```
* Panel Data Models in Stata
* Copyright 2013 by Ani Katchova

clear all
set more off

use C:\Econometrics\Data\panel_wage

global id id
global t t
global ylist lwage
global xlist exp exp2 wks ed

describe $id $t $ylist $xlist
summarize $id $t $ylist $xlist

* Set data as panel data
sort $id $t
xtset $id $t
xtdescribe
xtsum $id $t $ylist $xlist

* Pooled OLS estimator
reg $ylist $xlist

* Population-averaged estimator
xtreg $ylist $xlist, pa

* Between estimator
xtreg $ylist $xlist, be

* Fixed effects or within estimator
xtreg $ylist $xlist, fe

* First-differences estimator
reg D.($ylist $xlist), noconstant

* Random effects estimator
xtreg $ylist $xlist, re theta

* Hausman test for fixed versus random effects model
quietly xtreg $ylist $xlist, fe
estimates store fixed
quietly xtreg $ylist $xlist, re
estimates store random
hausman fixed random

* Breusch-Pagan LM test for random effects versus OLS
quietly xtreg $ylist $xlist, re
xttest0

* Recovering individual-specific effects
quietly xtreg $ylist $xlist, fe
predict alphafehat, u
sum alphafehat
```

```
. * Panel Data Models in Stata
. * Copyright 2013 by Ani Katchova
.
. clear all

. set more off

.
. use C:\Econometrics\Data\panel_wage
(PSID wage data 1976-82 from Baltagi and Khanti-Akom (1990))

.
. global id id

. global t t

. global ylist lwage

. global xlist exp exp2 wks ed

.
. describe $id $t $ylist $xlist

              storage  display     value
variable name   type   format      label      variable label
------------------------------------------------------------------------------------------
------------------
id              float  %9.0g
t               float  %9.0g
lwage           float  %9.0g                 log wage
exp             float  %9.0g                 years of full-time work experience
exp2            float  %9.0g
wks             float  %9.0g                 weeks worked
ed              float  %9.0g                 years of education

. summarize $id $t $ylist $xlist

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
          id |      4165         298    171.7821          1        595
           t |      4165           4     2.00024          1          7
       lwage |      4165    6.676346    .4615122    4.60517      8.537
         exp |      4165    19.85378    10.96637          1         51
        exp2 |      4165     514.405    496.9962          1       2601
-------------+--------------------------------------------------------
         wks |      4165    46.81152    5.129098          5         52
          ed |      4165    12.84538    2.787995          4         17

.
. * Set data as panel data
. sort $id $t

. xtset $id $t
       panel variable:  id (strongly balanced)
```

```
        time variable:  t, 1 to 7
                delta:  1 unit

. xtdescribe

      id:  1, 2, ..., 595                                 n =        595
       t:  1, 2, ..., 7                                   T =          7
           Delta(t) = 1 unit
           Span(t)  = 7 periods
           (id*t uniquely identifies each observation)

Distribution of T_i:   min      5%     25%       50%       75%     95%     max
                         7       7       7         7         7       7       7

     Freq.  Percent    Cum. |  Pattern
 ---------------------------+---------
      595    100.00  100.00 |  1111111
 ---------------------------+---------
      595    100.00         |  XXXXXXX

. xtsum $id $t $ylist $xlist

Variable         |      Mean   Std. Dev.       Min        Max |    Observations
-----------------+--------------------------------------------+----------------
id       overall |       298   171.7821          1        595 |     N =    4165
         between |              171.906          1        595 |     n =     595
         within  |                   0        298        298 |     T =       7
                 |                                           |
t        overall |         4    2.00024          1          7 |     N =    4165
         between |                   0          4          4 |     n =     595
         within  |              2.00024          1          7 |     T =       7
                 |                                           |
lwage    overall |  6.676346   .4615122    4.60517      8.537 |     N =    4165
         between |              .3942387     5.3364   7.813596 |     n =     595
         within  |              .2404023   4.781808   8.621092 |     T =       7
                 |                                           |
exp      overall |  19.85378   10.96637          1         51 |     N =    4165
         between |              10.79018          4         48 |     n =     595
         within  |              2.00024   16.85378   22.85378 |     T =       7
                 |                                           |
exp2     overall |   514.405   496.9962          1       2601 |     N =    4165
         between |              489.0495         20       2308 |     n =     595
         within  |              90.44581    231.405    807.405 |     T =       7
                 |                                           |
wks      overall |  46.81152   5.129098          5         52 |     N =    4165
         between |              3.284016   31.57143   51.57143 |     n =     595
         within  |              3.941881    12.2401   63.66867 |     T =       7
                 |                                           |
ed       overall |  12.84538   2.787995          4         17 |     N =    4165
         between |              2.790006          4         17 |     n =     595
         within  |                   0   12.84538   12.84538 |     T =       7

.
