HW6

Omair Shafi Ahmed

A

```
. use "E:\Master 2\ECONOMETRICS\Problem set 6\insurance.dta"

. regress insured selfemp, vce(robust)

Linear regression                               Number of obs     =      8,802
                                                F(1, 8800)        =      74.23
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0109
                                                Root MSE          =     .39697

------------------------------------------------------------------------------
             |               Robust
     insured |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     selfemp |  -.1276363   .0148142    -8.62   0.000    -.1566755   -.0985971
       _cons |   .8167119   .0044008   185.58   0.000     .8080853    .8253386
------------------------------------------------------------------------------
```

The coefficient of selfemp, being negative shows that self employed are less likely to have health insurance. This is statistically significant at the 1% level.

B

```
. regress insured age deg_nd deg_ged deg_hs deg_ba deg_ma deg_phd selfemp, vce(robust)

Linear regression                               Number of obs     =      8,802
                                                F(8, 8793)        =     128.94
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1159
                                                Root MSE          =     .37546

------------------------------------------------------------------------------
             |               Robust
     insured |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |    .005762   .0003793    15.19   0.000     .0050184    .0065056
      deg_nd |  -.3291785   .0194873   -16.89   0.000    -.3673781   -.2909789
     deg_ged |  -.1590603    .026564    -5.99   0.000    -.2111319   -.1069887
      deg_hs |  -.0500131   .0142351    -3.51   0.000    -.0779173    -.022109
      deg_ba |   .0252242   .0151998     1.66   0.097    -.0045709    .0550194
      deg_ma |   .0453013   .0168376     2.69   0.007     .0122956    .0783069
     deg_phd |   .0745985   .0238487     3.13   0.002     .0278494    .1213475
     selfemp |  -.1684386   .0141335   -11.92   0.000    -.1961436   -.1407336
       _cons |   .6628445   .0200892    33.00   0.000      .623465    .7022241
------------------------------------------------------------------------------
```

Controlling for other factors does not change the negative effect of being self employed on insurance.

C

Increasing the age by 1 year will lead to an increase in the probability of having insurance by 0.57%.

D

```
. gen ageself=age*selfemp

. regress insured ageself age deg_nd deg_ged deg_hs deg_ba deg_ma deg_phd selfemp, vce(robust)

Linear regression                               Number of obs     =      8,802
                                                F(9, 8792)        =     115.00
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1159
                                                Root MSE          =     .37548

------------------------------------------------------------------------------
             |               Robust
     insured |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     ageself |  -.0005264   .0013917    -0.38   0.705    -.0032545    .0022016
         age |   .0058137    .000394    14.76   0.000     .0050413     .006586
      deg_nd |   -.329265    .019492   -16.89   0.000    -.3674739   -.2910561
     deg_ged |  -.1590831   .0265716    -5.99   0.000    -.2111698   -.1069965
      deg_hs |  -.0500587   .0142355    -3.52   0.000    -.0779635   -.0221539
      deg_ba |   .0252024   .0151994     1.66   0.097    -.0045919    .0549968
      deg_ma |   .0452726   .0168391     2.69   0.007     .0122641    .0782811
     deg_phd |   .0743943   .0238483     3.12   0.002      .027646    .1211425
     selfemp |    -.14568   .0632744    -2.30   0.021    -.2697126   -.0216475
       _cons |     .66091   .0205355    32.18   0.000     .6206556    .7011644
------------------------------------------------------------------------------
```

A p value of 0.705 indicates that the effect of self employment on insurance is about the same for older workers as it is for younger workers.

E

```
. regress healthy ageself selfemp age  deg_nd deg_ged deg_hs deg_ba deg_ma deg_phd, vce(robust)

Linear regression                               Number of obs     =      8,802
                                                F(9, 8792)        =      18.31
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0287
                                                Root MSE          =     .25401
```

healthy	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
ageself	-.0007351	.000799	-0.92	0.358	-.0023012	.0008311
selfemp	.0492291	.0329363	1.49	0.135	-.0153338	.1137921
age	-.0018452	.0002739	-6.74	0.000	-.0023822	-.0013083
deg_nd	-.1151463	.0138056	-8.34	0.000	-.1422084	-.0880841
deg_ged	-.0700683	.0186157	-3.76	0.000	-.1065595	-.0335771
deg_hs	-.0163468	.0091456	-1.79	0.074	-.0342743	.0015808
deg_ba	.0076078	.009766	0.78	0.436	-.0115358	.0267514
deg_ma	.0134788	.0123512	1.09	0.275	-.0107324	.03769
deg_phd	.0242518	.0167961	1.44	0.149	-.0086724	.0571761
_cons	1.021649	.0130736	78.15	0.000	.9960214	1.047276

There doesn't appear to be a significant difference in health for people who are wage earners and people who are self employed. This could be a two way causality problem as self employed workers may have worse health as a result of not being insured and not being insured could affect productivity keeping them out of the formal labour markets.

