

```
-----
-----
name: <unnamed>
log: C:\Users\nw\Documents\HILDA Project\Training\Melbourne Nov
2011\Exercises\ex3.log
log type: text
opened on: 8 Nov 2011, 09:30:32
```

```
.
.
. *****
. * Load the panel data and tell Stata it is panel data *
. *****
.
. use "`working'\panelex3.dta" /*,replace*/

.
. xtset id wave
. panel variable: id (unbalanced)
. time variable: wave, 1 to 9, but with gaps
. delta: 1 unit

.
. *****
. * We are going to estimate the wage equation as a random *
. * effects model and compare the results to previous FE and *
. * and between estimates.
*
. * First, estimate the RE model. How important are the *
. * individual effects? Are they statistically significant *
. * (use a Breusch-Pagan test)?
*
. *****
. browse lwage_hr hgage agesq married female jbemp degree further ///
> tucov permanent nsw

.
.
. xtreg lwage_hr hgage agesq married female jbemp degree further ///
> tucov permanent nsw, re
```

```
Random-effects GLS regression                Number of obs      =
59755
Group variable: id                          Number of groups   =
13367

R-sq:  within = 0.1661                       Obs per group: min =
1                                             between = 0.2778   avg =
4.5                                         overall = 0.2515   max =
9

                                             Wald chi2(10)      =
11902.18
corr(u_i, X) = 0 (assumed)                   Prob > chi2        =
0.0000
```

```

-----
-----
      l wage_hr |      Coef.   Std. Err.      z    P>|z|      [95% Conf.
Interval]-----
-----
      hgage |      .0648374   .0014048    46.16   0.000    .0620842
.0675907
      agesq |     -.0638761   .0018107   -35.28   0.000    -.0674251  -
.0603271
      married |     .0053538   .0057606     0.93   0.353    -.0059368
.0166445
      female |    -.1096707   .0072999   -15.02   0.000    -.1239783  -
.0953631
      j bempt |     .0052257   .0003811    13.71   0.000    .0044788
.0059726
      degree |     .2170291   .0086203    25.18   0.000    .2001335
.2339246
      further |     .1264972   .0070027    18.06   0.000    .1127721
.1402223
      tucov |     .0198811   .004877     4.08   0.000    .0103224
.0294398
      permanent |     .0046802   .003976     1.18   0.239    -.0031125
.0124729
      nsw |     .0204337   .0070193     2.91   0.004    .006676
.0341913
      _cons |     1.432353   .0238173    60.14   0.000    1.385672
1.479034
-----
-----
      sigma_u |     .35287937
      sigma_e |     .30106476
      rho |     .57873966   (fraction of variance due to u_i)
-----
-----

```

```

. xttest0

```

Breusch and Pagan Lagrangian multiplier test for random effects

$$l\ wage\_hr[id,t] = Xb + u[id] + e[id,t]$$

Estimated results:

	Var	sd = sqrt(Var)
l wage_hr	.275198	.5245931
e	.09064	.3010648
u	.1245238	.3528794

Test: Var(u) = 0

chibar2(01) = 25947.48  
 Prob > chibar2 = 0.0000

```

. *****
. * Estimate the FE and between models to compare with RE(GLS) *

```

```

. * What do you notice?
*
. *****
.
. xtreg lwage_hr hgage agesq cohort married female jbemp degree further
///
>          tucov permanent nsw, fe
note: cohort omitted because of collinearity
note: female omitted because of collinearity

```

```

Fixed-effects (within) regression          Number of obs      =
59755                                     Number of groups   =
Group variable: id                        Number of groups   =
13367                                     Obs per group: min =
                                           avg =
R-sq:  within = 0.2412                    4.5                 max =
                                           9
                                           between = 0.1494
                                           overall = 0.1330
                                           F(9,46379)         =
1638.44                                   Prob > F            =
corr(u_i, Xb) = -0.8513
0.0000

