Roland Locke

5/7/24

Econ210

Professor Dorian Owen

Computing Assignment

1)

$$Sleep = 3836.766 - .162totwrk - 9.94educ - 6.566age + 0.102agesq + 87.972male - 2.014spwrk75 - 67.506gdhlth$$

(235.667) (0.018) (6.033) (11.333) (0.136) (34.619) (32.263) (53.127)

$$n = 706 \qquad R^2 = 0.1248 \qquad Adj\text{–}R^2 = 0.1160$$

Stata output:

```
. . reg sleep totwrk educ age agesq male spwrk75 gdhlth

      Source |       SS           df       MS      Number of obs   =       706
-------------+----------------------------------   F(7, 698)       =     14.22
       Model |  17378940.6         7  2482705.79   Prob > F        =    0.0000
    Residual |   121860895       698   174585.81   R-squared       =    0.1248
-------------+----------------------------------   Adj R-squared   =    0.1160
       Total |   139239836       705  197503.313   Root MSE        =    417.83

------------------------------------------------------------------------------
       sleep | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      totwrk |  -.1619808   .0182252    -8.89   0.000    -.1977636    -.126198
        educ |  -9.940112   6.032809    -1.65   0.100    -21.78474    1.904515
         age |  -6.565807   11.33305    -0.58   0.563    -28.81677    15.68515
       agesq |   .1019037   .1355199     0.75   0.452    -.1641719    .3679793
        male |   87.97167   34.61854     2.54   0.011     20.00272    155.9406
     spwrk75 |  -2.014303   32.26267    -0.06   0.950    -65.35782    61.32921
      gdhlth |  -67.50613   53.12671    -1.27   0.204    -171.8134    36.80118
       _cons |   3836.766   235.6672    16.28   0.000     3374.064    4299.467
------------------------------------------------------------------------------
```

2) The question: Do men Sleep more than women can be answered by examining the *male* dummy variable. Consider the following:

$H_0: \beta_5 = 0$ (no difference in the minutes of sleep for men) vs $H_1: \beta_5 > 0$ (Positive difference in sleep for men) Where β_5 is the coefficient on the *male* dummy variable. Performing a one-tailed test with 706-7-1=698 *df,* the critical value at a 1% significance level is 2.32. The t-statistic for the *male* dummy variable is 2.54. Therefore, we reject H_0 at the 1% significance level in favor of the alternative. That is, men on average, other factors equal sleep more than women.

3) According to this data, on average for every minute worked per week, an individual will lose .162 minutes of sleep per week, holding all other things equal. That's 0.023 minutes of sleep lost each day per minute spent working, other factors equal. To determine if this is statistically significant tradeoff, consider:

$H_0\!:\beta_1 = 0$ (No tradeoff between working and sleeping) vs $H_1\!:\beta_1 < 0$ (There is tradeoff of sleep for total hours worked). Where β_1 is the coefficient on the *totwrk* variable (total mins spent working per week). When performing a one-tailed test with 06-7-1=698 *df,* at a 0.05% significance level, the critical value is -3.307. The t-statistic for the *totwrk* variable is -8.89. Therefore, we reject H_0 at the 0.05% significance level in favor of the alternative. That is, 0.023 minutes of sleep lost each day per minute spent working or .162 minutes of sleep per week is a statistically significant amount.

4) Consider the following:

$H_0\!:\beta_3 = \ \beta_4 = 0$ (*age* and *agesq* have no effect on sleep time) vs $H_1\!:\beta_3 = \ \beta_4 \neq 0$ (*age* and *agesq* have an effect on sleep time). Running this test in Stata requires the following commands:

```
.  reg sleep totwrk educ male spwrk gdhlth

      Source |       SS           df       MS      Number of obs   =       706
-------------+----------------------------------   F(5, 700)       =     19.45
       Model |  16983690.4         5  3396738.07   Prob > F        =    0.0000
    Residual |   122256145       700  174651.636   R-squared       =    0.1220
-------------+----------------------------------   Adj R-squared   =    0.1157
       Total |   139239836       705  197503.313   Root MSE        =    417.91

------------------------------------------------------------------------------
       sleep | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      totwrk |  -.1652087   .0180488    -9.15   0.000    -.2006449   -.1297725
        educ |  -11.76178   5.843448    -2.01   0.045    -23.23456    -.288991
        male |   90.65388   34.57151     2.62   0.009      22.7776    158.5302
     spwrk75 |  -4.989384   32.19909    -0.15   0.877    -68.20774    58.22897
      gdhlth |  -75.54146   52.50475    -1.44   0.151    -178.6271     27.5442
       _cons |   3785.738   85.57317    44.24   0.000     3617.727    3953.749
```