1

```
. mean age deg_nd deg_ged deg_hs deg_ba deg_ma deg_phd

Mean estimation                   Number of obs    =       8,802
```

	Mean	Std. Err.	[95% Conf.	Interval]
age	38.93683	.1184283	38.70469	39.16898
deg_nd	.1271302	.0035509	.1201697	.1340907
deg_ged	.0424903	.0021501	.0382757	.046705
deg_hs	.5037491	.0053296	.493302	.5141963
deg_ba	.1759827	.0040592	.1680258	.1839397
deg_ma	.0595319	.0025222	.0545878	.064476
deg_phd	.0153374	.0013099	.0127696	.0179052

```
. scalar c0=_b[_cons]+_b[age]*38.93683+_b[ deg_nd ]*0.1271302+_b[ deg_ged ]*0.04249
> 03+_b[ deg_hs ]*0.5037491+_b[ deg_ba ]*0.1759827+_b[ deg_ma ]*0.0595319+_b[ deg_p
> hd ]*0.0153374

. scalar c1=_b[_cons]+_b[age]*38.93683+_b[ deg_nd ]*0.1271302+_b[ deg_ged ]*0.04249
> 03+_b[ deg_hs ]*0.5037491+_b[ deg_ba ]*0.1759827+_b[ deg_ma ]*0.0595319+_b[ deg_p
> hd ]*0.0153374+_b[ selfemp ]

. scalar dir
       c1 =  .65323801
       c0 =  .82167663
```

The probability of having health insurance for someone who is not self employed and average in every other sense is 82.16% whereas the probability of someone who is self employed having health insurance is 65.32%.

2

```
. logit insured age deg_nd deg_ged deg_hs deg_ba deg_ma deg_phd selfemp, vce(robust)

Iteration 0:   log pseudolikelihood = -4390.0862
Iteration 1:   log pseudolikelihood = -3939.1256
Iteration 2:   log pseudolikelihood = -3900.6588
Iteration 3:   log pseudolikelihood = -3900.3411
Iteration 4:   log pseudolikelihood = -3900.3405
Iteration 5:   log pseudolikelihood = -3900.3405

Logistic regression                             Number of obs     =      8,802
                                                Wald chi2(8)      =     871.05
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -3900.3405               Pseudo R2         =     0.1116

------------------------------------------------------------------------------
             |               Robust
     insured |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .0401036   .0027563    14.55   0.000     .0347014    .0455058
      deg_nd |  -1.802655   .1339189   -13.46   0.000    -2.065131   -1.540179
     deg_ged |  -1.030621   .1669408    -6.17   0.000    -1.357819   -.7034228
      deg_hs |  -.3845205   .1245314    -3.09   0.002    -.6285975   -.1404434
      deg_ba |   .2370634   .1441664     1.64   0.100    -.0454975    .5196244
      deg_ma |   .6646674   .2159877     3.08   0.002     .2413392    1.087996
     deg_phd |   .9635204   .3818707     2.52   0.012     .2150676    1.711973
     selfemp |  -1.089629    .080422   -13.55   0.000    -1.247253   -.9320045
       _cons |   .5363559   .1530594     3.50   0.000     .2363649    .8363468
------------------------------------------------------------------------------
```

```
. logit insured age deg_ged  deg_nd deg_hs deg_ba deg_ma deg_phd selfemp, vce(robus
> t) or

Iteration 0:   log pseudolikelihood = -4390.0862
Iteration 1:   log pseudolikelihood = -3939.1256
Iteration 2:   log pseudolikelihood = -3900.6588
Iteration 3:   log pseudolikelihood = -3900.3411
Iteration 4:   log pseudolikelihood = -3900.3405
Iteration 5:   log pseudolikelihood = -3900.3405

Logistic regression                             Number of obs     =      8,802
                                                Wald chi2(8)      =     871.05
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -3900.3405               Pseudo R2         =     0.1116

------------------------------------------------------------------------------
             |               Robust
     insured | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   1.040919   .0028691    14.55   0.000      1.03531    1.046557
     deg_ged |   .3567854    .059562    -6.17   0.000     .2572212    .4948885
      deg_nd |   .1648606   .0220779   -13.46   0.000     .1268016    .2143427
      deg_hs |    .680777   .0847781    -3.09   0.002     .5333393    .8689728
      deg_ba |   1.267522    .182734     1.64   0.100      .955522    1.681396
      deg_ma |   1.943844   .4198465     3.08   0.002     1.272953    2.968318
     deg_phd |   2.620907   1.000848     2.52   0.012     1.239946    5.539882
     selfemp |   .3363413   .0270493   -13.55   0.000     .2872929    .3937636
       _cons |   1.709765   .2616956     3.50   0.000     1.266636     2.30792
------------------------------------------------------------------------------
```

The logic model shows that the probability of someone who is self employed having health insurance is about one-thirds (33.63) as that of someone who is self employed.