Name: Imisi Raphael Aiyetan
Course: Econometrics 512

Problem Set 5

Solution

Problem 1

Question (a):

```
. * The regular robust standard errors
.
. xtreg lrpdi lsales, fe vce(robust)

Fixed-effects (within) regression               Number of obs      =        28
Group variable: id                              Number of groups   =         7

R-sq:  within  = 0.0070                         Obs per group: min =         3
       between = 0.4037                                        avg =       4.0
       overall = 0.0449                                        max =         5

                                                F(1,6)             =      0.40
corr(u_i, Xb)  = 0.1999                         Prob > F           =    0.5510

                                     (Std. Err. adjusted for 7 clusters in id)
------------------------------------------------------------------------------
             |               Robust
       lrpdi |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lsales |   .0042463   .0067238     0.63   0.551    -.0122061    .0206988
       _cons |   4.512929   .0306475   147.25   0.000     4.437938    4.587921
-------------+----------------------------------------------------------------
     sigma_u |  .06705105
     sigma_e |   .0135427
         rho |  .96080464   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

```
. * Clustering  by i, the of cross-sectional observation
.
. xtreg lrpdi lsales, fe vce(cluster id)

Fixed-effects (within) regression               Number of obs      =        28
Group variable: id                              Number of groups   =         7

R-sq:  within  = 0.0070                         Obs per group: min =         3
       between = 0.4037                                        avg =       4.0
       overall = 0.0449                                        max =         5

                                                F(1,6)             =      0.40
corr(u_i, Xb)  = 0.1999                         Prob > F           =    0.5510

                                   (Std. Err. adjusted for 7 clusters in id)
------------------------------------------------------------------------------
             |               Robust
       lrpdi |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lsales |   .0042463   .0067238     0.63   0.551    -.0122061    .0206988
       _cons |   4.512929   .0306475   147.25   0.000     4.437938    4.587921
-------------+----------------------------------------------------------------
     sigma_u |  .06705105
     sigma_e |   .0135427
         rho |  .96080464   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Interpretation:
Ignoring in fixed effect, we realized that regression for robust standard and when we grouped individuals in cluster the result of the two analyses are the same. By implication, when we account for unknown structure of variation and heteroskedasticity/differences across clusters of observation the results are the same However, the estimated parameter for *lsales* is insignificant for both methods. We present the STATA output in Appendix 1and 2. Also, the programming codes are shown in Appendix 13.

Question (b):

```
. egen lsale_bar = mean(lsales)

.
. egen lrpdi_bar =mean(lrpdi)

.
. gen z = lsales - lsale_bar

.
. gen q = lrpdi - lrpdi_bar

.
. reg z q

      Source |       SS       df       MS              Number of obs =      28
-------------+------------------------------           F(  1,    26) =    1.22
       Model |  .071949325     1  .071949325           Prob > F      =  0.2789
    Residual |  1.52932909    26   .05882035           R-squared     =  0.0449
-------------+------------------------------           Adj R-squared =  0.0082
       Total |  1.60127842    27  .059306608           Root MSE      =  .24253

------------------------------------------------------------------------------
           z |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           q |   .7995552   .7229344     1.11   0.279    -.6864578    2.285568
       _cons |   6.82e-08   .0458337     0.00   1.000    -.0942124    .0942126
------------------------------------------------------------------------------
```

```
. egen l_sale_bar = mean(lsales), by(id)

.
. egen l_rpdi_bar =mean(lrpdi), by(id)

.
. gen x = lsales - l_sale_bar

.
. gen y = lrpdi - l_rpdi_bar

.
. reg y x
```

Source	SS	df	MS
Model	.000026032	1	.000026032
Residual	.003668096	26	.000141081
Total	.003694129	27	.00013682

Number of obs	= 28
F(1, 26)	= 0.18
Prob > F	= 0.6711
R-squared	= 0.0070
Adj R-squared	= -0.0311
Root MSE	= .01188

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
x	.0042463	.0098853	0.43	0.671	-.0160732	.0245659
_cons	-1.54e-07	.0022447	-0.00	1.000	-.0046142	.0046139

Interpretation:
Estimating the fixed effect model using demean approach, we find that the estimated parameter of *x and q* are different. Although the results for both approaches are insignificant, but it suggests that when the fixed effect model is estimated and we account for difference/heteroskedasticity across group, the results are different. We present the STATA output in Appendix 2 and 3. Also, the programming codes are shown in Appendix 13