In this example, *agesq* and *age* have been omitted from the regression, which effectively holds all other factors constant omitting the effect of *age* on *sleep*. Next, using the formula for the F-test we get the result:

$$\text{F}(2, 698) = \frac{(122256145-121860895)/2}{121860895/698} = 1.13$$

Checking this result against the Stata test command, we get the same results:

```
. test age agesq

 ( 1)  age = 0
 ( 2)  agesq = 0

       F(  2,   698) =    1.13
            Prob > F =    0.3230
```

To find the critical values of this F-test, we use 2 degrees of freedom in the numerator and 698 degrees of freedom in the denominator. With a %5 significance level the critical value is 3.00. Meaning, with a F-value of 1.13 these results don't pass at the %5 significance level so, we don't reject the null hypothesis. Therefore, age does have a statistically significant effect on sleep time according to this model.

5) a) Using the above model, consider the hypothesis $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$. That is, β_1 through β_7 all have no effect on sleep time. The alternative hypothesis is: H_1: at least one of β_1 through β_7 has a non-zero effect on sleep time. Running this test in Stata requires the following input:

```
. reg sleep

      Source |       SS           df       MS      Number of obs   =       706
-------------+----------------------------------   F(0, 705)       =      0.00
       Model |           0         0           .   Prob > F        =         .
    Residual |   139239836       705  197503.313   R-squared       =    0.0000
-------------+----------------------------------   Adj R-squared   =    0.0000
       Total |   139239836       705  197503.313   Root MSE        =    444.41

------------------------------------------------------------------------------
       sleep | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
```

```
-------------+----------------------------------------------------------------
       _cons |   3266.356   16.72572   195.29   0.000     3233.517    3299.194
------------------------------------------------------------------------------
```

All independent variables have been omitted from the regression. For the F-test we get the result:

$$\text{F}(7, 698) = \frac{0.1248/7}{(1-.1248)/698} = 14.22$$

Comparing to the Stata test command:

```
. test totwrk educ age agesq male spwrk75 gdhlth

 ( 1)  totwrk = 0
 ( 2)  educ = 0
 ( 3)  age = 0
 ( 4)  agesq = 0
 ( 5)  male = 0
 ( 6)  spwrk75 = 0
 ( 7)  gdhlth = 0

       F(  7,   698) =   14.22
            Prob > F =    0.0000
```

To find the critical values of this F-test, we use 7 degrees of freedom in the numerator and 698 degrees of freedom in the denominator. With a %5 significance level the critical value is 2.01. Meaning, with an F-value of 14.22 the result far exceeds the critical value. We reject the null hypothesis in favor of the alternative. Therefore, β_1 through β_7 have a statistically significant non-zero effect on *sleep*.

5) b) Consider the hypothesis $H_0: \beta_1 = -0.15, \beta_2 = -10,\ \beta_3 = \beta_4, \beta_7 = 0$. That is, each unit of β_1 (*totwrk*) has on average a -0.15 effect on sleep time, β_2 (*educ*) has on average -10 effect on sleep time. The effects of β_3 and $\beta_4$are on average equal, and β_7 has no statistically significant effect on sleep time. The alternative hypothesis is: $H_1: \beta_1 \neq -0.15, \beta_2 \neq -10,\ \beta_3 \neq \beta_4,\ \beta_7$ has a non-zero effect on sleep time. Running this test in Stata requires the following input:

```
. gen newdepvar = sleep + 0.15*totwrk + 10*educ

. gen agecomb = age + agesq
```

```
. reg newdepvar agecomb male spwrk75

      Source |       SS           df       MS      Number of obs   =       706
-------------+----------------------------------   F(3, 702)       =      2.95
       Model |  1541681.74         3  513893.913   Prob > F        =    0.0322
    Residual |   122387661       702  174341.397   R-squared       =    0.0124
-------------+----------------------------------   Adj R-squared   =    0.0082
       Total |   123929343       705  175786.302   Root MSE        =    417.54

------------------------------------------------------------------------------
   newdepvar | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
     agecomb |   .0272646   .0164226     1.66   0.097    -.0049787    .0595079
        male |   75.62578   32.13031     2.35   0.019     12.54276    138.7088
     spwrk75 |  -4.512802   31.98306    -0.14   0.888    -67.30672    58.28112
       _cons |   3626.278   41.45892    87.47   0.000     3544.879    3707.676
------------------------------------------------------------------------------
```

