(R)
Statistics/Data Analysis

User: Nond Prueksiri
Project: Assignment 5

(R)
Statistics/Data Analysis **15.0**
Special Edition

Copyright 1985-2017 StataCorp LLC
StataCorp
4905 Lakeway Drive
College Station, Texas 77845 USA
800-STATA-PC http://www.stata.com
979-696-4600 stata@stata.com
979-696-4601 (fax)

1150-user Stata network license expires 23 Jul 2019:
Serial number: 401509213955
Licensed to: Nond Prueksiri
Duke University

Notes:
1. Unicode is supported; see help unicode_advice.
2. Maximum number of variables is set to 5000; see help set_maxvar.

```
 1 . do "C:\Users\NONDP~1\AppData\Local\Temp\STDd40_000000.tmp"

 2 . clear all

 3 . set more off, perm
   (set more preference recorded)

 4 . set scrollbufsize 2000000
   (set scrollbufsize will take effect the next time you launch Stata)

 5 . set obs 10000
   number of observations (_N) was 0, now 10,000

 6 . set seed 12345

 7 .
 8 . * REDO: Assignment 2
 9 . * Exercise 1 Data Creation
10 .
11 . gen x1 = runiform(1,3)

12 . gen x2 = rgamma(3,2)

13 . gen x3 = rnbinomial(10000,0.3)

14 . gen eps = rnormal(2,1)

15 . gen y = 0.5 + 1.2*x1 - 0.9*x2 + 0.1*x3 + eps

16 . gen ydum = 0

17 . egen mean_y = mean(y)
```

```
18 . replace ydum = 1 if y > mean_y
   (4,981 real changes made)

19 .
20 . * Exercise 2 OLS
21 . corr y x1
   (obs=10,000)
```

	y	x1
y	**1.0000**	
x1	**0.0256**	**1.0000**

```
22 . reg y x1 x2 x3
```

Source	SS	df	MS		
Model	**7849360.69**	**3**	**2616453.56**	Prob > F	**0.0000**
Residual	**9980.08148**	**9,996**	**.998407511**	R-squared	**0.9987**
				Adj R-squared	**0.9987**
Total	**7859340.77**	**9,999**	**786.012679**	Root MSE	**.9992**

Number of obs	=	**10,000**			
F(3, 9996)	>	**99999.00**			

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
x1	**1.195803**	**.0173609**	**68.88**	**0.000**	**1.161772**	**1.229834**
x2	**-.9019605**	**.0028824**	**-312.92**	**0.000**	**-.9076106**	**-.8963105**
x3	**.0999977**	**.0000359**	**2788.29**	**0.000**	**.0999274**	**.100068**
_cons	**2.569322**	**.8376162**	**3.07**	**0.002**	**.9274259**	**4.211219**

```
23 . bootstrap, reps(49) seed(12345) : reg y x1 x2 x3
   (running regress on estimation sample)

   Bootstrap replications (49)
   ----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
   .................................................

   Linear regression                               Number of obs     =     10,000
                                                   Replications      =         49
                                                   Wald chi2(3)      =  8336991.55
                                                   Prob > chi2       =     0.0000
                                                   R-squared         =     0.9987
                                                   Adj R-squared     =     0.9987
                                                   Root MSE          =     0.9992
```

y	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
x1	**1.195803**	**.0199619**	**59.90**	**0.000**	**1.156679**	**1.234928**
x2	**-.9019605**	**.0031712**	**-284.43**	**0.000**	**-.9081759**	**-.8957452**
x3	**.0999977**	**.0000347**	**2885.49**	**0.000**	**.0999298**	**.1000657**
_cons	**2.569322**	**.8118018**	**3.16**	**0.002**	**.9782199**	**4.160425**

```
24 . bootstrap, reps(499) seed(12345) : reg y x1 x2 x3
   (running regress on estimation sample)

   Bootstrap replications (499)
   ----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
   .................................................    50
   .................................................   100
   .................................................   150
   .................................................   200
   .................................................   250
   .................................................   300
   .................................................   350
   .................................................   400
   .................................................   450
   .................................................
```

Linear regression

Number of obs	**10,000**
Replications	**499**
Wald chi2(**3**)	**7712839.19**
Prob > chi2	**0.0000**
R-squared	**0.9987**
Adj R-squared	**0.9987**
Root MSE	**0.9992**

y	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
x1	**1.195803**	**.017773**	**67.28**	**0.000**	**1.160969**	**1.230638**
x2	**-.9019605**	**.0028753**	**-313.69**	**0.000**	**-.907596**	**-.896325**
x3	**.0999977**	**.0000363**	**2751.71**	**0.000**	**.0999265**	**.100069**
_cons	**2.569322**	**.848007**	**3.03**	**0.002**	**.9072591**	**4.231385**

```
25 .
26 . * Exercise 3 Probit
27 . probit ydum x1 x2 x3

   Iteration 0:   log likelihood = -6931.3996
   Iteration 1:   log likelihood = -287.39437
   Iteration 2:   log likelihood =  -262.3666
   Iteration 3:   log likelihood = -261.19541
   Iteration 4:   log likelihood = -261.19416
   Iteration 5:   log likelihood = -261.19416
```