```

```

-----
-----

```

	lwage_hr	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
	hgage	.1345099	.002121	63.42	0.000	.1303528
.1386671	agesq	-.0921478	.002585	-35.65	0.000	-.0972144 -
.0870812	cohort	(omitted)				
	married	-.0018379	.0069057	-0.27	0.790	-.0153731
.0116973	female	(omitted)				
	jbempt	.0012362	.0004347	2.84	0.004	.0003842
.0020883	degree	.0152046	.014657	1.04	0.300	-.0135233
.0439325	further	.062598	.0102969	6.08	0.000	.042416
.08278	tucov	.0275785	.0052327	5.27	0.000	.0173223
.0378348	permanent	-.0341137	.004087	-8.35	0.000	-.0421243 -
.0261032	nsw	.0252218	.0132444	1.90	0.057	-.0007373
.0511809	_cons	-.6066431	.040114	-15.12	0.000	-.6852671 -
.5280191						

```

-----
-----
sigma_u | .85089062

```

```

sigma_e | .30106476
rho | .88873817 (fraction of variance due to u_i)

```

```

-----
F test that all u_i=0:      F(13366, 46379) =      5.70      Prob > F =
0.0000

```

```

.
. xtreg lwage_hr hgage agesq cohort married female jbempt degree further
///
>          tucov permanent nsw, be

```

```

Between regression (regression on group means)  Number of obs      =
59755
Group variable: id                             Number of groups    =
13367

```

```

R-sq:  within = 0.1957      Obs per group: min =
1
      between = 0.3437      avg =
4.5
      overall = 0.3177      max =
9

```

```

635.69      F(11,13355)      =
sd(u_i + avg(e_i.))= .3961964      Prob > F      =
0.0000

```

```

-----
-----
lwage_hr |      Coef.   Std. Err.   t    P>|t|   [95% Conf.
Interval]
-----+-----
hgage |      .0921246   .0025491   36.14   0.000   .087128
.0971212
agesq |     -.0652708   .002319   -28.15   0.000   -.0698164   -
.0607253
cohort |      .0361822   .0016601   21.80   0.000   .0329282
.0394363
married |      .056993   .0089566    6.36   0.000   .0394368
.0745493
female |     -.1040327   .0069516  -14.97   0.000   -.1176588   -
.0904066
jbempt |      .0061516   .0006798    9.05   0.000   .0048191
.007484
degree |      .2739894   .0099528   27.53   0.000   .2544804
.2934984
further |      .0917343   .0087095   10.53   0.000   .0746624
.1088061
tucov |      .073737   .0099336    7.42   0.000   .0542657
.0932082
permanent |      .1064556   .0094627   11.25   0.000   .0879072
.1250039
nsw |      .0368062   .0076559    4.81   0.000   .0217995
.051813

```

```

      _cons | -70.89103   3.332567  -21.27   0.000   -77.42333   -
64.35873
-----
-----

```

```

.
. *****
. * Recall that the GLS estimator uses a weighted average of      *
. * between and within variation. The GLS transform consists of *
. * deviating variables from a fraction of their individual mean*
. *. Variables are transformed as devw(it)=w(it)-theta(i)*meanw *
. * (it) where theta(i)=1-sqroot[sigmasq_e/sigmasq_e+T(i)*      *
. * sigmasq_u]
. *
. * Note theta(i) is bounded between 0 and 1. If theta(i)=1, we *
. * have within estimator and if theta(0) we have the OLS      *
. * estimator(equal weight given to within and between          *
. * variation).
. *
. * What are the average theta values used in the wage eqn(use *
. * theta option)? what do you conclude?
. *
. * Would it be ok to use OLS instead of GLS? why (not)?      *
. *****
.

```

```

. xtreg lwage_hr hgage agesq cohort married female jbemp degree further
///
>          tucov permanent nsw, re theta

```

```

Random-effects GLS regression              Number of obs      =
59755                                     Number of groups     =
Group variable: id                        Number of groups     =
13367                                     Obs per group: min =
                                           1
                                           between = 0.3220          avg =
4.5                                       overall = 0.3156        max =
9
                                           Wald chi2(11)         =
20906.70
corr(u_i, X) = 0 (assumed)                Prob > chi2          =
0.0000

```

```

----- theta -----
   min      5%      median      95%      max
0.3424   0.3424   0.6000   0.7206   0.7206

```

```

-----
-----
      lwage_hr |          Coef.   Std. Err.      z    P>|z|      [95% Conf.
Interval]
-----+-----
      hgage |   .1193194   .0014617    81.63   0.000   .1164546
.1221843