The Sum of square residuals from this regression as well as the original, are used to calculate the F-test for these restrictions:

$$F(4, 701) = \frac{(122387661 - 121860895)/4}{(121860895)/701} = 0.75$$

Comparing to the Stata test command:

```
test (totwrk=-0.15) (educ=-10) (age-agesq=0) (gdhlth=0)

 ( 1)  totwrk = -.15
 ( 2)  educ = -10
 ( 3)  age - agesq = 0
 ( 4)  gdhlth = 0

       F(  4,   698) =    0.75
            Prob > F =    0.5553
```

The critical values of this F-test, we use 4 degrees of freedom in the numerator and 701 degrees of freedom in the denominator. With a %5 significance level, the critical value is 2.37. With an F-test result of 0.75 the result doesn't exceed the critical value. We do not reject the null hypothesis. Therefore, each unit of β_1 (*totwrk*) has on average a -0.15 effect on sleep time, β_2 (*educ*) has on average -10 effect on sleep time. The effects of β_3 and $\beta_4$are on average equal, and β_7 has no statistically significant effect on sleep time.

6) Stata output for testing this equation separately for both men and women:

```
. reg sleep totwrk educ age agesq spwrk75 gdhlth yngkid if male == 1

      Source |       SS           df       MS      Number of obs   =       400
-------------+----------------------------------   F(7, 392)       =     10.85
       Model |  12269936.7         7   1752848.1   Prob > F        =    0.0000
    Residual |  63300203.9       392  161480.112   R-squared       =    0.1624
-------------+----------------------------------   Adj R-squared   =    0.1474
       Total |  75570140.6       399  189398.849   Root MSE        =    401.85

------------------------------------------------------------------------------
       sleep | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      totwrk |  -.1776878   .0245981    -7.22   0.000    -.2260486    -.129327
        educ |  -10.16235   7.606884    -1.34   0.182    -25.11775    4.793038
         age |    10.0324   14.52225     0.69   0.490    -18.51884    38.58363
       agesq |   -.076999   .1705628    -0.45   0.652    -.4123313    .2583333
     spwrk75 |  -24.64174   41.14681    -0.60   0.550    -105.5378    56.25429
      gdhlth |  -111.0472    71.7303    -1.55   0.122    -252.0714    29.97699
      yngkid |   58.51314   59.02528     0.99   0.322    -57.53258    174.5589
       _cons |   3651.975   311.3274    11.73   0.000     3039.894    4264.055
------------------------------------------------------------------------------

. reg sleep totwrk educ age agesq spwrk75 gdhlth yngkid if male == 0

      Source |       SS           df       MS      Number of obs   =       306
-------------+----------------------------------   F(7, 298)       =      4.75
       Model |  6374582.21         7  910654.601   Prob > F        =    0.0000
    Residual |  57115569.9       298  191662.986   R-squared       =    0.1004
-------------+----------------------------------   Adj R-squared   =    0.0793
       Total |  63490152.1       305  208164.433   Root MSE        =    437.79

------------------------------------------------------------------------------
       sleep | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      totwrk |  -.1344332   .0283125    -4.75   0.000     -.190151   -.0787154
        educ |  -10.10285   9.830513    -1.03   0.305    -29.44888    9.243167
         age |  -31.55383   18.97683    -1.66   0.097    -68.89941    5.791746
       agesq |   .3838169    .229881     1.67   0.096     -.068579    .8362127
     spwrk75 |   48.12266   52.95735     0.91   0.364    -56.09511    152.3404
      gdhlth |  -29.10627    80.8454    -0.36   0.719    -188.2065     129.994
      yngkid |  -118.3917   93.45786    -1.27   0.206    -302.3127    65.52934
       _cons |   4246.316   385.9974    11.00   0.000      3486.69    5005.942
```

Based on these results alone, the coefficient on *yngkid, age,* and *spwrks* are vastly different in the two models. The *yngkid* coefficient specifically implies all things equal, men gain sleep in a relationship when a child under three is present, while women, all things equal, loses sleep when a child under three is present. Important to note here is the sample size in between the different models (400 in the male model 306 for women). The R squared values are also very low meaning the given variables don't predict the dependent variable. According to the F statistic, the variance in the women's model is much lower than that of the men's, however both have very high Probability greater than F stats, indicating both models are highly statistically significant.