Probit regression

Number of obs	**10,000**
LR chi2(**3**)	**13340.41**
Prob > chi2	**0.0000**
Log likelihood = **-261.19416**	Pseudo R2 = **0.9623**

ydum	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
x1	**1.310577**	**.131925**	**9.93**	**0.000**	**1.052008**	**1.569145**
x2	**-.9257078**	**.0553151**	**-16.74**	**0.000**	**-1.034123**	**-.8172922**
x3	**.1013994**	**.0057763**	**17.55**	**0.000**	**.0900781**	**.1127207**
_cons	**-2363.215**	**134.6236**	**-17.55**	**0.000**	**-2627.073**	**-2099.358**

Note: 4182 failures and 4167 successes completely determined.

```
28 .
29 . * Exercise 4 Discrete Choice
30 . probit ydum x1 x2 x3

Iteration 0:   log likelihood = -6931.3996
Iteration 1:   log likelihood = -287.39437
Iteration 2:   log likelihood =  -262.3666
Iteration 3:   log likelihood = -261.19541
Iteration 4:   log likelihood = -261.19416
Iteration 5:   log likelihood = -261.19416

Probit regression                                 Number of obs     =     10,000
                                                  LR chi2(3)        =   13340.41
                                                  Prob > chi2       =     0.0000
Log likelihood = -261.19416                       Pseudo R2         =     0.9623

------------------------------------------------------------------------------
        ydum |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |    1.310577    .131925     9.93   0.000     1.052008    1.569145
          x2 |   -.9257078   .0553151   -16.74   0.000    -1.034123   -.8172922
          x3 |    .1013994   .0057763    17.55   0.000     .0900781    .1127207
       _cons |   -2363.215   134.6236   -17.55   0.000    -2627.073   -2099.358
------------------------------------------------------------------------------
Note: 4182 failures and 4167 successes completely determined.

31 . logit ydum x1 x2 x3

Iteration 0:   log likelihood = -6931.3996
Iteration 1:   log likelihood =   -280.023
Iteration 2:   log likelihood = -264.11024
Iteration 3:   log likelihood = -262.63364
Iteration 4:   log likelihood = -262.63171
Iteration 5:   log likelihood = -262.63171

Logistic regression                               Number of obs     =     10,000
                                                  LR chi2(3)        =   13337.54
                                                  Prob > chi2       =     0.0000
Log likelihood = -262.63171                       Pseudo R2         =     0.9621

------------------------------------------------------------------------------
        ydum |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |    2.375338   .2454155     9.68   0.000     1.894332    2.856343
          x2 |   -1.672413   .1088389   -15.37   0.000    -1.885734   -1.459093
          x3 |    .1828114   .0114137    16.02   0.000     .1604409    .2051819
       _cons |   -4260.635   266.0159   -16.02   0.000    -4782.016   -3739.253
------------------------------------------------------------------------------
Note: 3632 failures and 3646 successes completely determined.

32 . reg ydum x1 x2 x3

      Source |       SS           df       MS      Number of obs   =     10,000
-------------+----------------------------------   F(3, 9996)      =   5906.90
       Model |  1598.35475         3  532.784917   Prob > F        =     0.0000
    Residual |   901.60915     9,996  .090196994   R-squared       =     0.6394
-------------+----------------------------------   Adj R-squared   =     0.6392
       Total |  2499.9639      9,999  .250021392   Root MSE        =     .30033
```

ydum	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
x1	.0224658	.0052181	4.31	0.000	.0122372	.0326944
x2	-.0144045	.0008664	-16.63	0.000	-.0161028	-.0127063
x3	.0014245	.0000108	132.15	0.000	.0014034	.0014457
_cons	-32.70141	.2517603	-129.89	0.000	-33.19491	-32.20791

```
33 .
34 . * Exercise 5 Marginal Effects
35 . probit ydum x1 x2 x3

Iteration 0:   log likelihood = -6931.3996
Iteration 1:   log likelihood = -287.39437
Iteration 2:   log likelihood =  -262.3666
Iteration 3:   log likelihood = -261.19541
Iteration 4:   log likelihood = -261.19416
Iteration 5:   log likelihood = -261.19416

Probit regression                                 Number of obs     =     10,000
                                                  LR chi2(3)        =   13340.41
                                                  Prob > chi2       =     0.0000
Log likelihood = -261.19416                       Pseudo R2         =     0.9623
```

ydum	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
x1	1.310577	.131925	9.93	0.000	1.052008	1.569145
x2	-.9257078	.0553151	-16.74	0.000	-1.034123	-.8172922
x3	.1013994	.0057763	17.55	0.000	.0900781	.1127207
_cons	-2363.215	134.6236	-17.55	0.000	-2627.073	-2099.358

```
Note: 4182 failures and 4167 successes completely determined.

36 . margins, dydx(*)

Average marginal effects                          Number of obs     =     10,000
Model VCE    : OIM

Expression   : Pr(ydum), predict()
dy/dx w.r.t. : x1 x2 x3
```

	dy/dx	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
x1	.0189197	.0015744	12.02	0.000	.0158339	.0220056
x2	-.0133637	.0002488	-53.72	0.000	-.0138513	-.0128761
x3	.0014638	3.09e-06	473.46	0.000	.0014578	.0014699

```
37 . logit ydum x1 x2 x3

Iteration 0:   log likelihood = -6931.3996
Iteration 1:   log likelihood =   -280.023
Iteration 2:   log likelihood = -264.11024
Iteration 3:   log likelihood = -262.63364
Iteration 4:   log likelihood = -262.63171
Iteration 5:   log likelihood = -262.63171
```

