



```

-----
name: <unnamed>
log: C:\Users\afarou\Downloads\AbdullahFarouk.txt
log type: text
opened on: 30 Mar 2016, 13:08:00

```

```

1 .
2 . sysuse examdata1.dta
3 .
4 . *Question A
5 .
6 . summarize

```

Variable	Obs	Mean	Std. Dev.	Min	Max
id	54,876	9146.5	5280.494	1	18292
year	54,876	4	.816504	3	5
immigrant	54,876	.5258036	.4993383	0	1
married	54,876	.6543844	.4755729	0	1
schyears	54,876	14.16904	2.157363	10	18
cat	0				
cat1	54,876	.00687	.0826012	0	1
cat2	54,876	.1753772	.3802929	0	1
cat3	54,876	.1013922	.3018501	0	1
cat4	54,876	.6826846	.4654356	0	1
cat5	54,876	.0266966	.1611965	0	1
cat6	54,876	.0069794	.0832513	0	1
catnr	54,876	3.567899	.8559481	1	6
age	54,876	47.12224	11.51423	22	69
yten	54,876	14.01957	9.935795	1	49
nabs	54,876	1.63813	1.371656	0	5
dabs	54,876	2.600135	2.506181	0	19
wage	54,876	27.99927	14.87708	3.597399	114.3154

```

7 . gen lwage=log(wage)
8 . gen exp= (age-schyears-6)
9 . gen exp2=exp^2
10. regress lwage immigrant schyears exp exp2 if year==5, vce(robust)

```

```

Linear regression               Number of obs   =    18,292
                               F(4, 18287)        =   70308.21
                               Prob > F           =    0.0000
                               R-squared           =    0.9305
                               Root MSE        =    .14954

```

lwage	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
immigrant	-.1662029	.0022716	-73.17	0.000	-.1706554	-.1617504
schyears	.0621988	.0005249	118.49	0.000	.0611698	.0632277
exp	.1053557	.0004626	227.73	0.000	.1044489	.1062625
exp2	-.0011334	8.51e-06	-133.23	0.000	-.0011501	-.0011167
_cons	.6483137	.0091647	70.74	0.000	.6303501	.6662774

```

11.
12. *Immigrant is a dummy variable which indicates that an immigrant is likely to
> earn their non immigrant counter part. Its t stat is extremely high indicating
> that its coefficient is significant. Your wages are likely to go up by 6.2% f
> or every additional year of education and by 10.5% for every additional year o
> f experience. Notice however that after a certain level, experience has a nega
> tive effect(0.1%) indicated by the exp2 term. Note all these coefficients are
> statistically significant given the high t scores.
13.
14. *Question 2
15.
16. *lwage= exp exp2 immigrant Bcollar ImmB
17.
18. keep if year==5
    (36,584 observations deleted)
19. tabulate cat, gen(W)

```

cat	Freq.	Percent	Cum.
-----+			
bcollar1	503	2.75	2.75
bcollar2	113	0.62	3.37
manag1	130	0.71	4.08
manag2	3,297	18.02	22.10
manag3	1,833	10.02	32.12
wcollar	12,416	67.88	100.00
-----+			
Total	18,292	100.00	

```

20. generate Bcollar = W1+W2
21. generate office = W3+W4+W5+W6
22. generate ImmB = Bcollar*immigrant
23. generate ImmO = office*immigrant
24. estimate store tabulate
25. tabulate ImmB

```

ImmB	Freq.	Percent	Cum.
-----+			
0	17,984	98.32	98.32
1	308	1.68	100.00
-----+			
Total	18,292	100.00	

```

26. tabulate ImmO

```

ImmO	Freq.	Percent	Cum.
-----+			
0	8,982	49.10	49.10
1	9,310	50.90	100.00
-----+			
Total	18,292	100.00	

```

27. reg lwage exp exp2 immigrant Bcollar ImmB

```

Source	SS	df	MS	Number of obs	=	18,292
-----+				F(5, 18286)	=	27859.80
Model	5202.37717	5	1040.47543	Prob > F	=	0.0000
Residual	682.924249	18,286	.037346836	R-squared	=	0.8840
-----+				Adj R-squared	=	0.8839
Total	5885.30142	18,291	.321759413	Root MSE	=	.19325

