EE425: Homework 5

1. Score = 30

   a)

Dependent Variable: SALARY
Method: Least Squares
Date: 12/11/11 Time: 11:17
Sample: 1 11
Included observations: 11

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6472.291	1770.527	3.655574	0.0053
AGE	126.5793	35.65693	3.549922	0.0062

R-squared	0.583371	Mean dependent var	12427.27
Adjusted R-squared	0.537079	S.D. dependent var	2760.830
S.E. of regression	1878.422	Akaike info criterion	18.07722
Sum squared resid	31756231	Schwarz criterion	18.14956
Log likelihood	-97.42470	F-statistic	12.60195
Durbin-Watson stat	0.516743	Prob(F-statistic)	0.006217

All of the parameters are significant at 1% level of significant since p-values are below 0.01. In addition, the estimated coefficient on age has the expected positive sign because we would expect the one with more experience to have higher wage. Since age may be used as a proxy for experience, the constant term can be interpreted as the average salary for those with no experience.

Looking at the R^2, the model can explain the variation of salary by 58.3% and the F-statistic is high high and significant. However, the model has low DW. So, there may be an autocorrelation problem.

b) By assuming that the variance of the disturbance term is proportional to age square, then $\sigma_i^2 = age_i^2$. Therefore, $weight = \dfrac{1}{\sigma_i} = \dfrac{1}{age_i}$

Dependent Variable: SALARY
Method: Least Squares
Date: 12/11/11   Time: 22:46
Sample: 1 11
Included observations: 11
Weighting series: 1/(AGE)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4231.240	1096.699	3.858160	0.0039
AGE	177.0674	28.66570	6.176977	0.0002

Weighted Statistics

R-squared	0.809140	Mean dependent var	11513.14
Adjusted R-squared	0.787934	S.D. dependent var	2216.611
S.E. of regression	1434.245	Akaike info criterion	17.53763
Sum squared resid	18513525	Schwarz criterion	17.60997
Log likelihood	-94.45697	F-statistic	38.15505
Durbin-Watson stat	0.440189	Prob(F-statistic)	0.000163

By using the WLS, the coefficients still have the same signs, and all are significant. In addition the R^2 and F-statistic increase significantly. R^2 increases from 58% to 80% where F increases from 12 to 38. This may imply that the significance of the model improves tremendously after taking into account the heteroscedasticity problem.

c) By assuming that the variance of the disturbance term is proportional to age, then $\sigma_i^2 = age_i$. Therefore, $weight = \frac{1}{\sigma_i} = \frac{1}{\sqrt{age_i}}$

Dependent Variable: SALARY
Method: Least Squares
Date: 12/11/11 Time: 22:48
Sample: 1 11
Included observations: 11
Weighting series: AGE^(-0.5)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5163.313	1404.904	3.675207	0.0051
AGE	154.4030	31.93997	4.834162	0.0009

Weighted Statistics

R-squared	0.721957	Mean dependent var	11991.38
Adjusted R-squared	0.691064	S.D. dependent var	1463.129
S.E. of regression	1682.934	Akaike info criterion	17.85743
Sum squared resid	25490409	Schwarz criterion	17.92978
Log likelihood	-96.21587	F-statistic	23.36913
Durbin-Watson stat	0.470189	Prob(F-statistic)	0.000929

I would choose the model (b) because all statistics seems to be superior. The estimated coefficients in (b) have higher t-statistics. In addition, R-square, adjusted-R^2 and F-statistic in model (b) are significant and higher than model (c).

2. Total score = 20

a) It's simple to interpret the model (1). When the number of employees increases by 1, the average wage would increases by 0.09. The model also suggests that if firms have no employees, firms have to pay average wage of 7.5 which is not meaningful.

For model (2), we don't have to interpret it directly but we can transform model (2) to make it similar to (1) by multiplying through by N_t. Then, the interpretation will be pretty much the same as above.

b) The author worried about the heteroskedasticity because the transformation was quite similar to the weighted least squares or the GLS method using $W_i = \frac{1}{N_i}$.

c) Yes. By multiplying equation (2) through by N_t , then $W_i = 7.81 + 0.008N_i$ which is now comparable to model (1).

d) No. R^2 in model 1 measures the variation of W_i, but R^2 in model 2 measures the variation of $\frac{W_i}{N_i}$. So, the 2 R^2 's cannot be compared because the dependent variables are not the same.