Logistic regression

Number of obs	=	**10,000**			
LR chi2(**3**)	=	**13337.54**			
Prob > chi2	=	**0.0000**			
Log likelihood = **-262.63171**	Pseudo R2	=	**0.9621**		

ydum	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
x1	**2.375338**	**.2454155**	**9.68**	**0.000**	**1.894332**	**2.856343**
x2	**-1.672413**	**.1088389**	**-15.37**	**0.000**	**-1.885734**	**-1.459093**
x3	**.1828114**	**.0114137**	**16.02**	**0.000**	**.1604409**	**.2051819**
_cons	**-4260.635**	**266.0159**	**-16.02**	**0.000**	**-4782.016**	**-3739.253**

Note: 3632 failures and 3646 successes completely determined.

```
38 . margins, dydx(*)
```

Average marginal effects — Number of obs = **10,000**
Model VCE : **OIM**

Expression : **Pr(ydum), predict()**
dy/dx w.r.t. : **x1 x2 x3**

	Delta-method dy/dx	Std. Err.	z	P>|z|	[95% Conf.	Interval]
x1	**.0190283**	**.0015587**	**12.21**	**0.000**	**.0159734**	**.0220832**
x2	**-.0133973**	**.0002479**	**-54.04**	**0.000**	**-.0138833**	**-.0129114**
x3	**.0014645**	**4.41e-06**	**331.95**	**0.000**	**.0014558**	**.0014731**

```
39 .
40 . ** Delta Method
41 . probit ydum x1 x2 x3
```

```
Iteration 0:   log likelihood = -6931.3996
Iteration 1:   log likelihood = -287.39437
Iteration 2:   log likelihood =  -262.3666
Iteration 3:   log likelihood = -261.19541
Iteration 4:   log likelihood = -261.19416
Iteration 5:   log likelihood = -261.19416
```

Probit regression

Number of obs	=	**10,000**			
LR chi2(**3**)	=	**13340.41**			
Prob > chi2	=	**0.0000**			
Log likelihood = **-261.19416**	Pseudo R2	=	**0.9623**		

ydum	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
x1	**1.310577**	**.131925**	**9.93**	**0.000**	**1.052008**	**1.569145**
x2	**-.9257078**	**.0553151**	**-16.74**	**0.000**	**-1.034123**	**-.8172922**
x3	**.1013994**	**.0057763**	**17.55**	**0.000**	**.0900781**	**.1127207**
_cons	**-2363.215**	**134.6236**	**-17.55**	**0.000**	**-2627.073**	**-2099.358**

Note: 4182 failures and 4167 successes completely determined.

```
42 . margins, dydx(*) vce(delta)

Average marginal effects                        Number of obs     =     10,000
Model VCE    : OIM

Expression   : Pr(ydum), predict()
dy/dx w.r.t. : x1 x2 x3
```

	Delta-method dy/dx	Std. Err.	z	P>|z|	[95% Conf.	Interval]
x1	.0189197	.0015744	12.02	0.000	.0158339	.0220056
x2	-.0133637	.0002488	-53.72	0.000	-.0138513	-.0128761
x3	.0014638	3.09e-06	473.46	0.000	.0014578	.0014699

```
43 .
44 . ** Bootstrap
45 . bootstrap, reps(49) seed(12345): probit ydum x1 x2 x3
(running probit on estimation sample)

Bootstrap replications (49)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
.................................................

Probit regression                               Number of obs     =     10,000
                                                Replications      =         49
                                                Wald chi2(3)      =     343.64
                                                Prob > chi2       =     0.0000
Log likelihood = -261.19416                     Pseudo R2         =     0.9623
```

ydum	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
x1	1.310577	.1188293	11.03	0.000	1.077675	1.543478
x2	-.9257078	.0567094	-16.32	0.000	-1.036856	-.8145594
x3	.1013994	.0057884	17.52	0.000	.0900544	.1127445
_cons	-2363.215	134.8297	-17.53	0.000	-2627.476	-2098.954

```
Note: 4182 failures and 4167 successes completely determined.

46 .
47 . * REDO: Assignment 3
48 . * Exercise 1 Data Description
49 . clear all

50 . import delimited https://raw.githubusercontent.com/ms486/Econ613/master/Assignments/A3/product.
(13 vars, 4,470 obs)

51 .
52 . ** Average and dispersion
```

```
53 . sum(p*)
```

Variable	Obs	Mean	Std. Dev.	Min	Max
ppk_stk	4,470	.5184362	.1505174	.19	.67
pbb_stk	4,470	.5432103	.1203319	.19	1.01
pfl_stk	4,470	1.01502	.0428952	.95	1.16
phse_stk	4,470	.4371476	.1188312	.19	.64
pgen_stk	4,470	.3452819	.0351661	.25	.55
pimp_stk	4,470	.7807785	.1146461	.33	2.3
pss_tub	4,470	.8250895	.0612116	.5	.98
ppk_tub	4,470	1.077409	.0297261	.98	1.24
pfl_tub	4,470	1.189376	.0140545	.69	1.47
phse_tub	4,470	.5686734	.072455	.33	1.27

```
54 .
55 . g sales =.
(4,470 missing values generated)

56 . replace sales = ppk_stk if choice == 1
(1,766 real changes made)

57 . replace sales = pbb_stk  if choice == 2
(699 real changes made)

58 . replace sales = pfl_stk if choice == 3
(243 real changes made)

59 . replace sales = phse_stk if choice == 4
(593 real changes made)

60 . replace sales = pgen_stk if choice == 5
(315 real changes made)

61 . replace sales = pimp_stk if choice == 6
(74 real changes made)

62 . replace sales = pss_tub if choice == 7
(319 real changes made)

63 . replace sales = ppk_tub if choice == 8
(203 real changes made)

64 . replace sales = pfl_tub if choice == 9
(225 real changes made)

65 . replace sales = phse_tub if choice == 10
(33 real changes made)

66 .
67 . ** Market Share
68 . collapse (sum) sales , by(choice)
```

```
69 . egen totsales = sum(sales)

70 .  g mktshare = sales / totsales

71 . list choice mktshare
```

