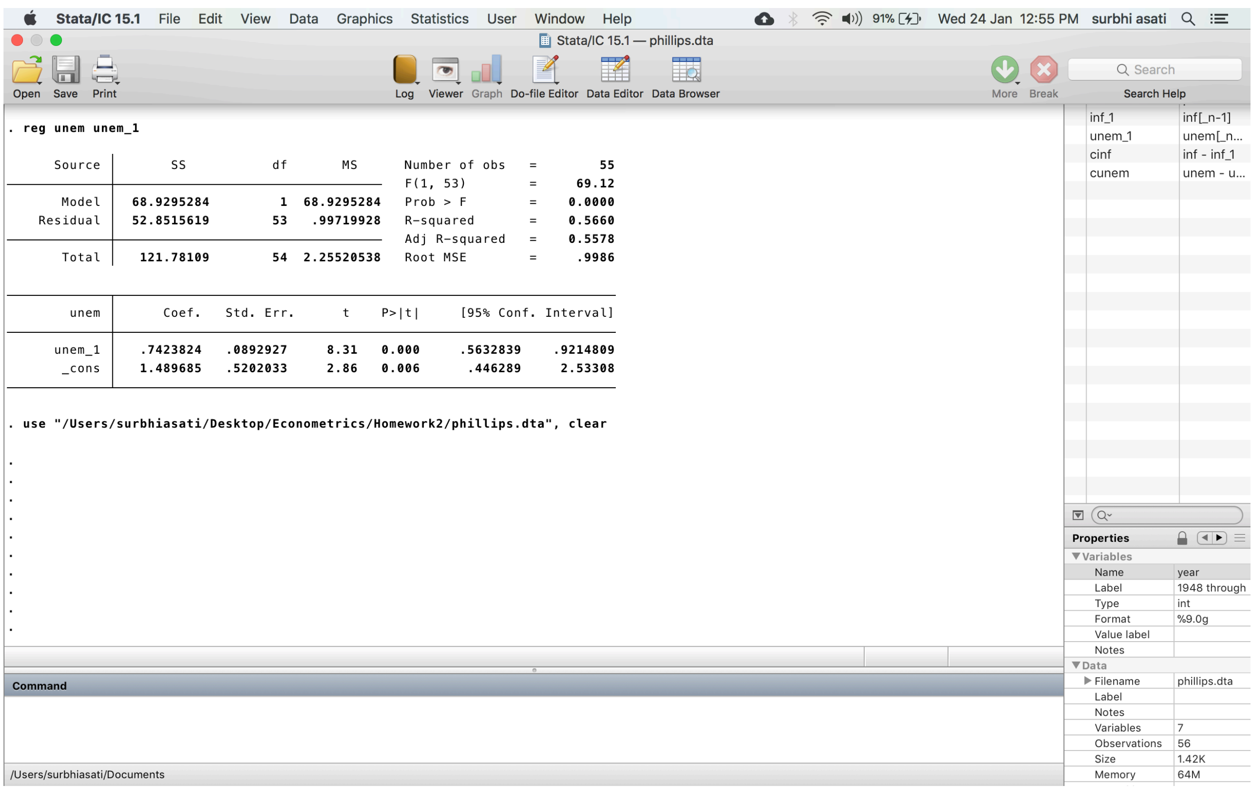
Plug into the estimated model:

$gprice_t = \alpha + (LRP - q1 - q2 - q3 - \ldots - q12)\, gwage_t + q_1 gwage_{t-1} + q_2 gwage_{t-2} + q_3 gwage_{t-3} + q_4 gwage_{t-4} + q_5 gwage_{t-5} + q_6 gwage_{t-6} + q_7 gwage_{t-7} + q_8 gwage_{t-8} + q_9 gwage_{t-9} + q_{10} gwage_{t-10} + q_{11} gwage_{t-11} + q_{12} gwage_{t-12} + u_t$

3. Use the data in PHILLIPS.dta for this exercise.

(a) Estimate an AR(1) model for the unemployment rate. Use this equation to predict the unemployment rate for 2004. Compare this with the actual unemployment rate for 2004. (You can find this information in a recent Economic Report of the President.)