7) a) If we consider that the being a man or a woman may influence sleep time. We need to determine whether a model pooling data collected from both men and women is adequate specification. To do so we will perform a Chow Test. Chow test for the given model performed in Stata:

```
. reg sleep totwrk educ age agesq spwrk75 gdhlth yngkid

      Source |       SS           df       MS      Number of obs   =       706
-------------+----------------------------------   F(7, 698)       =     13.19
       Model |  16270817.8         7  2324402.54   Prob > F        =    0.0000
    Residual |   122969018       698  176173.378   R-squared       =    0.1169
-------------+----------------------------------   Adj R-squared   =    0.1080
       Total |   139239836       705  197503.313   Root MSE        =    419.73

------------------------------------------------------------------------------
       sleep | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      totwrk |   -.145276   .0170752    -8.51   0.000    -.1788009   -.1117511
        educ |  -9.323139   6.056106    -1.54   0.124    -21.21351    2.567228
         age |  -6.050771   11.49615    -0.53   0.599    -28.62196    16.52041
       agesq |   .1000146   .1368081     0.73   0.465      -.16859    .3686192
     spwrk75 |   -11.9582   32.17956    -0.37   0.710    -75.13853    51.22213
      gdhlth |  -64.71073   53.35679    -1.21   0.226    -169.4698    40.04831
      yngkid |   16.56249   50.07271     0.33   0.741    -81.74868    114.8737
       _cons |   3826.507   240.9338    15.88   0.000     3353.465    4299.549
------------------------------------------------------------------------------
```

```
. reg sleep totwrk educ age agesq spwrk75 gdhlth yngkid if male == 1

      Source |       SS           df       MS      Number of obs   =       400
-------------+----------------------------------   F(7, 392)       =     10.85
       Model |  12269936.7         7   1752848.1   Prob > F        =    0.0000
    Residual |  63300203.9       392  161480.112   R-squared       =    0.1624
-------------+----------------------------------   Adj R-squared   =    0.1474
       Total |  75570140.6       399  189398.849   Root MSE        =    401.85

------------------------------------------------------------------------------
       sleep | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      totwrk |  -.1776878   .0245981    -7.22   0.000    -.2260486    -.129327
        educ |  -10.16235   7.606884    -1.34   0.182    -25.11775    4.793038
         age |    10.0324   14.52225     0.69   0.490    -18.51884    38.58363
       agesq |   -.076999   .1705628    -0.45   0.652    -.4123313    .2583333
     spwrk75 |  -24.64174   41.14681    -0.60   0.550    -105.5378    56.25429
      gdhlth |  -111.0472    71.7303    -1.55   0.122    -252.0714    29.97699
      yngkid |   58.51314   59.02528     0.99   0.322    -57.53258    174.5589
       _cons |   3651.975   311.3274    11.73   0.000     3039.894    4264.055
------------------------------------------------------------------------------

. reg sleep totwrk educ age agesq spwrk75 gdhlth yngkid if male == 0

      Source |       SS           df       MS      Number of obs   =       306
-------------+----------------------------------   F(7, 298)       =      4.75
       Model |  6374582.21         7  910654.601   Prob > F        =    0.0000
    Residual |  57115569.9       298  191662.986   R-squared       =    0.1004
-------------+----------------------------------   Adj R-squared   =    0.0793
       Total |  63490152.1       305  208164.433   Root MSE        =    437.79

------------------------------------------------------------------------------
       sleep | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      totwrk |  -.1344332   .0283125    -4.75   0.000     -.190151   -.0787154
        educ |  -10.10285   9.830513    -1.03   0.305    -29.44888    9.243167
         age |  -31.55383   18.97683    -1.66   0.097    -68.89941    5.791746
       agesq |   .3838169    .229881     1.67   0.096     -.068579    .8362127
     spwrk75 |   48.12266   52.95735     0.91   0.364    -56.09511    152.3404
      gdhlth |  -29.10627    80.8454    -0.36   0.719    -188.2065     129.994
      yngkid |  -118.3917   93.45786    -1.27   0.206    -302.3127    65.52934
       _cons |   4246.316   385.9974    11.00   0.000      3486.69    5005.942
------------------------------------------------------------------------------
```