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
exp	.1123617	.0006399	175.58	0.000	.1111073	.113616
exp2	-.0012622	.0000114	-110.50	0.000	-.0012846	-.0012398
immigrant	-.2259165	.0029221	-77.31	0.000	-.231644	-.2201889
Bcollar	-.2818831	.0112359	-25.09	0.000	-.3039065	-.2598598
ImmB	.1012171	.0158439	6.39	0.000	.0701616	.1322727
_cons	1.489959	.0084054	177.26	0.000	1.473484	1.506434

28. estimate store BlueI

29.

30. reg lwage exp exp2 immigrant office ImmO

Source	SS	df	MS	Number of obs	=	18,292
Model	5202.37717	5	1040.47543	F(5, 18286)	=	27859.80
Residual	682.924249	18,286	.037346836	Prob > F	=	0.0000
				R-squared	=	0.8840
				Adj R-squared	=	0.8839
Total	5885.30142	18,291	.321759413	Root MSE	=	.19325

lwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
exp	.1123617	.0006399	175.58	0.000	.1111073	.113616
exp2	-.0012622	.0000114	-110.50	0.000	-.0012846	-.0012398
immigrant	-.1246993	.0155733	-8.01	0.000	-.1552245	-.0941742
office	.2818831	.0112359	25.09	0.000	.2598598	.3039065
ImmO	-.1012171	.0158439	-6.39	0.000	-.1322727	-.0701616
_cons	1.208076	.01331	90.76	0.000	1.181987	1.234165

31. estimate store OfficeI

32.

33. estimate table OfficeI BlueI

Variable	OfficeI	BlueI
exp	.11236167	.11236167
exp2	-.00126222	-.00126222
immigrant	-.12469934	-.22591648
office	.28188314	
ImmO	-.10121714	
Bcollar		-.28188314
ImmB		.10121714
_cons	1.2080759	1.489959

34.

35. *the t stats for all the explanatory variable are high indicating that our results are statistically significant.

36. *Being a blue immigrant (whose coefficient is given by the sum of the coefficients of the immigrant, blue collar and the interaction dummy term) suggest a negative effect of being a blue collar immigrant worker. on the other hand being an officeimmigrant your wages are likely to go up. This is consistent with the fact that there are more whitecollar immigrant workers than there are blue collar ones. Hence we observe white

```

37.
38. *Question 3
39. Clear
    program clear already defined
    (error occurred while loading Clear.ado)
    r(110);

    end of do-file

    r(110);

40. do "C:\Users\afarou\AppData\Local\Temp\STD00000000.tmp"

41. clear

42.
43. sysuse examdata1.dta

44. gen lwage=log(wage)

45. gen exp= (age-schyears-6)

46. gen exp2=exp^2

47. tabulate cat, gen(W)

```

cat	Freq.	Percent	Cum.
-----+-----			
bcollar1	1,465	2.67	2.67
bcollar2	383	0.70	3.37
manag1	377	0.69	4.05
manag2	9,624	17.54	21.59
manag3	5,564	10.14	31.73
wcollar	37,463	68.27	100.00
-----+-----			
Total	54,876	100.00	

```

48. generate Bcollar = W1+W2

49. generate office = W3+W4+W5+W6

50. generate ImmB = Bcollar*immigrant

51. generate ImmO = office*immigrant

52. estimate store tabulate

53. xtset id year
    panel variable: id (strongly balanced)
    time variable: year, 3 to 5
    delta: 1 unit

54.
55. xtreg lwage exp exp2 immigrant Bcollar office, fe vce(cluster id)
    note: immigrant omitted because of collinearity
    note: Bcollar omitted because of collinearity
    note: office omitted because of collinearity

```

Fixed-effects (within) regression	Number of obs	=	54,876
Group variable: id	Number of groups	=	18,292
R-sq:	Obs per group:		
within = 0.7009	min =		3
between = 0.7337	avg =		3.0
overall = 0.6674	max =		3
corr(u_i, Xb) = -0.9875	F(2,18291)	=	50475.42
	Prob > F	=	0.0000