```



corr(u\_i, Xb) = -0.8513 Prob > F =  
 0.0000

```
-----
```

	l wage_hr	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
	hgage	.1345099	.002121	63.42	0.000	.1303528
.1386671	agesq	-.0921478	.002585	-35.65	0.000	-.0972144 -
.0870812	cohort	(omitted)				
	married	-.0018379	.0069057	-0.27	0.790	-.0153731
.0116973	female	(omitted)				
	j bempt	.0012362	.0004347	2.84	0.004	.0003842
.0020883	degree	.0152046	.014657	1.04	0.300	-.0135233
.0439325	further	.062598	.0102969	6.08	0.000	.042416
.08278	tucov	.0275785	.0052327	5.27	0.000	.0173223
.0378348	permanent	-.0341137	.004087	-8.35	0.000	-.0421243 -
.0261032	nsw	.0252218	.0132444	1.90	0.057	-.0007373
.0511809	_cons	-.6066431	.040114	-15.12	0.000	-.6852671 -
.5280191						

```
-----
```

	sigma_u	.85089062				
	sigma_e	.30106476				
	rho	.88873817	(fraction of variance due to u_i)			

```
-----
```

F test that all u\_i=0: F(13366, 46379) = 5.70 Prob > F =  
 0.0000

. estimates store fixed //store results under name fixed

```
. xtreg l wage_hr hgage agesq cohort married female j bempt degree further
///
>          tucov permanent nsw, re
```

Random-effects GLS regression Number of obs =  
 59755  
 Group variable: id Number of groups =  
 13367

R-sq: within = 0.2351 Obs per group: min =  
 1 between = 0.3220 avg =  
 4.5

```

overall = 0.3156
max =
9
Wald chi2(11) =
20906.70
corr(u_i, X) = 0 (assumed)
Prob > chi2 =
0.0000

```

```

-----
-----
      lwage_hr |      Coef.   Std. Err.      z    P>|z|      [95% Conf.
Interval]
-----+-----
-----
      hgage |   .1193194   .0014617    81.63   0.000   .1164546
.1221843
      agesq |  -.0791819   .0017062   -46.41   0.000  -.0825261  -
.0758378
      cohort |   .0497645   .0005771    86.23   0.000   .0486334
.0508955
      married |   .0284287    .00542     5.25   0.000   .0178058
.0390517
      female |  -.1165087   .0067836   -17.17   0.000  -.1298044  -
.103213
      jbempt |   .0039446   .0003589    10.99   0.000   .0032411
.004648
      degree |   .2073784   .0080542    25.75   0.000   .1915925
.2231643
      further |   .0799398   .0065793    12.15   0.000   .0670446
.092835
      tucov |   .042346    .0046043     9.20   0.000   .0333217
.0513704
      permanent |  -.0131109   .0037594    -3.49   0.000  -.0204793  -
.0057426
      nsw |   .038013    .0065551     5.80   0.000   .0251652
.0508608
      _cons | -98.27032   1.156513   -84.97   0.000  -100.537  -
96.00359
-----+-----
-----
      sigma_u |   .34489511
      sigma_e |   .30106476
      rho |   .56754209   (fraction of variance due to u_i)
-----
-----

```

```

. estimates store random //store results under name random
.
. hausman fixed random

```

```

----- Coefficients -----
      |      (b)      (B)      (b-B)      sqrt(diag(V_b-
V_B))
      |      fixed      random      Difference      S.E.
-----+-----
-----

```

hgage		.1345099	.1193194	.0151905	.0015369
agesq		-.0921478	-.0791819	-.0129659	.0019419
married		-.0018379	.0284287	-.0302666	.0042793
jbempt		.0012362	.0039446	-.0027083	.0002453
degree		.0152046	.2073784	-.1921738	.0122457
further		.062598	.0799398	-.0173418	.0079207
tucov		.0275785	.042346	-.0147675	.0024863
permanent		-.0341137	-.0131109	-.0210028	.0016032
nsw		.0252218	.038013	-.0127912	.0115084

-----  
-----

b = consistent under Ho and Ha; obtained from  
xtreg  
B = inconsistent under Ha, efficient under Ho; obtained from  
xtreg