	choice	mktshare
1.	1	.3164004
2.	2	.1230866
3.	3	.0988726
4.	4	.0931612
5.	5	.0447412
6.	6	.0224712
7.	7	.0998426
8.	8	.0875344
9.	9	.1075665
10.	10	.0063232

```
72 .
73 . ** Merge Data
74 .
75 .
76 . import delimited https://raw.githubusercontent.com/ms486/Econ613/master/Assignments/A3/product.
(13 vars, 4,470 obs)

77 . merge m:1 hhid using demos.dta
```

Result	# of obs.	
not matched	0	
matched	4,470	(_merge==3)

```
78 .
79 . * Exercise 2 + 4 First Model (Conditional Logit) + Marginal Effects
80 .
81 . g sales =.
(4,470 missing values generated)

82 . replace sales = ppk_stk if choice == 1
(1,766 real changes made)

83 . replace sales = pbb_stk  if choice == 2
(699 real changes made)

84 . replace sales = pfl_stk if choice == 3
(243 real changes made)

85 . replace sales = phse_stk if choice == 4
(593 real changes made)
```

```
 86 . replace sales = pgen_stk if choice == 5
    (315 real changes made)

 87 . replace sales = pimp_stk if choice == 6
    (74 real changes made)

 88 . replace sales = pss_tub if choice == 7
    (319 real changes made)

 89 . replace sales = ppk_tub if choice == 8
    (203 real changes made)

 90 . replace sales = pfl_tub if choice == 9
    (225 real changes made)

 91 . replace sales = phse_tub if choice == 10
    (33 real changes made)

 92 .
 93 . bysort hhid: gen set = _n

 94 .
 95 . local j = 10

 96 . forval j = 1 / 10 {
      2. gen chosen`j' = 0
      3.
 97 . }

 98 .
 99 . local j = 10

100 . forval j = 1 / 10 {
      2. replace chosen`j' = 1 if choice == `j'
      3. }
    (1,766 real changes made)
    (699 real changes made)
    (243 real changes made)
    (593 real changes made)
    (315 real changes made)
    (74 real changes made)
    (319 real changes made)
    (203 real changes made)
    (225 real changes made)
    (33 real changes made)

101 .
102 . reshape long chosen, i(v1) j(c)
    (note: j = 1 2 3 4 5 6 7 8 9 10)

    Data                               wide   ->   long
    -----------------------------------------------------------------------------
    Number of obs.                     4470   ->   44700
    Number of variables                  33   ->      25
    j variable (10 values)                    ->   c
    xij variables:
               chosen1 chosen2 ... chosen10   ->   chosen
    -----------------------------------------------------------------------------
```

```
103 .
104 . g price =.
(44,700 missing values generated)

105 . replace price = ppk_stk if c == 1
(4,470 real changes made)

106 . replace price = pbb_stk  if c == 2
(4,470 real changes made)

107 . replace price = pfl_stk if c == 3
(4,470 real changes made)

108 . replace price = phse_stk if c == 4
(4,470 real changes made)

109 . replace price = pgen_stk if c == 5
(4,470 real changes made)

110 . replace price = pimp_stk if c == 6
(4,470 real changes made)

111 . replace price = pss_tub if c == 7
(4,470 real changes made)

112 . replace price = ppk_tub if c == 8
(4,470 real changes made)

113 . replace price = pfl_tub if c == 9
(4,470 real changes made)

114 . replace price = phse_tub if c == 10
(4,470 real changes made)

115 .
116 . egen gid = group(set hhid)

117 . clogit chosen price, group(gid) nolog

Conditional (fixed-effects) logistic regression

                                                Number of obs     =     44,700
                                                LR chi2(1)        =    1774.26
                                                Prob > chi2       =     0.0000
Log likelihood = -9405.4251                     Pseudo R2         =     0.0862

------------------------------------------------------------------------------
      chosen |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       price |  -2.428201   .0626529   -38.76   0.000    -2.550999   -2.305404
------------------------------------------------------------------------------

118 .
```

```
119 . margins, dydx(*)

Average marginal effects                        Number of obs     =     44,700
Model VCE    : OIM

Expression   : Pr(chosen|fixed effect is 0), predict(pu0)
dy/dx w.r.t. : price
```

	Delta-method dy/dx	Std. Err.	z	P>|z|	[95% Conf.	Interval]
price	-.3164552	.0001988	-1592.02	0.000	-.3168448	-.3160656

```
120 .
121 . * Exercise 3 + 4 Second Model (Multinomial Logit) + Marginal Effects
122 .
123 . mlogit choice income if chosen == 1, nolog

Multinomial logistic regression                 Number of obs     =      4,470
                                                LR chi2(9)        =      98.20
                                                Prob > chi2       =     0.0000
Log likelihood =  -8236.757                     Pseudo R2         =     0.0059
```