(Std. Err. adjusted for 18,292 clusters in id)

lwage	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
exp	.1484698	.0016335	90.89	0.000	.145268	.1516716
exp2	.0012956	.0000312	41.54	0.000	.0012344	.0013567
immigrant	0	(omitted)				
Bcollar	0	(omitted)				
office	0	(omitted)				
_cons	-1.931057	.022525	-85.73	0.000	-1.975208	-1.886906
sigma_u	2.0052595					
sigma_e	.14387637					
rho	.99487838	(fraction of variance due to u_i)				

56. estimates store FE

57. predict alpha, u

58. summarize alpha

Variable	Obs	Mean	Std. Dev.	Min	Max
alpha	54,876	-1.50e-09	2.005223	-6.030703	3.406896

59.

60. xtsum

Variable		Mean	Std. Dev.	Min	Max	Observations	
id	overall	9146.5	5280.494	1	18292	N =	54876
	between		5280.59	1	18292	n =	18292
	within		0	9146.5	9146.5	T =	3
year	overall	4	.816504	3	5	N =	54876
	between		0	4	4	n =	18292
	within		.816504	3	5	T =	3
immigr~t	overall	.5258036	.4993383	0	1	N =	54876
	between		.4993474	0	1	n =	18292
	within		0	.5258036	.5258036	T =	3
married	overall	.6543844	.4755729	0	1	N =	54876
	between		.4755816	0	1	n =	18292
	within		0	.6543844	.6543844	T =	3
schyears	overall	14.16904	2.157363	10	18	N =	54876
	between		2.157402	10	18	n =	18292
	within		0	14.16904	14.16904	T =	3
cat	overall	N =	0
	between		.	.	.	n =	0
	within		.	.	.	T =	.
cat1	overall	.00687	.0826012	0	1	N =	54876
	between		.0819382	0	1	n =	18292
	within		.0104565	-.6597966	.6735367	T =	3
cat2	overall	.1753772	.3802929	0	1	N =	54876
	between		.3771077	0	1	n =	18292
	within		.0491693	-.4912895	.8420439	T =	3
cat3	overall	.1013922	.3018501	0	1	N =	54876
	between		.3009487	0	1	n =	18292
	within		.0233816	-.5652744	.7680589	T =	3
cat4	overall	.6826846	.4654356	0	1	N =	54876
	between		.4635478	0	1	n =	18292
	within		.0419712	.0160179	1.349351	T =	3

