Computer Exercise 0 – Outline Answers

Read the data set in

```
u cex0
```

To summarise the data

```
su unr inf
```

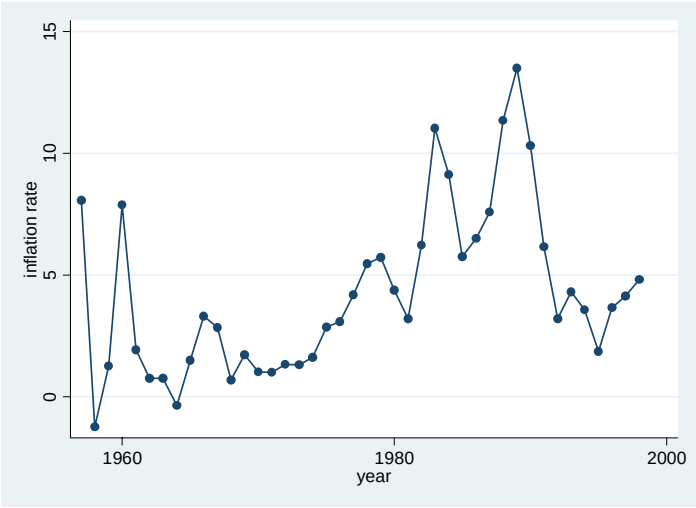
```
Variable |     Obs        Mean    Std. Dev.       Min        Max
---------+--------------------------------------------------------
     unr |      42    5.652381    1.645873         2.9        9.7
     inf |      42    4.223176    3.412677  -1.244818   13.49863
```

so the mean unemployment rate (% of labour force unemployed) is 5.5% and the mean inflation rate (annual %age change in prices)is 4.2% with a minimum of -1.2% and maximum of 13.5%

To graph the data use the commands

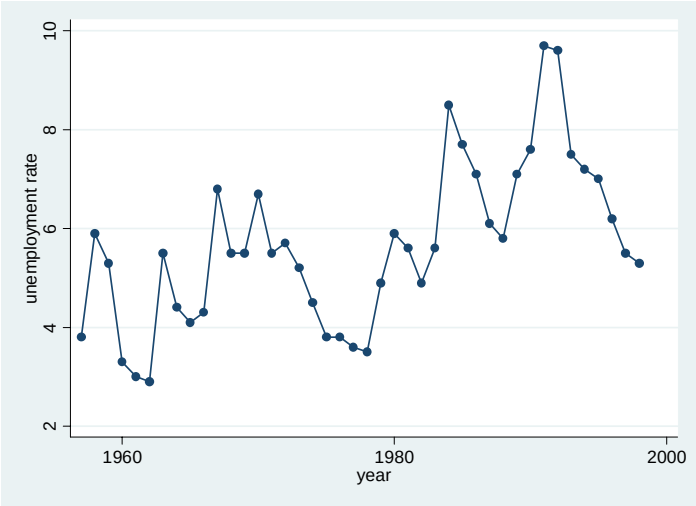
```
twoway connect inf year, xlabel(1960 1980 2000)
```

which gives



and

```
twoway connect unr year, xlabel(1960 1980 2000)
```



```
reg inf unr
```

```
      Source |       SS       df       MS              Number of obs =      42
-------------+------------------------------           F(  1,    40) =    2.78
       Model |   31.058555     1   31.058555           Prob > F      =  0.1031
    Residual |  446.442404    40  11.1610601           R-squared     =  0.0650
-------------+------------------------------           Adj R-squared =  0.0417
       Total |  477.500959    41  11.6463649           Root MSE      =  3.3408
------------------------------------------------------------------------------
         inf |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         unr |   .5288134   .3170039     1.67   0.103    -.1118753    1.169502
       _cons |   1.234121   1.864506     0.66   0.512    -2.534186    5.002428
------------------------------------------------------------------------------
```