choice		Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
1		(base outcome)					
2							
	income	-.0030887	.003114	-0.99	0.321	-.009192	.0030145
	_cons	-.8453242	.0931355	-9.08	0.000	-1.027866	-.662782
3							
	income	.0145862	.0038255	3.81	0.000	.0070885	.022084
	_cons	-2.399858	.1335802	-17.97	0.000	-2.66167	-2.138045
4							
	income	.0040504	.0030926	1.31	0.190	-.0020109	.0101118
	_cons	-1.201326	.0971021	-12.37	0.000	-1.391643	-1.01101
5							
	income	-.0012536	.0042024	-0.30	0.765	-.0094901	.0069829
	_cons	-1.690582	.1269952	-13.31	0.000	-1.939488	-1.441676
6							
	income	.030612	.004674	6.55	0.000	.0214512	.0397729
	_cons	-4.139767	.210989	-19.62	0.000	-4.553298	-3.726237
7							
	income	-.0069326	.0044161	-1.57	0.116	-.015588	.0017228
	_cons	-1.531042	.1280434	-11.96	0.000	-1.782002	-1.280081
8							
	income	.0228862	.0036217	6.32	0.000	.0157879	.0299846
	_cons	-2.848353	.1393848	-20.44	0.000	-3.121543	-2.575164
9							
	income	.017743	.0037623	4.72	0.000	.010369	.0251169
	_cons	-2.575597	.13614	-18.92	0.000	-2.842427	-2.308768
10							
	income	.0107909	.01013	1.07	0.287	-.0090636	.0306455
	_cons	-4.28227	.345792	-12.38	0.000	-4.96001	-3.60453

```
124 .
125 . mfx

Marginal effects after mlogit
      y  = Pr(choice==1) (predict)
         =  .39801714
```

variable	dy/dx	Std. Err.	z	P>|z|	[	95% C.I.	]	X
income	-.0010625	.00049	-2.18	0.029	-.002017	-.000108		27.6639

```
126 .
127 . * Exercise 5 IIA
128 .
129 . asmixlogit chosen price, case(gid) alternatives(c) casevars(income) nolog

Alternative-specific mixed logit                Number of obs      =     44,700
Case variable: gid                              Number of cases    =      4,470

Alternative variable: c                         Alts per case: min =         10
                                                               avg =       10.0
                                                               max =         10

Integration points:              0              Wald chi2(10)      =    1538.10
Log likelihood =        -7417.9325              Prob > chi2        =     0.0000
```

chosen	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
c						
price	-6.659669	.1747698	-38.11	0.000	-7.002212	-6.317127
1	(base alternative)					
2						
income	-.0042599	.0034392	-1.24	0.215	-.0110007	.0024808
_cons	-.8406734	.1038446	-8.10	0.000	-1.044205	-.6371417
3						
income	.014344	.0039221	3.66	0.000	.0066568	.0220311
_cons	.8886069	.1594585	5.57	0.000	.5760739	1.20114
4						
income	.0040998	.0032042	1.28	0.201	-.0021803	.01038
_cons	-1.828492	.103218	-17.71	0.000	-2.030795	-1.626188
5						
income	-.0011829	.0042971	-0.28	0.783	-.009605	.0072393
_cons	-2.87341	.1347573	-21.32	0.000	-3.13753	-2.609291
6						
income	.029809	.0047267	6.31	0.000	.0205448	.0390731
_cons	-2.457119	.215426	-11.41	0.000	-2.879346	-2.034891
7						
income	-.0092456	.0045935	-2.01	0.044	-.0182487	-.0002425
_cons	.4968692	.1424824	3.49	0.000	.2176089	.7761295
8						
income	.0219965	.0038203	5.76	0.000	.0145088	.0294841
_cons	.80306	.1709199	4.70	0.000	.4680631	1.138057

9						
income	**.0169911**	**.0039155**	**4.34**	**0.000**	**.0093169**	**.0246653**
_cons	**1.864125**	**.1799469**	**10.36**	**0.000**	**1.511436**	**2.216815**
10						
income	**.0087596**	**.0103007**	**0.85**	**0.395**	**-.0114295**	**.0289487**
_cons	**-4.142386**	**.3506563**	**-11.81**	**0.000**	**-4.829659**	**-3.455112**

```
130 . estimates store bf

131 .
132 . drop if c == 1
(4,470 observations deleted)

133 . asmixlogit chosen price, case(gid) alternatives(c) casevars(income) nolog
Note: 1766 cases (15894 obs) dropped due to no positive outcome, multiple positive outcomes, or a
      observation per case
```

Alternative-specific mixed logit — Number of obs = **24,336**
Case variable: **gid** — Number of cases = **2,704**

Alternative variable: **c** — Alts per case: min = **9**, avg = **9.0**, max = **9**

Integration points: **0** — Wald chi2(**9**) = **758.59**
Log likelihood = **-4884.1755** — Prob > chi2 = **0.0000**

