Panel Data Methods
1: Short Panels - Basics

© A. Colin Cameron
Univ. of Calif. Davis

.

Based on A. Colin Cameron and Pravin K. Trivedi (2009, 2010), Microeconometrics using Stata (MUS), Stata Press. and A. Colin Cameron and Pravin K. Trivedi (2005), Microeconometrics: Methods and Applications (MMA), C.U.P.

November 2013

1. Introduction

- Panel data are repeated measures on individuals (i) over time (t).
  - Regress y_{it} on $\mathbf{x}_{it}$ for $i = 1, ..., N$ and $t = 1, ..., T$.
- Here consider basic linear models for short panels.
- Presentation follows Cameron and Trivedi (2010) chapter 8.

Outline

2. Panel Data Summary: Wages

- PSID wage data 1976-82 on 595 individuals. Balanced.
- Source: Baltagi and Khanti-Akom (1990).
  - Corrected version of Cornwell and Rupert (1998).
- Goal: estimate causative effect of education on wages.
- Complication: education is time-invariant in these data.
  - Rules out fixed effects estimation for coefficient of education.
  - Need to use IV methods (Hausman-Taylor).
  - These more advanced methods are presented in day 2.

Reading in Panel Data

- Data organization may be
  - long form: each observation is an individual-time (i, t) pair
  - wide form: each observation is data on i for all time periods
  - wide form: each observation is data on t for all individuals
- `xt` commands require data in long form
  - use `reshape long` command to convert from wide to long form
  - see Cameron and Trivedi (2010) chapter 8.11.
- Data here are already in long form.

```
.  * Read in data set
.  use mus08psidextract.dta, clear
(PSID wage data 1976-82 from Baltagi and Khanti-Akom
(1990))
```

Summarize Data using Non-panel Commands

```
. * Describe dataset
. describe

Contains data from mus08psidextract.dta
  obs:         4,165                          PSID wage data 1976-82 from Baltagi and Khanti-Akom
 vars:            15                          16 Aug 2007 16:29
 size:       283,220 (97.5% of memory free)   (_dta has notes)
-------------------------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------------------------
exp             float  %9.0g                  years of full-time work experience
wks             float  %9.0g                  weeks worked
occ             float  %9.0g                  occupation; occ==1 if in a blue-collar occupation
ind             float  %9.0g                  industry; ind==1 if working in a manufacturing indu
south           float  %9.0g                  residence; south==1 if in the South area
smsa            float  %9.0g                  smsa==1 if in the Standard metropolitan statistical
ms              float  %9.0g                  marital status
fem             float  %9.0g                  female or male
union           float  %9.0g                  if wage set be a union contract
ed              float  %9.0g                  years of education
blk             float  %9.0g                  black
lwage           float  %9.0g                  log wage
id              float  %9.0g
t               float  %9.0g
exp2            float  %9.0g
-------------------------------------------------------------------------------------------------
```

- Summary statistics combine variation over i and t.

```
. * Summarize dataset
. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
exp	4165	19.85378	10.96637	1	51
wks	4165	46.81152	5.129098	5	52
occ	4165	.5111645	.4999354	0	1
ind	4165	.3954382	.4890033	0	1
south	4165	.2902761	.4539442	0	1
smsa	4165	.6537815	.475821	0	1
ms	4165	.8144058	.3888256	0	1
fem	4165	.112605	.3161473	0	1
union	4165	.3639856	.4812023	0	1
ed	4165	12.84538	2.787995	4	17
blk	4165	.0722689	.2589637	0	1
lwage	4165	6.676346	.4615122	4.60517	8.537
id	4165	298	171.7821	1	595
t	4165	4	2.00024	1	7
exp2	4165	514.405	496.9962	1	2601

- Since 4165 $(= 7 \times 595)$ observations for all variables the dataset is balanced and complete.

- Listing the first few observations is useful

```
. * Organization of data set
. list id t exp wks occ in 1/3, clean

       id   t   exp   wks   occ
  1.    1   1     3    32     0
  2.    1   2     4    43     0
  3.    1   3     5    40     0
```

- Data are in long form, sorted by `id` and then by `t`

Stata Commands for Panel Data Summary

- Commands `describe`, `summarize` and `tabulate` confound cross-section and time series variation.
- Instead use specialized panel commands after `xtset`:
  - `xtdescribe`: extent to which panel is unbalanced
  - `xtsum`: separate within (over time) and between (over individuals) variation
  - `xttab`: tabulations within and between for discrete data e.g. binary
  - `xttrans`: transition frequencies for discrete data
  - `xtline`: time series plot for each individual on one chart
  - `xtdata`: scatterplots for within and between variation.

Summarize Data using Panel Commands

- xtset command defines i and t.
  - Allows use of panel commands and some time series operators

```
.  * Declare individual identifier and time identifier
.  xtset id t
 panel variable:  id (strongly balanced)
 time variable:  t, 1 to 7
 delta:  1 unit
```

- `xtdescribe` command summarizes number of time periods each individual is observed.