Test: Ho: difference in coefficients not systematic

chi2(9) = (b-B)'[(V\_b-V\_B)^(-1)](b-B)  
= 821.39  
Prob>chi2 = 0.0000

```
.
. *****
. * Now test the zero correlation using the mundlak formulation.*
. * Recall that this involves augmenting the RE model with the *
. * individual means of the time varying characteristics.      *
. *****
. drop mlwage mage mcohort myear //from ex2

.
. * Create individual means
. foreach x in lwage_hr hgage agesq cohort married female jbemp degree
further ///
>   tucov permanent nsw {
2.   egen m`x'=mean(`x'), by(id) //individual means
3. }

.
. xtreg lwage_hr hgage agesq cohort married female jbemp degree further
///
>   tucov permanent nsw ///
>   mhgage magesq mmarried mjbemp mdegree mfurther ///
>   mtucov mpermanent, re
```

Random-effects GLS regression	Number of obs	=
59755		
Group variable: id	Number of groups	=
13367		

R-sq: within = 0.2412	Obs per group: min =
1	
between = 0.3436	avg =
4.5	
overall = 0.3333	max =
9	

22018.41  
 corr(u\_i, X) = 0 (assumed)      Wald chi2(19) =  
 0.0000      Prob > chi2 =

```
-----
```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
hwage_hr	.134567	.0021078	63.84	0.000	.1304358
hagesq	-.0921945	.0025697	-35.88	0.000	-.097231 -
hcohort	.0354501	.0017543	20.21	0.000	.0320119
hmarried	-.0018382	.0068665	-0.27	0.789	-.0152963
hfemale	-.1036957	.0068014	-15.25	0.000	-.1170262 -
hjbempt	.0012281	.0004321	2.84	0.004	.0003811
hdegree	.0151351	.0145735	1.04	0.299	-.0134284
hfurther	.0625815	.0102384	6.11	0.000	.0425146
htucov	.02756	.005203	5.30	0.000	.0173624
hpermanent	-.0341191	.0040638	-8.40	0.000	-.042084 -
hns	.0346717	.0065135	5.32	0.000	.0219054
hmhgage	-.0448031	.0033545	-13.36	0.000	-.0513778 -
hmagesq	.0290456	.0034515	8.42	0.000	.0222808
hmmarried	.0621542	.0110843	5.61	0.000	.0404293
hmjbemp	.0048158	.0007873	6.12	0.000	.0032727
hmdegree	.2621503	.017461	15.01	0.000	.2279273
hmfurther	.0290278	.0133399	2.18	0.030	.0028822
hmtucov	.0422629	.0110555	3.82	0.000	.0205946
hmpermanent	.1389208	.0104154	13.34	0.000	.118507
hcons	-69.39042	3.521336	-19.71	0.000	-76.29211 -

```
-----
```

sigma_u	.34489511				
sigma_e	.30106476				
rho	.56754209	(fraction of variance due to u_i)			

```
-----
```

```
.
. test mhgage magesq mmarried mjbemp mdegree murtherer ///
>          mtucov mpermanent
```

- ( 1) mhgage = 0
- ( 2) magesq = 0
- ( 3) mmarried = 0
- ( 4) mjbemp = 0
- ( 5) mdegree = 0
- ( 6) murtherer = 0
- ( 7) mtucov = 0
- ( 8) mpermanent = 0

```
          chi2( 8) = 825.59
          Prob > chi2 = 0.0000
```

```
.
. *****
. * Based on the results of the two tests, what do you conclude *
. * about the correlation about the appropriateness of the RE *
. * model? what do the differences between FE and RE estimates *
. * say about possible omitted variable biases? *
. *****
```