chosen	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
c						
price	**-6.4221**	**.2446946**	**-26.25**	**0.000**	**-6.901692**	**-5.942507**
2	(base alternative)					
3						
income	**.0184166**	**.0045476**	**4.05**	**0.000**	**.0095035**	**.0273298**
_cons	**1.636494**	**.1943569**	**8.42**	**0.000**	**1.255561**	**2.017426**
4						
income	**.0075214**	**.0039386**	**1.91**	**0.056**	**-.000198**	**.0152408**
_cons	**-.9428326**	**.1239508**	**-7.61**	**0.000**	**-1.185772**	**-.6998935**
5						
income	**.0030968**	**.0048094**	**0.64**	**0.520**	**-.0063295**	**.0125231**
_cons	**-1.968933**	**.1514053**	**-13.00**	**0.000**	**-2.265682**	**-1.672184**
6						
income	**.0336065**	**.0053594**	**6.27**	**0.000**	**.0231022**	**.0441108**
_cons	**-1.647682**	**.2327193**	**-7.08**	**0.000**	**-2.103803**	**-1.19156**
7						
income	**-.0043201**	**.0050706**	**-0.85**	**0.394**	**-.0142582**	**.0056181**
_cons	**1.223079**	**.1651403**	**7.41**	**0.000**	**.8994096**	**1.546748**
8						
income	**.0259724**	**.0044753**	**5.80**	**0.000**	**.017201**	**.0347438**
_cons	**1.555129**	**.2084066**	**7.46**	**0.000**	**1.14666**	**1.963599**
9						
income	**.0210461**	**.0045316**	**4.64**	**0.000**	**.0121643**	**.0299278**
_cons	**2.576022**	**.2254343**	**11.43**	**0.000**	**2.134179**	**3.017865**

10

income	**.0126704**	**.0103331**	**1.23**	**0.220**	**-.0075821**	**.0329228**
_cons	**-3.243888**	**.3515806**	**-9.23**	**0.000**	**-3.932973**	**-2.554802**

```
134 . estimates store br

135 .
136 . di "chi2(10) = " 2*( _est_bf - _est_br )
chi2(10) = 2

137 . di "Prob > chi2 = "chi2tail(10, 2*( _est_bf - _est_br ))
Prob > chi2 = .99634015

138 .
139 .
140 .
141 . * REDO: Assignment 4
142 . clear all

143 .
144 . *Exercise 1 Data
145 . import delimited https://raw.githubusercontent.com/ms486/Econ613/master/Assignments/A4/Koop-Tob
(10 vars, 17,919 obs)

146 . xtset personid timetrnd
       panel variable:  personid (unbalanced)
        time variable:  timetrnd, 0 to 14, but with gaps
                delta:  1 unit

147 . reshape wide educ logwage potexper , i(personid) j( timetrnd)
(note: j = 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14)

Data                               long   ->   wide
-----------------------------------------------------------------------------
Number of obs.                    17919   ->    2178
Number of variables                  10   ->      51
j variable (15 values)         timetrnd   ->   (dropped)
xij variables:
                                    educ   ->   educ0 educ1 ... educ14
                                 logwage   ->   logwage0 logwage1 ... logwage14
                                potexper   ->   potexper0 potexper1 ... potexper14
-----------------------------------------------------------------------------

148 . sample 5, count
(2,173 observations deleted)

149 . list
```

1.

personid	educ0	logwage0	potex~r0	educ1	logwage1	potex~r1	educ2	logwage2
1034	**12**	**2.89**	**4**	**12**	**3.11**	**5**	**.**	**.**

educ3	logwage3	potex~r3	educ4	logwage4	potex~r4	educ5	logwage5	potexp~5
.	.	.	.	.	.	.	.	.

logwage6	potexp~6	educ7	logwage7	potexp~7	educ8	logwage8	potexp~8	educ9
2.98	**10**	**12**	**2.79**	**11**	**12**	**2.73**	**12**	**12**

potexp~9	educ10	logwa~10	potex~10	educ11	logwa~11	potex~11	educ12
13	**12**	**3.08**	**14**	**12**	**2.84**	**15**	**12**

potex~12	educ13	logwa~13	potex~13	educ14	logwa~14	potex~14	ability
16	**12**	**2.95**	**17**	**12**	**2.87**	**18**	**.26**

fathered	brknhome	siblings
12	**0**	**4**

2.

personid	educ0	logwage0	potex~r0	educ1	logwage1	potex~r1	educ2	logwage2	
1808	.	.	.	.	.	.	.	.	

educ3	logwage3	potex~r3	educ4	logwage4	potex~r4	educ5	logwage5	potexp~5
.	.	.	**12**	**1.82**	**3**	**12**	**2.11**	**4**

logwage6	potexp~6	educ7	logwage7	potexp~7	educ8	logwage8	potexp~8	educ9	
.	.	.	.	.	.	.	.	.	

potexp~9	educ10	logwa~10	potex~10	educ11	logwa~11	potex~11	educ12	
.	.	.	.	.	.	.	.	

potex~12	educ13	logwa~13	potex~13	educ14	logwa~14	potex~14	ability	
.	.	.	.	.	.	.	**-.53**	

fathered	brknhome	siblings
8	**0**	**5**

3.

personid	educ0	logwage0	potex~r0	educ1	logwage1	potex~r1	educ2	logwage2	
2089	.	.	.	.	.	.	.	.	

educ3	logwage3	potex~r3	educ4	logwage4	potex~r4	educ5	logwage5	potexp~5
11	**1.72**	**2**	.	.	.	.	.	.

logwage6	potexp~6	educ7	logwage7	potexp~7	educ8	logwage8	potexp~8	educ9	
1.92	**4**	**12**	**2.5**	**5**	**12**	**2.45**	**6**	**12**	

potexp~9	educ10	logwa~10	potex~10	educ11	logwa~11	potex~11	educ12	
7	**12**	**2.72**	**8**	**12**	**2.72**	**9**	**12**	

potex~12	educ13	logwa~13	potex~13	educ14	logwa~14	potex~14	ability	
10	**12**	**2.64**	**11**	**12**	**2.66**	**12**	**-.24**	

fathered	brknhome	siblings
17	**0**	**3**

4.