```
. * Panel description of data set
. 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
```

- Data are balanced with every individual i having 7 time periods of data.

- `xtsum` command splits overall variation into
  - between variation: variation in $\bar{x}_i = T_i^{-1} \sum_i x_{it}$ across individuals
  - within variation: variation in x_{it} around $\bar{x}_i$

```
. * Panel summary statistics: within and between variation
. xtsum lwage exp ed t

Variable         |      Mean   Std. Dev.       Min        Max |    Observations
-----------------+--------------------------------------------+----------------
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
                 |                                            |
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
                 |                                            |
t        overall |         4    2.00024          1          7 |     N =    4165
         between |                    0          4          4 |     n =     595
         within  |              2.00024          1          7 |     T =       7
```

- For time-invariant variable `ed` the within variation is zero.
  For individual-invariant variable `t` the between variation is zero.
  For `lwage` the within variation $<$ between variation.

- xttab command provides more detail for discrete-valued variable.

```
. * Panel tabulation for a variable
. xttab south

                  Overall             Between            Within
    south |    Freq.  Percent      Freq.  Percent        Percent
----------+-----------------------------------------------------
        0 |     2956    70.97        428    71.93          98.66
        1 |     1209    29.03        182    30.59          94.90
----------+-----------------------------------------------------
    Total |     4165   100.00        610   102.52          97.54
                                (n = 595)
```

- 29.03% on average were in the south.
  30.59% were ever in the south.
  94.9% of those ever in the south were always in the south.

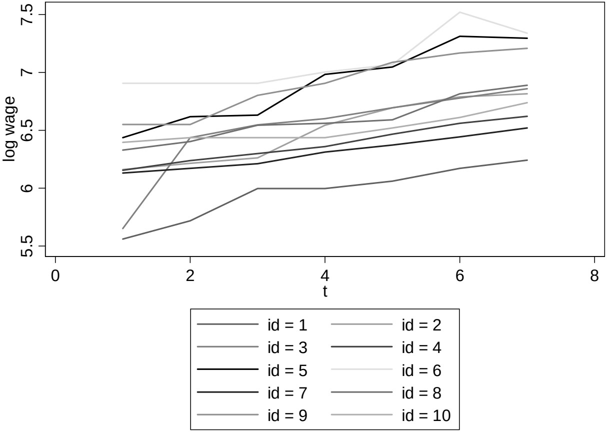
- `xttrans` provides transition probabilities for discrete-valued variable.

```
. * Transition probabilities for a variable
. xttrans south, freq

residence; |
  south==1 |   residence; south==1
  if in the |  if in the South area
South area |         0          1 |     Total
-----------+----------------------+----------
         0 |     2,527          8 |     2,535 
           |     99.68       0.32 |    100.00 
-----------+----------------------+----------
         1 |         8      1,027 |     1,035 
           |      0.77      99.23 |    100.00 
-----------+----------------------+----------
     Total |     2,535      1,035 |     3,570 
           |     71.01      28.99 |    100.00 
```

- For the 28.99% of the sample ever in the south, 99.23% remained in the south the next period.

```
.   * Time series plots of log wage for first 10 individuals
.   xtline lwage if id<=10, overlay
```



- Much autocorrelation in each person's wage.

Autocorrelations

- Because `xtset` set a time variable can use time series commands
  - `Lj.x` gives `x` lagged j periods.
- Can compute autocorrelations for a variable.

```
. * First-order autocorrelation in a variable
. sort id t

. correlate lwage L.lwage L2.lwage L3.lwage L4.lwage L5.lwage L6.lwage
(obs=595)
```

	lwage	L. lwage	L2. lwage	L3. lwage	L4. lwage	L5. lwage	L6. lwage
lwage							
--.	1.0000						
L1.	0.9238	1.0000					
L2.	0.9083	0.9271	1.0000				
L3.	0.8753	0.8843	0.9067	1.0000			
L4.	0.8471	0.8551	0.8833	0.8990	1.0000		
L5.	0.8261	0.8347	0.8721	0.8641	0.8667	1.0000	
L6.	0.8033	0.8163	0.8518	0.8465	0.8594	0.9418	1.0000

- High serial correlation: $\text{Cor}[y_t, y_{t-6}] = 0.80$ and not AR(1).
- Note that estimated autocorrelations without imposing stationarity.
- Weakness: Only 595 observations used as needed `L6.lwage`

Autocorrelations - better

- Command `pwcorr` uses all the available data