\*

```
.
. *****
. *
. * Additional Exercise
. *
```

```
.
. *****
. * Manual RE estimation
. *
```

```
. xtreg lwage_hr hgage agesq cohort married female jbemp degree further
///
>          tucov permanent nsw, re
```

```
Random-effects GLS regression                Number of obs    =
59755
Group variable: id                          Number of groups  =
13367

R-sq:  within = 0.2351                       Obs per group: min =
1                                             between = 0.3220    avg =
4.5                                          overall = 0.3156    max =
9

                                             Wald chi2(11)      =
20906.70
corr(u_i, X) = 0 (assumed)                   Prob > chi2        =
0.0000
```

```
-----
-----
```

lwage_hr	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
hgage	.1193194	.0014617	81.63	0.000	.1164546
agesq	-.0791819	.0017062	-46.41	0.000	-.0825261 -
cohort	.0497645	.0005771	86.23	0.000	.0486334
married	.0284287	.00542	5.25	0.000	.0178058
female	-.1165087	.0067836	-17.17	0.000	-.1298044 -
jbempt	.0039446	.0003589	10.99	0.000	.0032411
degree	.2073784	.0080542	25.75	0.000	.1915925
further	.0799398	.0065793	12.15	0.000	.0670446
tucov	.042346	.0046043	9.20	0.000	.0333217
permanent	-.0131109	.0037594	-3.49	0.000	-.0204793 -
nsw	.038013	.0065551	5.80	0.000	.0251652
_cons	-98.27032	1.156513	-84.97	0.000	-100.537 -
sigma_u	.34489511				
sigma_e	.30106476				
rho	.56754209	(fraction of variance due to u_i)			

```

. *****
. * We are going to try to replicate the RE estimates manually *
. * using the GLS transform followed by OLS. Needs a bit of *
. * fiddly code! There are several steps:
. *
. * 1. Create a variable for T(i)
. *
. * 2. Calculate theta(i) using the appropriate scalars returned*
. * by Stata above.
. *
. * 3. Calculate the deviations for each variables (including *
. * constant!)
. *
. * 4. Do OLS regression using deviations (but no constant) *
. * 5. Compare to previous RE estimates
. *****
.
. by id: gen Ti=_N // the number of obs (by person)
.

```

```

. gen theta=1-sqrt(e(sigma_e)^2/(e(sigma_e)^2+ Ti*e(sigma_u)^2))

.
. drop devlwage devhgage devagesq devcohort devyear // from ex2

.
. foreach x in lwage_hr hgage agesq cohort married female jrbemp degree
further ///
>     tucov permanent nsw {
2.     gen dev`x'=`x'-theta*m`x' //deviation
3. }

. gen transconst=1-theta //transfomred constant

.
. reg devlwage_hr devhgage devagesq devcohort devmarried devfemale
devjrbemp ///
>     devdegree devfurther devtucov devpermanent devnsw transconst ///
>     , nocons

```

Source	SS	df	MS	Number of obs =
				59755
-----+-----				F( 12, 59743)
Model	69595.2176	12	5799.60147	Prob > F =
Residual	5427.04766	59743	.090839892	R-squared =
-----+-----				Adj R-squared =
Total	75022.2652	59755	1.2554977	Root MSE =
				.3014

devlwage_hr	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
devhgage	.1193194	.0014617	81.63	0.000	.1164545
devagesq	-.079182	.0017062	-46.41	0.000	-.0825262
devcohort	.0497644	.0005771	86.23	0.000	.0486334
devmarried	.0284288	.00542	5.25	0.000	.0178056
devfemale	-.1165089	.0067836	-17.17	0.000	-.1298048
devjrbemp	.0039446	.0003589	10.99	0.000	.0032411
devdegree	.2073785	.0080542	25.75	0.000	.1915923
devfurther	.0799398	.0065793	12.15	0.000	.0670443
devtucov	.0423461	.0046043	9.20	0.000	.0333216

```
devpermanent | -.0131109 .0037594 -3.49 0.000 -.0204794 -
.0057424
      devnsw | .038013 .0065551 5.80 0.000 .025165
.0508611
      transconst | -98.27026 1.156513 -84.97 0.000 -100.537 -
96.00349
```

-----  
-----

```
.
.
. *****
. *                               End of additional exercise                               *
. *****
.
. *****
. * Save the new data set for ex4
*
. *****
.
. save "`working'\panelex4.dta", replace
file C:\Users\nw\Documents\HILDA Project\Training\Melbourne Nov
2011\Exercises\panelex4.dta saved
```

```
.
. set more on

.
. log c
      name: <unnamed>
      log: C:\Users\nw\Documents\HILDA Project\Training\Melbourne Nov
2011\Exercises\ex3.log
      log type: text
      closed on: 8 Nov 2011, 09:30:38
```

-----  
-----  
-----