Performing the math for the Chow test result:

$$F(8, 690) = \frac{(122969018 - (63300203.9+57115569.9))/8}{(63300203.9+57115569.9)/690} = 1.83$$

7) b) Another way to find the define the difference in sleep time between men and women through terms in the model is using interaction terms and a dummy variable. The following is a Stata test using the *male* dummy variable and a set of interaction terms involving the *male* term:

```
. reg sleep totwrk educ age agesq spwrk75 gdhlth yngkid male

      Source |       SS           df       MS      Number of obs   =       706
-------------+----------------------------------   F(8, 697)       =     12.43
       Model |  17378940.8         8   2172367.6   Prob > F        =    0.0000
    Residual |   121860895       697  174836.291   R-squared       =    0.1248
-------------+----------------------------------   Adj R-squared   =    0.1148
       Total |   139239836       705  197503.313   Root MSE        =    418.13

------------------------------------------------------------------------------
       sleep | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      totwrk |  -.1619798   .0182585    -8.87   0.000     -.197828   -.1261316
        educ |   -9.94024   6.038058    -1.65   0.100     -21.7952    1.914721
         age |   -6.56386   11.45426    -0.57   0.567    -29.05284    15.92512
       agesq |   .1018873   .1362899     0.75   0.455    -.1657008    .3694753
     spwrk75 |  -2.013155   32.29968    -0.06   0.950    -65.42949    61.40318
      gdhlth |  -67.50647   53.16552    -1.27   0.205    -171.8902     36.8773
      yngkid |   .0610275   50.31112     0.00   0.999    -98.71849    98.84055
        male |   87.96615   34.94117     2.52   0.012      19.3636    156.5687
       _cons |   3836.711    240.052    15.98   0.000       3365.4    4308.023
------------------------------------------------------------------------------


. gen maletotwrk = male*totwrk

. gen maleeduc = male*educ

. gen maleage = male*age

. gen maleagesq = male*agesq

. gen malespwrk75 = male*spwrk75

. gen malegdhlth = male*gdhlth

. gen maleyngkid = male*yngkid
```

```
. reg sleep male maletotwrk totwrk maleeduc educ maleage age maleagesq agesq
malespwrk75 spwrk75 malegdhlth gdhlth maleyngkid yngkid

      Source |       SS           df       MS      Number of obs   =       706
-------------+----------------------------------   F(15, 690)      =      7.19
       Model |    18824062        15  1254937.47   Prob > F        =    0.0000
    Residual |   120415774       690  174515.614   R-squared       =    0.1352
-------------+----------------------------------   Adj R-squared   =    0.1164
       Total |   139239836       705  197503.313   Root MSE        =    417.75

------------------------------------------------------------------------------
       sleep | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
        male |   -594.341   490.3194    -1.21   0.226    -1557.038     368.356
  maletotwrk |  -.0432546   .0371994    -1.16   0.245    -.1162922     .029783
      totwrk |  -.1344332   .0270163    -4.98   0.000    -.1874773   -.0813891
    maleeduc |  -.0595001   12.26902    -0.00   0.996    -24.14859    24.02959
        educ |  -10.10285   9.380461    -1.08   0.282    -28.52053    8.314818
     maleage |   41.58623   23.57587     1.76   0.078    -4.702832    87.87529
         age |  -31.55383   18.10805    -1.74   0.082    -67.10732     3.99966
   maleagesq |  -.4608159   .2820594    -1.63   0.103    -1.014614    .0929819
       agesq |   .3838169   .2193568     1.75   0.081    -.0468701    .8145038
 malespwrk75 |  -72.76439   66.20655    -1.10   0.272    -202.7549    57.22607
     spwrk75 |   48.12266   50.53291     0.95   0.341    -51.09406    147.3394
  malegdhlth |  -81.94095   107.2931    -0.76   0.445    -292.6011    128.7192
      gdhlth |  -29.10627   77.14421    -0.38   0.706    -180.5718    122.3593
  maleyngkid |   176.9048   108.2505     1.63   0.103    -35.63508    389.4447
      yngkid |  -118.3917   89.17926    -1.33   0.185    -293.4869     56.7036
       _cons |   4246.316    368.326    11.53   0.000     3523.141     4969.49
------------------------------------------------------------------------------
```