Question (C):

```
. reg lrpdi lsales d2 d3 d4 d5 d6 d7
```

Source	SS	df	MS
Model	.108981639	7	.015568806
Residual	.003564294	20	.000178215
Total	.112545933	27	.004168368

Number of obs	= 28
F(7, 20)	= 87.36
Prob > F	= 0.0000
R-squared	= 0.9683
Adj R-squared	= 0.9572
Root MSE	= .01335

lrpdi	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lsales	.0050718	.0109675	0.46	0.649	-.017806	.0279495
d2	.0241732	.0095558	2.53	0.020	.0042403	.0441062
d3	.0675644	.0095742	7.06	0.000	.047593	.0875359
d4	.1227823	.0097271	12.62	0.000	.1024918	.1430727
d5	.1309612	.0097047	13.49	0.000	.1107177	.1512048
d6	.1532011	.0096012	15.96	0.000	.1331732	.1732289
d7	.177454	.0095976	18.49	0.000	.1574336	.1974743
_cons	4.412576	.0488529	90.32	0.000	4.310671	4.514481

```
. reg lrpdi lsales i.id

      Source |       SS       df       MS              Number of obs =      28
-------------+------------------------------           F(  7,    20) =   84.81
       Model |  .108877836     7  .015553977           Prob > F      =  0.0000
    Residual |  .003668096    20  .000183405           R-squared     =  0.9674
-------------+------------------------------           Adj R-squared =  0.9560
       Total |  .112545933    27  .004168368           Root MSE      =  .01354

------------------------------------------------------------------------------
       lrpdi |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lsales |   .0042463    .011271     0.38   0.710    -.0192646    .0277572
             |
          id |
          2  |    .024285    .009697     2.50   0.021     .0040575    .0445125
          3  |   .0676848   .0097162     6.97   0.000     .0474172    .0879524
          4  |   .1245132   .0104325    11.94   0.000     .1027514     .146275
          5  |   .1285638   .0095564    13.45   0.000     .1086295    .1484982
          6  |   .1533331   .0097443    15.74   0.000     .1330068    .1736594
          7  |   .1775845   .0097406    18.23   0.000      .157266    .1979029
             |
       _cons |   4.416218   .0501929    87.98   0.000     4.311518    4.520919
------------------------------------------------------------------------------
```

Interpretation:
The most interesting part of these analyses is that when we account for individual and group effects, lsales has a significant impact on lrpdi. Comparing the two methods, we realize that the estimated parameters of the indicator variables are similar, and they are all significant at 5% level of significance. Therefore, estimating the fixed effect model with individual dummy variables and in clusters, the results are similar and significant. We present the STATA output in Appendix 3 and 4. Also, the programming codes are shown in Appendix 13

How do all of these results compare?
Using different approaches to estimate the fixed effect model, we realize that they generate the same results based on the estimated parameter. However, the most efficient method in this case is the fixed effect regression when we account for individual dummy variables or differences across groups. Note that these results could be as a result of the data we considered in this analysis. Both demean and robust standard errors approach give insignificant results.

Problem 2

Question a(i):

```
. egen E = max(grant), by(fcode)

.
. mean (hrsemp) if year==1987 & E==0

Mean estimation                    Number of obs    =      98

--------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
      hrsemp |   9.296744   1.663396      5.995363    12.59813
--------------------------------------------------------------

. mean (hrsemp) if year==1988 & E==0

Mean estimation                    Number of obs    =      96

--------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
      hrsemp |   9.671083   1.855213       5.98802    13.35415
--------------------------------------------------------------

. mean (hrsemp) if year==1987 & E==1

Mean estimation                    Number of obs    =      31

--------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
      hrsemp |   7.591087   3.650933      .1348872    15.04729
--------------------------------------------------------------

. mean (hrsemp) if year==1988 & E==1

Mean estimation                    Number of obs    =      31

--------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
      hrsemp |   35.97834   6.637497      22.42276    49.53392
--------------------------------------------------------------
```

```
. ttest hrsemp, by(E)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |     194    9.481984    1.241371     17.2903    7.033589    11.93038
       1 |      62    21.78471    4.172989    32.85815    13.44031    30.12912
---------+--------------------------------------------------------------------
combined |     256    12.46155    1.414802    22.63684    9.675366    15.24774
---------+--------------------------------------------------------------------
    diff |           -12.30273    3.217664               -18.63943   -5.966032
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  -3.8235
Ho: diff = 0                                     degrees of freedom =      254

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0001         Pr(|T| > |t|) = 0.0002          Pr(T > t) = 0.9999

. ttest hrsemp if year==1987, by(E)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
---------+--------------------------------------------------------------------
       0 |      98    9.296744    1.663396    16.46678    5.995363    12.59813
       1 |      31    7.591087    3.650933    20.32753    .1348872    15.04729
---------+--------------------------------------------------------------------
combined |     129    8.886858    1.532252    17.40304    5.855035    11.91868
---------+--------------------------------------------------------------------
    diff |            1.705658    3.597038               -5.412231    8.823546
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =   0.4742
Ho: diff = 0                                     degrees of freedom =      127

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.6819         Pr(|T| > |t|) = 0.6362          Pr(T > t) = 0.3181
```

```
. ttest hrsemp if year==1988, by(E)

Two-sample t test with equal variances
```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	96	9.671083	1.855213	18.1773	5.98802	13.35415
1	31	35.97834	6.637497	36.95602	22.42276	49.53392
combined	127	16.09254	2.352765	26.51432	11.43649	20.74859
diff		-26.30726	4.970328		-36.14415	-16.47036

```
    diff = mean(0) - mean(1)                                      t =  -5.2929
Ho: diff = 0                                     degrees of freedom =      125

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 0.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 1.0000
```