Using stata:



The estimated AR model is:

$$\text{Unem}_t = 1.489 + 0.742\ \text{Unem}_{t-1}$$

The unemployment rate in 2003 was 6. The predicted unemployment rate for 2004 is

$$1.49 + 0.742(6) = 5.94$$

According to the Economic Report of the President, the unemployment rate was 5.5. Therefore, the equation over predicts the 2004 unemployment rate by 0.44 percentage.

(b) Add a lag of inflation to the AR(1) model from part (a). Is inf_{t-1} statistically significant?

```
. reg unem unem_1 inf_1

      Source |       SS           df       MS      Number of obs   =        55
-------------+----------------------------------   F(2, 52)        =     59.60
       Model |  84.7924146         2  42.3962073   Prob > F        =    0.0000
    Residual |  36.9886756        52  .711320685   R-squared       =    0.6963
-------------+----------------------------------   Adj R-squared   =    0.6846
       Total |   121.78109        54  2.25520538   Root MSE        =     .8434

------------------------------------------------------------------------------
        unem |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      unem_1 |   .6494787    .0779388     8.33   0.000     .4930831    .8058743
       inf_1 |   .1830086    .0387537     4.72   0.000     .1052436    .2607735
       _cons |   1.296416    .4412563     2.94   0.005     .4109709    2.181862
------------------------------------------------------------------------------
```

Adding inf_1 to the model we get:

$$Unem_t = 1.296 + 0.649\ Unem_{t-1} + 0.183\ inf_{t-1}$$

With p value = 0 (less than 0.05) and a t- statistic of 4.72, Lagged inflation is statistically significant.

(c) Use the equation from part (b) to predict the unemployment rate for 2004. Is the result better or worse than in the model from part (a)?

Equation in part (b):

$$Unem_t = 1.296 + 0.649\ Unem_{t-1} + 0.183\ inf_{t-1}$$

Inflation rate for 2003 = 2.3

Prediction of unemployment in 2003:

$Unem_t = 1.296 + 0.649(6) + 0.183(2.3) = 5.6$

This is very close to the actual rate of 5.5, and it is certainly better than the predication from part (a) that is 5.94

4. (Wooldridge, Chapter 12, C2)

(a) Using the data in WAGEPRC.dta, estimate the distributed lag model from Problem 2 above. Regress $\hat{u}_t$ on $\hat{u}_{t-1}$ to test for AR(1) serial correlation.

Estimated distributed lag model:

gprice = - .00 + .11 $gwage_t$ + .09 $gwage_{t-1}$ + .04 $gwage_{t-2}$ + .04 $gwage_{t-3}$ + .08 $gwage_{t-4}$ + .10 $gwage_{t-5}$ + .09 $gwage_{t-6}$ + .10 $gwage_{t-7}$ + .10 $gwage_{t-8}$ + .15 $gwage_{t-9}$ + .11 $gwage_{t-10}$ + .10 $gwage_{t-11}$ + .01 $gwage_{t-12}$

Regression $\hat{u}_t$ on $\hat{u}_{t-1}$ gives t = 9.60, AR(1) correlation is positive

(b) Reestimate the model using iterated Cochrane-Orcutt (CO) estimation. What is your new estimate of the long-run propensity?

The new estimate of the long-run propensity is = 1.110

(c) Using iterated CO, find the standard error for the LRP. Determine whether the estimated LRP is statistically different from one at the 5% level.

$gprice_t$ = 0 + 0 $gwage_t$ + 1 ($gwage_{t-1}$ – $gwage_t$) + 2 ($gwage_{t-2}$ – $gwage_t$) + + 12 ($gwage_{t-12}$ – $gwage_t$) + u_t

Estimating this equation by CO gives LRP =1.110

The t statistic for testing is (1.110 – 1)/.191 = 0.57. This is greater than 5%. Therefore, LRP is not significant.