3. Total score = 30

a)

Dependent Variable: Y
Method: Least Squares
Date: 12/11/11 Time: 12:30
Sample: 1 15
Included observations: 15

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	246.2438	5.847661	42.10979	0.0000
X	15.18036	0.643157	23.60288	0.0000

R-squared	0.977197	Mean dependent var	367.6867
Adjusted R-squared	0.975443	S.D. dependent var	68.67615
S.E. of regression	10.76207	Akaike info criterion	7.713500
Sum squared resid	1505.689	Schwarz criterion	7.807906
Log likelihood	-55.85125	F-statistic	557.0959
Durbin-Watson stat	0.416232	Prob(F-statistic)	0.000000

$$\therefore DW = 0.4162$$

b) According to the estimation output, DW is 0.4162. To detect the autocorrelation, we test the following hypothesis,

$$H_0 : \text{There is no autocorrelation}$$
$$H_a : \text{There is a positive autocorrelation}$$

According to DW-table, when n = 15 and number of explanatory is 1, $d_L = 0.811$ and $d_U = 1.070$. Since computed DW (=0.4162) is less than d_L , we can reject H_0 . So, there is a positive autocorrelation in the model.

c) **i) Durbin's 2-step procedure:**

- Step 1: $Y_t = \beta_1 + \beta_2 X_t + u_t ---(1)$

$$\rho Y_{t-1} = \rho\beta_1 + \rho\beta_2 X_{t-1} + \rho u_{t-1} ---(2)$$

$$(1)-(2); Y_t - \rho Y_{t-1} = \beta_1(1-\rho) + \beta_2(X_t - \rho X_{t-1}) + \varepsilon_t$$

$$\text{or } Y_t = \beta_1(1-\rho) + \beta_2(X_t - \rho X_{t-1}) + \rho Y_{t-1} + \varepsilon_t$$

So, in the first step, we estimate $Y_t = \beta_1 + \beta_2 X_t - \beta_3 X_{t+1} + \rho Y_{t-1} + \varepsilon_t$. Then we keep the value $\hat{\rho}$ to proceed to the second step.

- Step 2: Use $\hat{\rho}$ to transform the data

$$Y_t - \hat{\rho} Y_{t-1} = \beta_1(1-\hat{\rho}) + \beta_2(X_t - \hat{\rho} X_{t-1}) + \varepsilon_t$$

Or we estimate the equation, $Y_t^* = \beta_1^* + \beta_2^* X_t^* + \varepsilon_t$

However, in this particular case, we cannot perform Durbin's 2-step procedure. Since X_t is a time variable, so $X_t - X_{t-1} = 1 \ \forall t = 2,...15$. In other words, we face the perfect-multicollinearity problem. As a result, we cannot obtain ρ from the first step. This is an example of why we need the Cochrane – Orcutt method.

ii) Cochrane – Orcutt

```
Iteration 0:  rho = 0.0000                       Iteration 11:  rho = 0.6684
Iteration 1:  rho = 0.6603                       Iteration 12:  rho = 0.6684
Iteration 2:  rho = 0.6634                       Iteration 13:  rho = 0.6684
Iteration 3:  rho = 0.6653                       Iteration 14:  rho = 0.6684
Iteration 4:  rho = 0.6664                       Iteration 15:  rho = 0.6685
Iteration 5:  rho = 0.6672                       Iteration 16:  rho = 0.6685
Iteration 6:  rho = 0.6676                       Iteration 17:  rho = 0.6685
Iteration 7:  rho = 0.6679                       Iteration 18:  rho = 0.6685
Iteration 8:  rho = 0.6681                       Iteration 19:  rho = 0.6685
Iteration 9:  rho = 0.6682                       Iteration 20:  rho = 0.6685
Iteration 10:  rho = 0.6683
Cochrane-Orcutt AR(1) regression -- iterated estimates
      Source |       SS       df       MS              Number of obs =      14
-------------+------------------------------           F(  1,    12) =  290.12
       Model |  7538.52178     1  7538.52178           Prob > F      =  0.0000
    Residual |  311.811971    12  25.9843309           R-squared     =  0.9603
-------------+------------------------------           Adj R-squared =  0.9570
       Total |  7850.33375    13  603.871827           Root MSE      =  5.0975

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   17.36284   1.019373    17.03   0.000     15.14181    19.58386
       _cons |   219.4135   11.48059    19.11   0.000     194.3995    244.4276
-------------+----------------------------------------------------------------
         rho |   .6684632
------------------------------------------------------------------------------
Durbin-Watson statistic (original)    0.416232
Durbin-Watson statistic (transformed) 1.904417
```

Therefore, by Cochrane-Orcutt method, ρ =0.6685 with 20 iterations.

4. Total score = 10

a) To see if there is serial correlation, we can look at DW statistic. Given an annual data, we have 16 observations.

$$H_0 : \text{There is no autocorrelation}$$
$$\mathrm{H}_a : \text{There is a positive autocorrelation}$$

- Model 1 with one explanatory variable:

From DW = 0.8252, $d_L = 0.844$ and $d_U = 1.086$, DW is less than d_L. So, we reject H_0 and conclude that there is a serial autocorrelation in model 1.

Or we can use modified-d test where we compare the estimated-DW with d_U only. In this case, we still reject H_0 and conclude that there is a positive serial autocorrelation.