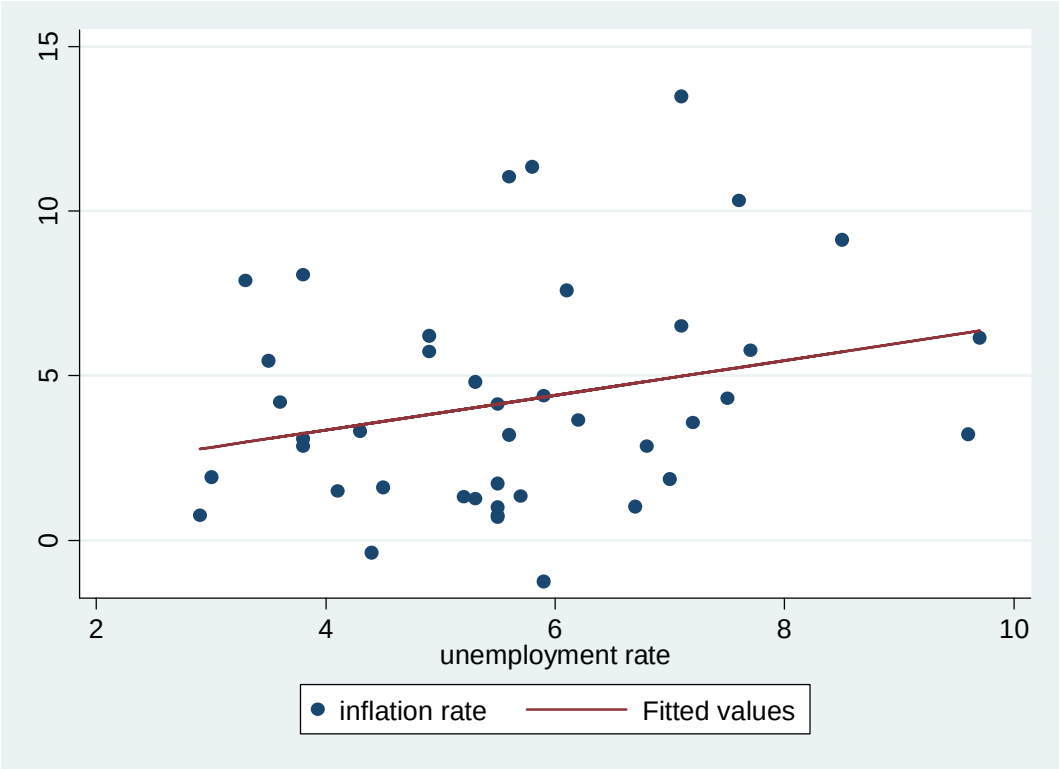
So estimates suggest a 1 **percentage point** rise in the unemployment rate leads to a 0.53 **percentage point** rise in inflation

To graph fitted and actual values, first save predicted values from the regression using the command

predict infhat

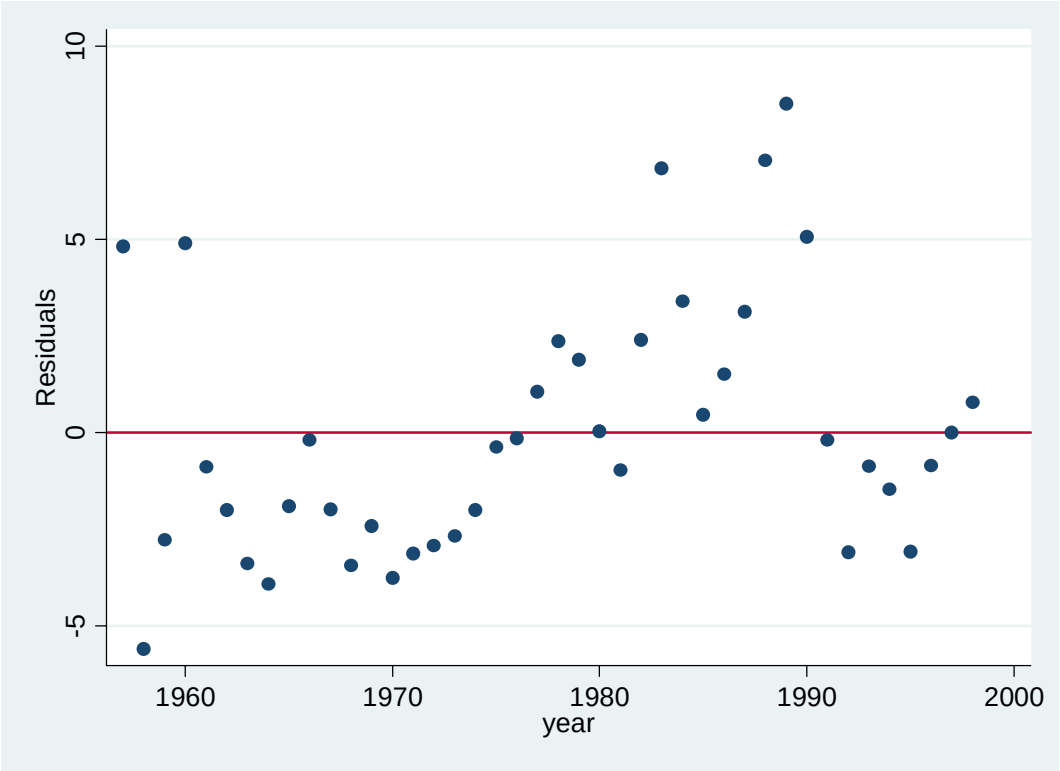
and then graph the results

graph twoway (scatter inf unr) (line yhat unr)