personid	educ0	logwage0	potex~r0	educ1	logwage1	potex~r1	educ2	logwage2	
48	**12**	**1.82**	**2**	**12**	**1.77**	**3**	**12**	**1.89**	

educ3	logwage3	potex~r3	educ4	logwage4	potex~r4	educ5	logwage5	potexp~5
12	**1.8**	**5**	**12**	**2.08**	**6**	**12**	**1.74**	**7**

logwage6	potexp~6	educ7	logwage7	potexp~7	educ8	logwage8	potexp~8	educ9	
1.42	**8**	**12**	**1.79**	**9**	**12**	**1.77**	**10**	**12**	

potexp~9	educ10	logwa~10	potex~10	educ11	logwa~11	potex~11	educ12	
11	**12**	**1.92**	**12**	**12**	**2.61**	**13**	**12**	

potex~12	educ13	logwa~13	potex~13	educ14	logwa~14	potex~14	ability	
14	**12**	**2.36**	**15**	**12**	**2.05**	**16**	**-.59**	

fathered	brknhome	siblings
14	**1**	**0**

5.

personid	educ0	logwage0	potex~r0	educ1	logwage1	potex~r1	educ2	logwage2
356	**12**	**2.07**	**3**	**12**	**2.27**	**4**	**12**	**1.97**

educ3	logwage3	potex~r3	educ4	logwage4	potex~r4	educ5	logwage5	potexp~5
12	**2.06**	**6**	**12**	**2.24**	**7**	**12**	**2.4**	**8**

logwage6	potexp~6	educ7	logwage7	potexp~7	educ8	logwage8	potexp~8	educ9
2.56	**9**	**12**	**2.63**	**10**	**12**	**2.79**	**11**	**12**

potexp~9	educ10	logwa~10	potex~10	educ11	logwa~11	potex~11	educ12
12	**12**	**2.6**	**13**	**12**	**2.55**	**14**	**12**

potex~12	educ13	logwa~13	potex~13	educ14	logwa~14	potex~14	ability
15	**12**	**2.48**	**16**	**12**	**2.65**	**17**	**-.25**

fathered	brknhome	siblings
9	**0**	**2**

```
150 .
151 . * Exercise 2 Random Effects
152 . clear all

153 . import delimited https://raw.githubusercontent.com/ms486/Econ613/master/Assignments/A4/Koop-Tob
    (10 vars, 17,919 obs)

154 . xtset personid timetrnd
           panel variable:  personid (unbalanced)
            time variable:  timetrnd, 0 to 14, but with gaps
                    delta:  1 unit

155 . xtreg logwage educ potexper, re

Random-effects GLS regression                   Number of obs     =     17,919
Group variable: personid                        Number of groups  =      2,178

R-sq:                                           Obs per group:
     within  = 0.1961                                         min =          1
     between = 0.1533                                         avg =        8.2
     overall = 0.1578                                         max =         15

                                                Wald chi2(2)      =    4209.96
corr(u_i, X)   = 0 (assumed)                    Prob > chi2       =     0.0000

------------------------------------------------------------------------------
     logwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |    .107938   .0033832    31.90   0.000     .1013071     .114569
    potexper |   .0387645   .0007178    54.00   0.000     .0373576    .0401714
       _cons |   .5635206   .0438846    12.84   0.000     .4775083    .6495328
-------------+----------------------------------------------------------------
     sigma_u |  .37207276
     sigma_e |  .33545728
         rho |   .5516129   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

```
156 .
157 . * Exercise 3 Fixed Effects Model
158 . ** Between Estimator
159 . collapse (mean) logwage potexper educ , by(personid)

160 . reg logwage potexper educ
```

Source	SS	df	MS	Number of obs	=	2,178
				F(2, 2175)	=	200.01
Model	63.7247799	2	31.8623899	Prob > F	=	0.0000
Residual	346.490052	2,175	.159305771	R-squared	=	0.1553
				Adj R-squared	=	0.1546
Total	410.214832	2,177	.18843125	Root MSE	=	.39913

logwage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
potexper	.0259987	.0036049	7.21	0.000	.0189294	.0330681
educ	.0930999	.0046685	19.94	0.000	.0839447	.1022551
_cons	.8455688	.0770179	10.98	0.000	.6945324	.9966052

```
161 .
162 . ** Within Estimator
163 . clear all

164 . import delimited https://raw.githubusercontent.com/ms486/Econ613/master/Assignments/A4/Koop-Tob
    (10 vars, 17,919 obs)

165 . xtset personid timetrnd
           panel variable:  personid (unbalanced)
            time variable:  timetrnd, 0 to 14, but with gaps
                    delta:  1 unit

166 .
167 . egen mean_wage = mean(logwage), by(personid)

168 . egen mean_exper = mean(potexper), by(personid)

169 . egen mean_educ = mean(educ), by(personid)

170 . g fe_wage = logwage - mean_wage

171 . g fe_exper = potexper - mean_exper

172 . g fe_educ = educ - mean_educ

173 . reg fe_wage fe_exper fe_educ, nocon
```

Source	SS	df	MS	Number of obs	=	17,919
				F(2, 17917)	=	2189.65
Model	432.903006	2	216.451503	Prob > F	=	0.0000
Residual	1771.13462	17,917	.098852186	R-squared	=	0.1964
				Adj R-squared	=	0.1963
Total	2204.03763	17,919	.123000035	Root MSE	=	.31441

fe_wage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
fe_exper	.0385611	.0007109	54.24	0.000	.0371677	.0399545
fe_educ	.123662	.0054003	22.90	0.000	.1130769	.1342472