cat5	overall	.0266966	.1611965	0	1	N =	54876
	between		.160065	0	1	n =	18292
	within		.019091	-.6399701	.6933632	T =	3
cat6	overall	.0069794	.0832513	0	1	N =	54876
	between		.0810343	0	1	n =	18292
	within		.019091	-.6596873	.673646	T =	3
catnr	overall	3.567899	.8559481	1	6	N =	54876
	between		.8512383	1	6	n =	18292
	within		.0898155	2.234565	4.901232	T =	3
age	overall	47.12224	11.51423	22	69	N =	54876
	between		11.48545	23	68	n =	18292
	within		.816504	46.12224	48.12224	T =	3
yten	overall	14.01957	9.935795	1	49	N =	54876
	between		9.902369	2	48	n =	18292
	within		.816504	13.01957	15.01957	T =	3
nabs	overall	1.63813	1.371656	0	5	N =	54876
	between		1.307631	0	5	n =	18292
	within		.4142489	-.0285371	3.304796	T =	3
dabs	overall	2.600135	2.506181	0	19	N =	54876
	between		2.195935	0	11.66667	n =	18292
	within		1.207886	-4.066532	12.2668	T =	3
wage	overall	27.99927	14.87708	3.597399	114.3154	N =	54876
	between		13.19305	4.540222	90.39766	n =	18292
	within		6.87585	-15.83267	62.11433	T =	3
lwage	overall	3.1773	.5841081	1.280211	4.738961	N =	54876
	between		.5431862	1.509025	4.441627	n =	18292
	within		.2148064	2.36152	3.763917	T =	3
exp	overall	26.9532	11.29999	2	53	N =	54876
	between		11.27066	3	52	n =	18292
	within		.816504	25.9532	27.9532	T =	3
exp2	overall	854.1627	610.913	4	2809	N =	54876
	between		609.0582	9.666667	2704.667	n =	18292
	within		47.7101	750.4961	958.4961	T =	3
W1	overall	.0266966	.1611965	0	1	N =	54876
	between		.160065	0	1	n =	18292
	within		.019091	-.6399701	.6933632	T =	3
W2	overall	.0069794	.0832513	0	1	N =	54876
	between		.0810343	0	1	n =	18292
	within		.019091	-.6596873	.673646	T =	3
W3	overall	.00687	.0826012	0	1	N =	54876
	between		.0819382	0	1	n =	18292
	within		.0104565	-.6597966	.6735367	T =	3
W4	overall	.1753772	.3802929	0	1	N =	54876
	between		.3771077	0	1	n =	18292
	within		.0491693	-.4912895	.8420439	T =	3
W5	overall	.1013922	.3018501	0	1	N =	54876
	between		.3009487	0	1	n =	18292
	within		.0233816	-.5652744	.7680589	T =	3
W6	overall	.6826846	.4654356	0	1	N =	54876
	between		.4635478	0	1	n =	18292
	within		.0419712	.0160179	1.349351	T =	3
Bcollar	overall	.0336759	.1803953	0	1	N =	54876
	between		.1803985	0	1	n =	18292
	within		0	.0336759	.0336759	T =	3

office	overall	.9663241	.1803953	0	1	N =	54876
	between		.1803985	0	1	n =	18292
	within		0	.9663241	.9663241	T =	3
ImmB	overall	.016838	.1286653	0	1	N =	54876
	between		.1286676	0	1	n =	18292
	within		0	.016838	.016838	T =	3
ImmO	overall	.5089657	.4999242	0	1	N =	54876
	between		.4999333	0	1	n =	18292
	within		0	.5089657	.5089657	T =	3
_est_FE	overall	1	0	1	1	N =	54876
	between		0	1	1	n =	18292
	within		0	1	1	T =	3
alpha	overall	-1.50e-09	2.005223	-6.030703	3.406896	N =	54876
	between		2.00526	-6.030703	3.406896	n =	18292
	within		0	-1.50e-09	-1.50e-09	T =	3

61.

62. *We cant control for immigrant status as it is time invariant.

63. *We cant control for age nor tenure as they increase by i unit of time and the
 > refore when including this with exp and exp2 which also change in the same fas
 > hion and as a result we will have perfect collinearity.

64.

65.

66. *Question 4

67. xtreg lwage exp exp2 immigrant Bcollar office, re
 note: office omitted because of collinearity

Random-effects GLS regression
 Group variable: id

Number of obs = 54,876
 Number of groups = 18,292

R-sq:

within = 0.4581
 between = 0.9130
 overall = 0.8184

Obs per group:

min = 3
 avg = 3.0
 max = 3

corr(u_i, X) = 0 (assumed)

Wald chi2(4) = 108860.20
 Prob > chi2 = 0.0000

lwage	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
exp	.1009583	.0006979	144.66	0.000	.0995904 .1023262
exp2	-.0010567	.0000129	-81.98	0.000	-.0010819 -.0010314
immigrant	-.2255575	.0033626	-67.08	0.000	-.2321481 -.2189669
Bcollar	-.2164731	.0093083	-23.26	0.000	-.234717 -.1982291
office	0 (omitted)				
_cons	1.484617	.0088886	167.03	0.000	1.467195 1.502038
sigma_u	.13237998				
sigma_e	.14387637				
rho	.458457	(fraction of variance due to u_i)			

68. estimates store RE

69. xtreg lwage exp exp2 immigrant Bcollar office, re vce(robust)
 note: office omitted because of collinearity