Can see even though this is the line of best fit (OLS) there are still large residual values for most observations. This is one way of finding out that the model does not summarise the variation in the dependent variable very well. The same impression is given by the graph of the residuals from the same regression.

predict reshat, resid
twoway scatter uhat year, yline(0)



The model tends to overpredict inflation in the early years of inflation when inflation is low, ie (y - $\hat{y}$) is negative.

3. To rescale the data use the command

```
replace unr=unr/100
```

and then

```
reg inf unr
```

```
      Source |       SS       df       MS              Number of obs =      42
-------------+------------------------------           F(  1,    40) =    2.78
       Model |  31.0585592     1  31.0585592           Prob > F      =  0.1031
    Residual |    446.4424    40    11.16106           R-squared     =  0.0650
-------------+------------------------------           Adj R-squared =  0.0417
       Total |  477.500959    41  11.6463649           Root MSE      =  3.3408
------------------------------------------------------------------------------
         inf |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         unr |   52.88135   31.70039     1.67   0.103    -11.18753    116.9502
       _cons |   1.234121   1.864506     0.66   0.512    -2.534186    5.002428
```

Rescaling the independent variable multiplies the estimate of the **slope coefficient** by the inverse of the constant of multiplication (ie 1/1/100 = 100). Note that the estimate of the constant is unchanged. Why?

```
replace inf=inf/100
reg inf unr
```

```
      Source |       SS       df       MS              Number of obs =      42
-------------+------------------------------           F(  1,    40) =    2.78
       Model |  .003105856     1  .003105856           Prob > F      =  0.1031
```

```
    Residual |  .044644239    40  .001116106           R-squared     =  0.0650
-------------+----------------------------           Adj R-squared =  0.0417
       Total |  .047750095    41  .001164636           Root MSE      =  .03341
------------------------------------------------------------------------------
         inf |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         unr |   .5288135   .3170039     1.67   0.103    -.1118752    1.169502
       _cons |   .0123412   .0186451     0.66   0.512    -.0253419    .0500243
```

Rescaling both explanatory and independent variable leaves both coefficients unchanged. Why?

4. Transforming independent variable into logs

```
g lu=log(unr)
```

```
. reg inf lu

      Source |       SS       df       MS              Number of obs =      42
-------------+------------------------------           F(  1,    40) =    2.40
       Model |  .002699954     1  .002699954           Prob > F      =  0.1294
    Residual |  .045050141    40  .001126254           R-squared     =  0.0565
-------------+------------------------------           Adj R-squared =  0.0330
       Total |  .047750095    41  .001164636           Root MSE      =  .03356
------------------------------------------------------------------------------
         inf |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          lu |    .027337   .0176559     1.55   0.129     -.008347    .0630209
       _cons |   .1219227   .0517292     2.36   0.023     .0173741    .2264713
------------------------------------------------------------------------------
```

sincec this is a semi-log model (dinf/dLn(unr) = dinf/(dunr/unr) = unit change in inflation wrt % change in unemployment/100).
So change in inflation = (0.027/100)*%change in unemployment

In this case a 1% increase in the unemployment rate would raise inflation by 0.027/100 (0.027 percentage points)

```
replace unr=unr*100
```

```
g lu2=log(unr)
```

```
reg inf lu2
```

```
      Source |       SS       df       MS              Number of obs =      42
-------------+------------------------------           F(  1,    40) =    2.40
       Model |  .002699954     1  .002699954           Prob > F      =  0.1294
    Residual |  .045050141    40  .001126254           R-squared     =  0.0565
-------------+------------------------------           Adj R-squared =  0.0330
       Total |  .047750095    41  .001164636           Root MSE      =  .03356
------------------------------------------------------------------------------
         inf |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         lu2 |    .027337   .0176559     1.55   0.129     -.008347    .0630209
       _cons |  -.0039687   .0302852    -0.13   0.896    -.0651773    .0572398
------------------------------------------------------------------------------
```

So variables measured in logs are invariant to the scale in which the original variable (unemployment) is measured. This is because the log specification implies a relative effect of the variable not an absolute one.
(see question 6 on exercise 1)