. * Pooled OLS estimator
. reg $ylist $xlist
```

```
      Source |       SS       df       MS              Number of obs =    4165
-------------+------------------------------           F(  4,  4160) =  411.62
       Model |  251.491445     4  62.8728613           Prob > F      =  0.0000
    Residual |  635.413457  4160  .152743619           R-squared     =  0.2836
-------------+------------------------------           Adj R-squared =  0.2829
       Total |  886.904902  4164  .212993492           Root MSE      =  .39082

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |    .044675   .0023929    18.67   0.000     .0399838    .0493663
        exp2 |  -.0007156   .0000528   -13.56   0.000    -.0008191   -.0006121
         wks |    .005827   .0011827     4.93   0.000     .0035084    .0081456
          ed |   .0760407   .0022266    34.15   0.000     .0716754     .080406
       _cons |   4.907961   .0673297    72.89   0.000     4.775959    5.039963
------------------------------------------------------------------------------

.
. * Population-averaged estimator
. xtreg $ylist $xlist, pa

Iteration 1: tolerance = .10073745
Iteration 2: tolerance = .10171516
Iteration 3: tolerance = .0939929
Iteration 4: tolerance = .05375245
Iteration 5: tolerance = .01851339
Iteration 6: tolerance = .00488708
Iteration 7: tolerance = .00118383
Iteration 8: tolerance = .00028047
Iteration 9: tolerance = .0000661
Iteration 10: tolerance = .00001556
Iteration 11: tolerance = 3.660e-06
Iteration 12: tolerance = 8.612e-07

GEE population-averaged model                   Number of obs      =      4165
Group variable:                         id      Number of groups   =       595
Link:                             identity      Obs per group: min =         7
Family:                           Gaussian                     avg =       7.0
Correlation:                  exchangeable                     max =         7
                                              Wald chi2(4)       =   6160.57
Scale parameter:                  .7476287    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .1079955   .0024527    44.03   0.000     .1031883    .1128026
        exp2 |  -.0005202   .0000543    -9.59   0.000    -.0006266   -.0004139
         wks |   .0008365   .0006042     1.38   0.166    -.0003477    .0020208
          ed |   .1378558   .0125814    10.96   0.000     .1131968    .1625149
       _cons |    2.98986   .1711799    17.47   0.000     2.654353    3.325366
------------------------------------------------------------------------------

.
. * Between estimator
```

```
. xtreg $ylist $xlist, be

Between regression (regression on group means)  Number of obs      =      4165
Group variable: id                              Number of groups   =       595

R-sq:  within  = 0.1357                         Obs per group: min =         7
       between = 0.3264                                        avg =       7.0
       overall = 0.2723                                        max =         7

                                                F(4,590)           =     71.48
sd(u_i + avg(e_i.))=   .324656                  Prob > F           =    0.0000

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |    .038153   .0056967     6.70   0.000     .0269647    .0493412
        exp2 |  -.0006313   .0001257    -5.02   0.000    -.0008781   -.0003844
         wks |   .0130903   .0040659     3.22   0.001     .0051048    .0210757
          ed |   .0737838   .0048985    15.06   0.000     .0641632    .0834044
       _cons |   4.683039   .2100989    22.29   0.000     4.270407    5.095672
------------------------------------------------------------------------------

.