Interpretation:

It is evident from the result that when considering E=0 for 1987 and 1988 the mean values are quite close. However, for E=1 the mean values are far from each other. This effect could be as a result different number of observations. Similarly, for the diff-in-diff estimation, we find that when considering 1987 the result is insignificant based on the t-test. However, for other estimations, the results are significant at 5% level of significance.

Question a(ii):

```
. reg hrsemp grant d88 E
```

Source	SS	df	MS
Model	19608.6794	3	6536.22646
Residual	111060.074	252	440.714578
Total	130668.753	255	512.426483

```
Number of obs =       256
F(  3,    252) =    14.83
Prob > F      =    0.0000
R-squared     =    0.1501
Adj R-squared =    0.1399
Root MSE      =    20.993
```

hrsemp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
grant	28.01292	6.125444	4.57	0.000	15.94933	40.0765
d88	.374339	3.014608	0.12	0.901	-5.562698	6.311376
E	-1.705658	4.325932	-0.39	0.694	-10.22524	6.813929
_cons	9.296744	2.120634	4.38	0.000	5.120321	13.47317

Interpretation:

The result shows that grant has significant positive impact on hrsemp. Other variables in this model are insignificant to hrsemp. We present the STATA output in Appendix 5, 6, and 7. Also, the programming codes are shown in Appendix 14

Question a(iii):

```
. xtreg hrsemp grant d88, fe

Fixed-effects (within) regression               Number of obs      =       256
Group variable: fcode                           Number of groups   =       131

R-sq:  within  = 0.4711                         Obs per group: min =         1
       between = 0.0692                                        avg =       2.0
       overall = 0.1495                                        max =         2

                                                F(2,123)           =     54.77
corr(u_i, Xb)  = -0.0235                        Prob > F           =    0.0000

------------------------------------------------------------------------------
      hrsemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       grant |   27.87793   3.129216     8.91   0.000     21.68384    34.07202
         d88 |   .5093233   1.558337     0.33   0.744     -2.57531    3.593956
       _cons |   8.833036   .9462108     9.34   0.000      6.96007      10.706
-------------+----------------------------------------------------------------
     sigma_u |  19.405599
     sigma_e |  10.683421
         rho |  .76740875   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(130, 123) =     6.55             Prob > F = 0.0000
```

Interpretation:

The result shows that when estimating fixed effect model grant has significant positive impact on hrsemp. Other variables in the model are insignificant to hrsemp. We present the STATA output in Appendix 7. Also, the programming codes are shown in Appendix 14

Do you get the answer in each of these cases? Why or why not?

The results show that when we account for dummy variable for being a firm that receives treatment the estimated parameters of grant are close to the one in the fixed effect regression and they both significant. Generally, the results from two approaches are the same. The reason for this result is that the fixed effect model is exactly the same with the regression model since q_i and E_i are considered to be dummy variables in the two model.