- Model 2 with 2 explanatory variables

From DW = 1.82, $d_L = 0.700$ and $d_U = 1.252$, DW is less more than d_U. So, we fail to reject H_0 and conclude that there is no serial autocorrelation in model 2.

Or we can use modified-d test where we compare the estimated-DW with d_U only. In this case, we still fail to reject H_0 and conclude that there is no serial autocorrelation.

b) Specification error in model(A) may cause problem in the estimated errors. Because by adding the square term in the model, we can solve the autocorrelation problem.

5. Total score = 20

a) From $u_t = \rho u_{t-1} + \varepsilon_t$,

$$u_t = \sum_{r=0}^{\infty} \rho^r \varepsilon_{t-r}$$

$$\operatorname{var}(u_t) = E\left(\sum_{r=0}^{\infty} \rho^r \varepsilon_{t-r}\right)^2 - \sum_{r=0}^{\infty} \rho^r E(\varepsilon_{t-r})$$

$$\operatorname{var}(u_t) = E\left(\sum_{r=0}^{\infty} \rho^r \varepsilon_{t-r}\right)\left(\sum_{r=0}^{\infty} \rho^r \varepsilon_{t-r}\right) = E(\varepsilon_t^2 + \rho^2 \varepsilon_{t-1}^2 + \ldots + \rho \varepsilon_t \varepsilon_{t-1} + \rho^2 \varepsilon_t \varepsilon_{t-2} + \ldots)$$

$$\operatorname{var}(u_t) = E(\varepsilon_t^2 + \rho^2 \varepsilon_{t-1}^2 + \ldots) = \sigma^2 + \rho^2 \sigma^2 + \rho^4 \sigma^2 \ldots$$

$$\operatorname{var}(u_t) = \sigma^2 (1 + \rho^2 + \rho^4 + \ldots) = \frac{\sigma^2}{1-\rho^2}$$

b) $\text{cov}(u_t u_{t-1}) = E(u_t u_{t-1})$

$$\text{cov}(u_t u_{t-1}) = E\left(\sum_{r=0}^{\infty} \rho^r \varepsilon_{t-r}\right)\left(\sum_{r=1}^{\infty} \rho^{r-1} \varepsilon_{t-r}\right)$$

$$\text{cov}(u_t u_{t-1}) = E(\varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + \ldots)(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \rho^2\varepsilon_{t-3} + \ldots)$$

$$\text{cov}(u_t u_{t-1}) = E(\rho\varepsilon_{t-1}^2 + \rho^3\varepsilon_{t-2}^2 + \ldots + \varepsilon_t\varepsilon_{t-1} + \rho\varepsilon_t\varepsilon_{t-2} + \ldots)$$

$$\text{cov}(u_t u_{t-1}) = E(\rho\varepsilon_{t-1}^2 + \rho^3\varepsilon_{t-2}^2 + \ldots)$$

$$\text{cov}(u_t u_{t-1}) = \rho\sigma^2 + \rho^3\sigma^2 + \ldots$$

$$\text{cov}(u_t u_{t-1}) = \frac{\rho\sigma^2}{(1-\rho^2)}$$

$$\text{cov}(u_t u_{t-2}) = \frac{\rho^2\sigma^2}{(1-\rho^2)}$$

$$\therefore \text{cov}(u_t u_{t-r}) = \frac{\rho^r\sigma^2}{(1-\rho^2)}$$

c)

$$\text{var}-\text{cov}(u) = \begin{bmatrix} \text{var}(u_1) & \text{cov}(u_1,u_2) & \ldots & \text{cov}(u_1,u_n) \\ \text{cov}(u_2,u_1) & \text{var}(u_2) & \cdots & \text{cov}(u_2,u_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(u_n,u_1) & \text{cov}(u_n,u_2) & \cdots & \text{var}(u_n) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\sigma^2}{1-\rho^2} & \frac{\rho\sigma^2}{(1-\rho^2)} & \cdots & \frac{\rho^{n-1}\sigma^2}{(1-\rho^2)} \\ \frac{\rho\sigma^2}{(1-\rho^2)} & \frac{\sigma^2}{1-\rho^2} & \cdots & \frac{\rho^{n-2}\sigma^2}{(1-\rho^2)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\rho^{n-1}\sigma^2}{(1-\rho^2)} & \frac{\rho^{n-2}\sigma^2}{(1-\rho^2)} & \cdots & \frac{\sigma^2}{1-\rho^2} \end{bmatrix}$$

d) From $\text{var}(u_t) = \frac{\sigma^2}{1-\rho^2}$, if $\rho = 1$, the var(u) is undefined.

If we do the first difference, then $Y_t - \rho Y_{t-1} = (1-\rho)\beta_1 + \beta_2(X_t - \rho X_{t-1}) + u_t - \rho u_{t-1}$.

When $\rho = 1$, $\Delta Y_t = \beta_2 \Delta X_t + \varepsilon_t$, or we have to run the first difference without intercept instead.