. * Fixed effects or within estimator
. xtreg $ylist $xlist, fe
note: ed omitted because of collinearity

Fixed-effects (within) regression               Number of obs      =      4165
Group variable: id                              Number of groups   =       595

R-sq:  within  = 0.6566                         Obs per group: min =         7
       between = 0.0276                                        avg =       7.0
       overall = 0.0476                                        max =         7

                                                F(3,3567)          =   2273.74
corr(u_i, Xb)  = -0.9107                        Prob > F           =    0.0000

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .1137879   .0024689    46.09   0.000     .1089473    .1186284
        exp2 |  -.0004244   .0000546    -7.77   0.000    -.0005315   -.0003173
         wks |   .0008359   .0005997     1.39   0.163    -.0003399    .0020116
          ed |          0  (omitted)
       _cons |   4.596396   .0389061   118.14   0.000     4.520116    4.672677
-------------+----------------------------------------------------------------
     sigma_u |  1.0362039
     sigma_e |  .15220316
         rho |  .97888036   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(594, 3567) =    56.52           Prob > F = 0.0000

.
. * First-differences estimator
. reg D.($ylist $xlist), noconstant
```

```
note: _delete omitted because of collinearity

      Source |       SS       df       MS              Number of obs =    3570
-------------+------------------------------           F(  3,  3567) =  337.12
       Model |  33.3371458     3  11.1123819           Prob > F      =  0.0000
    Residual |   117.57812  3567  .032962748           R-squared     =  0.2209
-------------+------------------------------           Adj R-squared =  0.2202
       Total |  150.915266  3570  .042273184           Root MSE      =  .18156

------------------------------------------------------------------------------
     D.lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |
          D1. |   .1170654   .0063106    18.55   0.000     .1046927    .1294381
              |
        exp2 |
          D1. |  -.0005321   .0001393    -3.82   0.000    -.0008052   -.0002591
              |
         wks |
          D1. |  -.0002683   .0005648    -0.47   0.635    -.0013757    .0008392
              |
          ed |
          D1. |          0  (omitted)
------------------------------------------------------------------------------

.
. * Random effects estimator
. xtreg $ylist $xlist, re theta

Random-effects GLS regression                   Number of obs      =      4165
Group variable: id                              Number of groups   =       595

R-sq:  within  = 0.6340                         Obs per group: min =         7
       between = 0.1716                                        avg =       7.0
       overall = 0.1830                                        max =         7

                                                Wald chi2(4)       =   3012.45
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000
theta          = .82280511

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .0888609   .0028178    31.54   0.000     .0833382    .0943837
        exp2 |  -.0007726   .0000623   -12.41   0.000    -.0008946   -.0006505
         wks |   .0009658   .0007433     1.30   0.194     -.000491    .0024226
          ed |   .1117099   .0060572    18.44   0.000     .0998381    .1235818
       _cons |   3.829366   .0936336    40.90   0.000     3.645848    4.012885
-------------+----------------------------------------------------------------
     sigma_u |  .31951859
     sigma_e |  .15220316
         rho |  .81505521   (fraction of variance due to u_i)
------------------------------------------------------------------------------

.
```

```
. * Hausman test for fixed versus random effects model
. quietly xtreg $ylist $xlist, fe

. estimates store fixed

. quietly xtreg $ylist $xlist, re

. estimates store random

. hausman fixed random

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed        random       Difference          S.E.
-------------+------------------------------------------------------------
         exp |    .1137879     .0888609        .0249269               .
        exp2 |   -.0004244    -.0007726        .0003482               .
         wks |    .0008359     .0009658       -.0001299               .
------------------------------------------------------------------------------
                           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(3) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =     6191.43
                Prob>chi2 =      0.0000
                (V_b-V_B is not positive definite)

.
. * Breusch-Pagan LM test for random effects versus OLS
. quietly xtreg $ylist $xlist, re

. 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 |   .2129935       .4615122
                       e |   .0231658       .1522032
                       u |   .1020921       .3195186

        Test:   Var(u) = 0
                             chibar2(01) =  5192.13
                          Prob > chibar2 =   0.0000

.
. * Recovering individual-specific effects
. quietly xtreg $ylist $xlist, fe

. predict alphafehat, u
```

```
. sum alphafehat

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
  alphafehat |      4165   -1.97e-10    1.035457  -3.700898   1.896135
```