Results for the interaction term test were found using the test command in Stata. The results are shown below:

```
. test male maletotwrk maleeduc maleage maleagesq malespwrk75 malegdhlth maleyngkid

 ( 1)  male = 0
 ( 2)  maletotwrk = 0
 ( 3)  maleeduc = 0
 ( 4)  maleage = 0
 ( 5)  maleagesq = 0
 ( 6)  malespwrk75 = 0
 ( 7)  malegdhlth = 0
 ( 8)  maleyngkid = 0

       F(  8,   690) =    1.83
            Prob > F =    0.0687
```

Both tests produced a result of 1.83. For the critical values of this F-test, we use 8 degrees of freedom in the numerator and 690 degrees of freedom in the denominator. With a %5 significance level the critical value is 1.88. With an F-test result of 1.83 the result doesn't exceed the critical value. Therefore, the pooled function is an adequate specification, and we don't need to run separate specifications for men and women. However, including the interaction terms and dummy variables may allow more insight into the variables effecting sleep time for men and women.

8) Testing whether, the interaction terms are jointly significant using the test command in Stata:

```
. test maletotwrk maleeduc maleage maleagesq malespwrk75 malegdhlth maleyngkid

 ( 1)  maletotwrk = 0
 ( 2)  maleeduc = 0
 ( 3)  maleage = 0
 ( 4)  maleagesq = 0
 ( 5)  malespwrk75 = 0
 ( 6)  malegdhlth = 0
 ( 7)  maleyngkid = 0

       F(  7,   690) =    1.18
            Prob > F =    0.3101
```

Again, the f-statistic doesn't exceed the critical value of a 5% significance level and therefore, including the *male* interaction terms in the model is not jointly statistically significant with the interaction terms.

9) Given the three models in 8 and 7a and 7b, the preferred model is number 8 that with a dummy variable allows for a different intercept for males and females. None of the models exceed the 5% critical value in the F-test and therefore, don't require separate specifications for men and women. However, in the original regression model with a dummy variable and without the interaction terms, the male variable was one of only two independent variables with a statistical significance of five percent. Therefore, with a high statistical significance, including it in the model allows the intercepts for men and women to be different.

10) Breusch-pagan LM test is performed to detect heteroscedasticity as with it present the OLS estimator (the modern standard estimator) is no longer the best linear unbiased estimator or BLUE. In the first step, we assume that MLR.1 through MLR.4 holds true and run the OLS estimator on the model. Then using the square of the residuals, as the dependent variable the model is run again. An LM Statistic used on a Chi-Squared distribution to determine the statistical significance. The second regression reveals whether the variance of the residuals matches the variance of the independent variables.