```
174 .
175 . ** First time difference
176 . g fd_wage = logwage - l.logwage
(4,235 missing values generated)

177 . g fd_exper = potexper - l.potexper
(4,235 missing values generated)

178 . g fd_educ = educ - l.educ
(4,235 missing values generated)

179 . reg fd_wage fd_exper fd_educ, nocon
```

Source	SS	df	MS		
Model	**38.7282937**	**2**	**19.3641469**	Number of obs	= **13,684**
Residual	**1541.54171**	**13,682**	**.112669326**	F(2, 13682)	= **171.87**
				Prob > F	= **0.0000**
Total	**1580.27001**	**13,684**	**.115483046**	R-squared	= **0.0245**
				Adj R-squared	= **0.0244**
				Root MSE	= **.33566**

fd_wage	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
fd_exper	**.0535369**	**.0029221**	**18.32**	**0.000**	**.0478092**	**.0592647**
fd_educ	**.0431084**	**.0151792**	**2.84**	**0.005**	**.0133551**	**.0728617**

```
180 .
181 . * Exercise 4 Understanding Fixed Effects
182 . clear all

183 . import delimited https://raw.githubusercontent.com/ms486/Econ613/master/Assignments/A4/Koop-Tob
(10 vars, 17,919 obs)

184 . xtset personid timetrnd
       panel variable:  personid (unbalanced)
        time variable:  timetrnd, 0 to 14, but with gaps
                delta:  1 unit

185 . reshape wide educ logwage potexper , i(personid) j( timetrnd)
(note: j = 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14)

Data                               long   ->   wide
-----------------------------------------------------------------------------
Number of obs.                    17919   ->    2178
Number of variables                  10   ->      51
j variable (15 values)         timetrnd   ->   (dropped)
xij variables:
                                   educ   ->   educ0 educ1 ... educ14
                                logwage   ->   logwage0 logwage1 ... logwage14
                               potexper   ->   potexper0 potexper1 ... potexper14
-----------------------------------------------------------------------------
```

```
186 . sample 100, count
    (2,078 observations deleted)

187 . reshape long educ logwage potexper , i(personid) j( timetrnd)
    (note: j = 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14)

    Data                               wide   ->   long
    -----------------------------------------------------------------------------
    Number of obs.                      100   ->    1500
    Number of variables                  51   ->      10
    j variable (15 values)                    ->   timetrnd
    xij variables:
                  educ0 educ1 ... educ14      ->   educ
            logwage0 logwage1 ... logwage14   ->   logwage
         potexper0 potexper1 ... potexper14   ->   potexper
    -----------------------------------------------------------------------------

188 .
189 . gen alpha = .
    (1,500 missing values generated)

190 . qui reg logwage educ potexper ibn.personid, noconst

191 . levelsof personid, local(levels)
    9 35 66 68 90 138 139 185 192 200 202 217 219 228 230 297 313 331 350 373 397 406 407 477 526 536
    >  654 738 766 812 815 846 848 865 885 920 939 956 976 1034 1037 1038 1064 1135 1143 1153 1162 11
    >  1211 1239 1242 1321 1335 1387 1389 1414 1433 1454 1459 1467 1480 1484 1501 1503 1544 1567 1584
    > 651 1706 1801 1803 1809 1823 1861 1873 1885 1887 1915 1964 1966 2009 2017 2076 2089 2098 2100 2
    > 0 2133 2154 2170 2178

192 . foreach l of local levels {
      2. replace alpha = _b[`l'.personid] if personid == `l'
      3. }
    (15 real changes made)
    (15 real changes made)
    (15 real changes made)
```

(15 real changes made)
(15 real changes made)
(15 real changes made)

```
(15 real changes made)
(15 real changes made)
(15 real changes made)

193 .
194 . reshape wide alpha educ logwage potexper , i(personid) j( timetrnd)
(note: j = 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14)
```

Data	long	->	wide
Number of obs.	**1500**	->	**100**
Number of variables	**11**	->	**66**
j variable (15 values)	**timetrnd**	->	(dropped)
xij variables:			
	alpha	->	**alpha0 alpha1 ... alpha14**
	educ	->	**educ0 educ1 ... educ14**
	logwage	->	**logwage0 logwage1 ... logwage14**
	potexper	->	**potexper0 potexper1 ... potexper14**

```
195 . reg alpha0 ability  mothered fathered brknhome siblings
```

Source	SS	df	MS		
				Number of obs	= **100**
				F(5, 94)	= **1.95**
Model	**1.74808852**	**5**	**.349617703**	Prob > F	= **0.0929**
Residual	**16.8318459**	**94**	**.179062191**	R-squared	= **0.0941**
				Adj R-squared	= **0.0459**
Total	**18.5799344**	**99**	**.187676106**	Root MSE	= **.42316**

alpha0	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ability	**.0955583**	**.0541129**	**1.77**	**0.081**	**-.0118841**	**.2030007**
mothered	**-.0232063**	**.0214727**	**-1.08**	**0.283**	**-.0658408**	**.0194283**
fathered	**.0258986**	**.0161378**	**1.60**	**0.112**	**-.0061434**	**.0579406**
brknhome	**-.125615**	**.1145152**	**-1.10**	**0.275**	**-.3529876**	**.1017577**
siblings	**.0035094**	**.0213934**	**0.16**	**0.870**	**-.0389677**	**.0459865**
_cons	**1.238214**	**.2771964**	**4.47**	**0.000**	**.6878343**	**1.788594**

```
196 .
197 .
198 .
end of do-file

199 .
```