Random-effects GLS regression
 Group variable: id

Number of obs = 54,876
 Number of groups = 18,292

R-sq:

within = 0.4581
 between = 0.9130
 overall = 0.8184

Obs per group:

min = 3
 avg = 3.0
 max = 3

corr(u_i, X) = 0 (assumed) Wald chi2(4) = 270746.22
 Prob > chi2 = 0.0000

(Std. Err. adjusted for 18,292 clusters in id)

lwage	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
exp	.1009583	.0004643	217.46	0.000	.1000484	.1018682
exp2	-.0010567	9.02e-06	-117.15	0.000	-.0010744	-.001039
immigrant	-.2255575	.0024322	-92.74	0.000	-.2303245	-.2207905
Bcollar	-.2164731	.0056458	-38.34	0.000	-.2275386	-.2054075
office	0	(omitted)				
_cons	1.484617	.0052881	280.75	0.000	1.474252	1.494981
sigma_u	.13237998					
sigma_e	.14387637					
rho	.458457	(fraction of variance due to u_i)				

70.

71. regress lwage exp exp2 immigrant Bcollar office, vce(cluster id)
 note: Bcollar omitted because of collinearity

Linear regression Number of obs = 54,876
 F(4, 18291) = 67312.76
 Prob > F = 0.0000
 R-squared = 0.8202
 Root MSE = .24767

(Std. Err. adjusted for 18,292 clusters in id)

lwage	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
exp	.1063136	.0004389	242.25	0.000	.1054534	.1071738
exp2	-.0012017	8.38e-06	-143.35	0.000	-.0012181	-.0011853
immigrant	-.226061	.0023468	-96.33	0.000	-.2306609	-.221461
Bcollar	0	(omitted)				
office	.2066794	.0052372	39.46	0.000	.1964141	.2169448
_cons	1.257411	.0071113	176.82	0.000	1.243472	1.27135

72. xttest0

last estimates not found
 r(301);

end of do-file

r(301);

73. do "C:\Users\afarou\AppData\Local\Temp\STD00000000.tmp"

74. xttest0

last estimates not found
 r(301);

end of do-file

r(301);

75. do "C:\Users\afarou\AppData\Local\Temp\STD00000000.tmp"

76. xtreg lwage exp exp2 immigrant Bcollar office, re
note: office omitted because of collinearity

Random-effects GLS regression	Number of obs	=	54,876
Group variable: id	Number of groups	=	18,292
R-sq:			
within	=	0.4581	
between	=	0.9130	
overall	=	0.8184	
Obs per group:			
	min	=	3
	avg	=	3.0
	max	=	3
Wald chi2(4) = 108860.20			
corr(u_i, X) = 0 (assumed)	Prob > chi2	=	0.0000

	lwage	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
exp		.1009583	.0006979	144.66	0.000	.0995904 .1023262
exp2		-.0010567	.0000129	-81.98	0.000	-.0010819 -.0010314
immigrant		-.2255575	.0033626	-67.08	0.000	-.2321481 -.2189669
Bcollar		-.2164731	.0093083	-23.26	0.000	-.234717 -.1982291
office		0 (omitted)				
_cons		1.484617	.0088886	167.03	0.000	1.467195 1.502038
sigma_u		.13237998				
sigma_e		.14387637				
rho		.458457	(fraction of variance due to u_i)			

77.
end of do-file

78. do "C:\Users\afarou\AppData\Local\Temp\STD00000000.tmp"

79. xttest0

Breusch and Pagan Lagrangian multiplier test for random effects

lwage[id,t] = Xb + u[id] + e[id,t]

Estimated results:

	Var	sd = sqrt(Var)
lwage	.3411822	.5841081
e	.0207004	.1438764
u	.0175245	.13238

Test: Var(u) = 0

chibar2(01) = 563.27
Prob > chibar2 = 0.0000

80.
end of do-file

81. do "C:\Users\afarou\AppData\Local\Temp\STD00000000.tmp"

82. xtreg lwage exp exp2 immigrant Bcollar office, re vce(robust)
note: office omitted because of collinearity