Performing a Breusch-pagan LM test for heteroskedasticity:

```
. reg sleep totwrk educ age agesq spwrk75 gdhlth yngkid male

      Source |       SS           df       MS      Number of obs   =       706
-------------+----------------------------------   F(8, 697)       =     12.43
       Model |  17378940.8         8   2172367.6   Prob > F        =    0.0000
    Residual |   121860895       697  174836.291   R-squared       =    0.1248
-------------+----------------------------------   Adj R-squared   =    0.1148
       Total |   139239836       705  197503.313   Root MSE        =    418.13

------------------------------------------------------------------------------
       sleep | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      totwrk |  -.1619798   .0182585    -8.87   0.000     -.197828   -.1261316
        educ |   -9.94024   6.038058    -1.65   0.100     -21.7952    1.914721
         age |   -6.56386   11.45426    -0.57   0.567    -29.05284    15.92512
       agesq |   .1018873   .1362899     0.75   0.455    -.1657008    .3694753
     spwrk75 |  -2.013155   32.29968    -0.06   0.950    -65.42949    61.40318
      gdhlth |  -67.50647   53.16552    -1.27   0.205    -171.8902     36.8773
      yngkid |   .0610275   50.31112     0.00   0.999    -98.71849    98.84055
        male |   87.96615   34.94117     2.52   0.012      19.3636    156.5687
       _cons |   3836.711    240.052    15.98   0.000       3365.4    4308.023
------------------------------------------------------------------------------

. predict uh, resid

. gen uhsq = uh^2

. reg uhsq totwrk educ age agesq spwrk75 gdhlth yngkid male

      Source |       SS           df       MS      Number of obs   =       706
-------------+----------------------------------   F(8, 697)       =      1.62
       Model |  1.6553e+12         8  2.0691e+11   Prob > F        =    0.1163
    Residual |  8.9178e+13       697  1.2794e+11   R-squared       =    0.0182
-------------+----------------------------------   Adj R-squared   =    0.0070
       Total |  9.0833e+13       705  1.2884e+11   Root MSE        =    3.6e+05
```

```
------------------------------------------------------------------------------
        uhsq | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      totwrk |   17.09085   15.61925     1.09   0.274    -13.57557    47.75727
        educ |  -9030.051    5165.27    -1.75   0.081     -19171.4    1111.302
         age |  -8175.251    9798.57    -0.83   0.404     -27413.5       11063
       agesq |   59.24164   116.5895     0.51   0.612    -169.6671    288.1504
     spwrk75 |  -29976.39   27630.83    -1.08   0.278    -84226.03    24273.26
      gdhlth |  -21977.36   45480.56    -0.48   0.629    -111272.7    67317.96
      yngkid |   1702.656   43038.76     0.04   0.968     -82798.5    86203.81
        male |  -41564.23    29890.5    -1.39   0.165    -100250.4    17121.97
       _cons |   529497.8     205353     2.58   0.010     126313.1    932682.4
------------------------------------------------------------------------------
```

Using the result from the second regression using the square of the residuals we calculate the LM statistic:

$$\text{LM} = nR_{Uhat^2}{}^2 = 706 * 0.0182 \approx 12.85$$

The test included eight terms making this test's null hypothesis: $H_0\!: \delta_2 = \delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = \delta_6 = \delta_7 = 0$. That is, the model is homoscedastic. While the alternative hypothesis was: The model is heteroscedastic. At a 5% significance level with 8 degrees of freedom the Chi-Squared distribution has a critical value at 15.5. The test's LM statistic is under that value and therefore we do not reject the null and instead assume the alternative; the model is homoscedastic.

11) Performing the same test with hetero-robust standard errors in Stata:

```
. reg sleep totwrk educ age agesq spwrk75 gdhlth yngkid male, robust

Linear regression                               Number of obs     =        706
                                                F(8, 697)         =      11.09
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1248
                                                Root MSE          =     418.13

------------------------------------------------------------------------------
             |               Robust
       sleep | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
      totwrk |  -.1619798    .020483    -7.91   0.000    -.2021955   -.1217641
```

```
        educ |   -9.94024   5.776834    -1.72   0.086    -21.28232    1.401842
         age |   -6.56386    11.9462    -0.55   0.583    -30.01871    16.89099
       agesq |   .1018873   .1380079     0.74   0.461    -.1690738    .3728484
     spwrk75 |  -2.013155   31.85884    -0.06   0.950    -64.56396    60.53765
      gdhlth |  -67.50647   57.94971    -1.16   0.244    -181.2834    46.27046
      yngkid |   .0610275   54.07073     0.00   0.999       -106.1    106.2221
        male |   87.96615   35.92353     2.45   0.015     17.43486    158.4974
       _cons |   3836.711   261.1577    14.69   0.000     3323.961    4349.461
------------------------------------------------------------------------------
```

This regression displays values that are heteroscedasticity-robust standard errors, if there was heteroscedasticity in the model the statistical significance of the variables would be preserved. However, as was discovered in result of question 10, the model was already homoscedastic. Moreover, when comparing this result to the non-hetero robust function, we see that the model displays statistical significance at the five-percent level for only the *totwrk* and *male* coefficients which matches the results we see in the non-hetero robust model. Therefore, there is no reason to compute the model with hetero robust standard errors.