Question 2b(i):

```
. xtreg hrsemp t, fe

Fixed-effects (within) regression               Number of obs      =       390
Group variable: fcode                           Number of groups   =       135

R-sq:  within  = 0.0872                         Obs per group: min =         1
       between = 0.0117                                        avg =       2.9
       overall = 0.0301                                        max =         3

                                                F(1,254)           =     24.27
corr(u_i, Xb)  = -0.0164                        Prob > F           =    0.0000

------------------------------------------------------------------------------
      hrsemp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   5.825537   1.182386     4.93   0.000     3.497008    8.154066
       _cons |   3.241778   2.566541     1.26   0.208    -1.812633    8.296189
-------------+----------------------------------------------------------------
     sigma_u |  20.071123
     sigma_e |  18.973517
         rho |  .52808937   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(134, 254) =     3.27            Prob > F = 0.0000

.
. predict hrsemp_res
(option xb assumed; fitted values)

. xtreg grant t, fe

Fixed-effects (within) regression               Number of obs      =       390
Group variable: fcode                           Number of groups   =       135

R-sq:  within  = 0.0774                         Obs per group: min =         1
       between = 0.0253                                        avg =       2.9
       overall = 0.0564                                        max =         3

                                                F(1,254)           =     21.31
corr(u_i, Xb)  = -0.0251                        Prob > F           =    0.0000

------------------------------------------------------------------------------
       grant |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   .1087379   .0235548     4.62   0.000     .0623502    .1551255
       _cons |  -.0675878   .0511292    -1.32   0.187    -.1682789    .0331034
-------------+----------------------------------------------------------------
     sigma_u |  .17032656
     sigma_e |  .37797988
         rho |  .16878723   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(134, 254) =     0.57            Prob > F = 0.9998

.
. predict grant_res
(option xb assumed; fitted values)
```

```
. xtreg d88 t, fe

Fixed-effects (within) regression               Number of obs      =       390
Group variable: fcode                           Number of groups   =       135

R-sq:  within  = 0.0000                         Obs per group: min =         1
       between = 0.3217                                        avg =       2.9
       overall = 0.0001                                        max =         3

                                                F(1,254)           =      0.00
corr(u_i, Xb)  = 0.0568                         Prob > F           =    0.9569

------------------------------------------------------------------------------
         d88 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |  -.0019417   .0358725    -0.05   0.957    -.0725872    .0687037
       _cons |   .3295494   .0778666     4.23   0.000     .1762031    .4828957
-------------+----------------------------------------------------------------
     sigma_u |  .08343789
     sigma_e |  .57563965
         rho |  .02057762   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(134, 254) =     0.03            Prob > F = 1.0000

.
. predict d88_res
(option xb assumed; fitted values)
. xtreg E t, fe
note: t omitted because of collinearity

Fixed-effects (within) regression               Number of obs      =       390
Group variable: fcode                           Number of groups   =       135

R-sq:  within  =      .                         Obs per group: min =         1
       between =      .                                        avg =       2.9
       overall =      .                                        max =         3

                                                F(1,254)           =         .
corr(u_i, Xb)  =      .                         Prob > F           =         .

------------------------------------------------------------------------------
           E |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |          0  (omitted)
       _cons |    .474359          .        .       .            .           .
-------------+----------------------------------------------------------------
     sigma_u |  .50074571
     sigma_e |          0
         rho |          1   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(134, 254) =        .            Prob > F =      .

.
. predict E_res
(option xb assumed; fitted values)
```

Question 2b(ii)

```
. reg hrsemp_res grant_res d88_res E_res
note: d88_res omitted because of collinearity
note: E_res omitted because of collinearity

      Source |       SS           df       MS      Number of obs   =       390
-------------+----------------------------------   F(1, 388)       =         .
       Model |  8923.22499         1  8923.22499   Prob > F        =         .
    Residual |           0       388           0   R-squared       =    1.0000
-------------+----------------------------------   Adj R-squared   =    1.0000
       Total |  8923.22499       389  22.9388817   Root MSE        =         0

------------------------------------------------------------------------------
  hrsemp_res |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   grant_res |   53.57414          .        .       .            .           .
     d88_res |          0  (omitted)
       E_res |          0  (omitted)
       _cons |   6.862733          .        .       .            .           .
------------------------------------------------------------------------------
```

Interpretation:

Generally, the result shows that when we considering firm-specific variables and trend for the three years, the most important residual is the one coming from grant. However, other residuals are not important to the firm. In conclusion, estimating fixed effect model using this approach is inefficient. We present the STATA output in Appendix 8,9,10,11, and 12. Also, the programming codes are shown in Appendix 14.