Random-effects GLS regression	Number of obs	=	54,876
Group variable: id	Number of groups	=	18,292
R-sq:			
within	=	0.4581	
between	=	0.9130	
overall	=	0.8184	
Obs per group:			
	min	=	3
	avg	=	3.0
	max	=	3
Wald chi2(4) = 270746.22			
corr(u_i, X) = 0 (assumed)	Prob > chi2	=	0.0000

(Std. Err. adjusted for 18,292 clusters in id)

lwage	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
exp	.1009583	.0004643	217.46	0.000	.1000484	.1018682
exp2	-.0010567	9.02e-06	-117.15	0.000	-.0010744	-.001039
immigrant	-.2255575	.0024322	-92.74	0.000	-.2303245	-.2207905
Bcollar	-.2164731	.0056458	-38.34	0.000	-.2275386	-.2054075
office	0	(omitted)				
_cons	1.484617	.0052881	280.75	0.000	1.474252	1.494981
sigma_u	.13237998					
sigma_e	.14387637					
rho	.458457	(fraction of variance due to u_i)				

83.

84. regress lwage exp exp2 immigrant Bcollar office, vce(cluster id)
note: Bcollar omitted because of collinearity

Linear regression	Number of obs	=	54,876
	F(4, 18291)	=	67312.76
	Prob > F	=	0.0000
	R-squared	=	0.8202
	Root MSE	=	.24767

(Std. Err. adjusted for 18,292 clusters in id)

lwage	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
exp	.1063136	.0004389	242.25	0.000	.1054534	.1071738
exp2	-.0012017	8.38e-06	-143.35	0.000	-.0012181	-.0011853
immigrant	-.226061	.0023468	-96.33	0.000	-.2306609	-.221461
Bcollar	0	(omitted)				
office	.2066794	.0052372	39.46	0.000	.1964141	.2169448
_cons	1.257411	.0071113	176.82	0.000	1.243472	1.27135

85.

end of do-file

86. do "C:\Users\afarou\AppData\Local\Temp\STD00000000.tmp"

87.

88. * A positive variation of the fixed effects term should result in

89. *The BP test gives us a statistically significant chi squared value in which
> case we reject the null and thus conclude that RE is more efficient, or put in
> another way the fixed effect has a positive variation.

90.

91. xi: xtreg lwage exp exp2 immigrant Bcollar office i.year*immigrant, re
i.year _Iyear_3-5 (naturally coded; _Iyear_3 omitted)
i.year*immigr~t _IyeaXimmig_# (coded as above)
note: office omitted because of collinearity
note: immigrant omitted because of collinearity

Random-effects GLS regression	Number of obs	=	54,876
Group variable: id	Number of groups	=	18,292

R-sq:	Obs per group:
within = 0.7230	min = 3
between = 0.9145	avg = 3.0
overall = 0.8885	max = 3

corr(u_i, X)	= 0 (assumed)	Wald chi2(8)	= 277829.76
		Prob > chi2	= 0.0000

lwage	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
exp	.0962684	.0005083	189.40	0.000	.0952722	.0972647
exp2	-.0010294	9.38e-06	-109.72	0.000	-.0010478	-.001011
immigrant	-.2282286	.0029672	-76.92	0.000	-.2340441	-.2224131
Bcollar	-.2166601	.006839	-31.68	0.000	-.2300642	-.203256
office	0	(omitted)				
_Iyear_4	.2921329	.0020636	141.56	0.000	.2880883	.2961775
_Iyear_5	.3561781	.0020727	171.84	0.000	.3521157	.3602405
immigrant	0	(omitted)				
_IyeaXimmig_4	-.003334	.0028419	-1.17	0.241	-.008904	.0022361
_IyeaXimmig_5	-.0021126	.0028422	-0.74	0.457	-.0076832	.003458
_cons	1.374014	.0065707	209.11	0.000	1.361136	1.386892
sigma_u	.13784429					
sigma_e	.12755641					
rho	.5387054	(fraction of variance due to u_i)				

92.

93. log close

name: <unnamed>

log: C:\Users\afarou\Downloads\AbdullahFarouk.txt

log type: text

closed on: 30 Mar 2016, 13:08:48