```
. * The previous uses few observations as needs L6.lwage
. * The following uses as much data as is available
. pwcorr lwage L.lwage L2.lwage L3.lwage L4.lwage L5.lwage L6.lwage
```

	lwage	L.lwage	L2.lwage	L3.lwage	L4.lwage	L5.lwage	L6.lwage
lwage	1.0000						
L.lwage	0.9189	1.0000					
L2.lwage	0.8858	0.9128	1.0000				
L3.lwage	0.8649	0.8748	0.9044	1.0000			
L4.lwage	0.8460	0.8565	0.8684	0.8988	1.0000		
L5.lwage	0.8220	0.8444	0.8602	0.8631	0.8944	1.0000	
L6.lwage	0.8033	0.8163	0.8518	0.8465	0.8594	0.9418	1.0000

Autocorrelations - nonstationary

- The preceding imposed stationarity: $\text{Corr}[y_{it}, y_{it-j}] = \text{Corr}[y_{it+k}, y_{it+k-j}]$.
- Following does not constrain e.g. correlation between years 1 and 2 to equal that between 2 and 3 (so no longer stationary)

```
. * First-order autocorrelation differs in different year pairs
. forvalues s = 2/7 {
  2.      quietly corr lwage L1.lwage if t == `s'
  3.      display "Autocorrelation at lag 1 in year `s' = " %6.3f r(rho)
  4.      }
Autocorrelation at lag 1 in year 2 =  0.942
Autocorrelation at lag 1 in year 3 =  0.867
Autocorrelation at lag 1 in year 4 =  0.899
Autocorrelation at lag 1 in year 5 =  0.907
Autocorrelation at lag 1 in year 6 =  0.927
Autocorrelation at lag 1 in year 7 =  0.924
```

3. Pooled OLS (a Population-Averaged Estimator)

- Pooled OLS is regular OLS of y_{it} on $\mathbf{x}_{it}$
  - Consistent if $\mathbf{x}_{it}$ is uncorrelated with the error u_{it}.

```
. * Pooled OLS with incorrect default standard errors
. regress lwage exp exp2 wks ed, noheader
```

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

- Important: The default standard errors are too small
  - they erroneously assume errors are independent over t for given i.
  - this assumes more information content from data then is the case.

Cluster-Robust Standard Errors

- Should instead use cluster-robust standard errors

```
. * Pooled OLS with cluster-robust standard errors
. regress lwage exp exp2 wks ed, noheader vce(cluster id)
                              (Std. Err. adjusted for 595 clusters in id)
```

lwage	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
exp	.044675	.0054385	8.21	0.000	.0339941	.055356
exp2	-.0007156	.0001285	-5.57	0.000	-.0009679	-.0004633
wks	.005827	.0019284	3.02	0.003	.0020396	.0096144
ed	.0760407	.0052122	14.59	0.000	.0658042	.0862772
_cons	4.907961	.1399887	35.06	0.000	4.633028	5.182894

- Cluster-robust standard errors here are twice as large as default! Cluster-robust t-statistics are half as large as default!
- Typical result. Need to use cluster-robust se's if use pooled OLS.

Pooled FGLS (a Population-Averaged Estimator)

- Do more efficient Feasible GLS assuming an error correlation model (over time) such as AR(1) or equicorrelated.
  - Here specify AR(2): $u_{it} = \rho_1 u_{i,t-1} + \rho_2 u_{i,t-2} + \varepsilon_{it}$

```
. xtreg lwage exp exp2 wks ed, pa corr(ar 2) vce(robust) nolog

GEE population-averaged model                   Number of obs      =      4165
Group and time vars:                 id t       Number of groups   =       595
Link:                            identity       Obs per group: min =         7
Family:                          Gaussian                      avg =       7.0
Correlation:                        AR(2)                      max =         7
                                                Wald chi2(4)       =    873.28
Scale parameter:                 .1966639       Prob > chi2        =    0.0000

                                   (Std. Err. adjusted for clustering on id)
------------------------------------------------------------------------------
             |             Semirobust
       lwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .0718915    .003999    17.98   0.000     .0640535    .0797294
        exp2 |  -.0008966   .0000933    -9.61   0.000    -.0010794   -.0007137
         wks |   .0002964   .0010553     0.28   0.779     -.001772    .0023647
          ed |   .0905069   .0060161    15.04   0.000     .0787156    .1022982
       _cons |   4.526381   .1056897    42.83   0.000     4.319233    4.733529
------------------------------------------------------------------------------
```

  - Same as `xtgee lwage $xlist, pa corr(ar 2) vce(robust) nolog`

4. Linear Short Panels Overview: Basic considerations

1. Regular time intervals assumed.
2. Unbalanced panel okay (`xt` commands handle unbalanced data). [Should then rule out selection/attrition bias].
3. Short panel assumed, with T small and $N \to \infty$. [Versus long panels, with $T \to \infty$ and N small or $N \to \infty$.]
4. Errors are correlated. [For short panel: correlated over t for given i, but not over i.]
5. Parameters may vary over individuals or time. Intercept: Individual-specific effects model (fixed or random effects). Slopes: Pooling and random coefficients models.
6. Regressors: time-invariant, individual-invariant, or vary over both.
7. Prediction. [Not always possible even if marginal effects computed.]
8. Dynamic models: possible. [Usually static models are estimated.]

Summary of Basic Linear Short Panel Models

- Pooled model (or population-averaged)

$$y_{it} = \alpha + \mathbf{x}_{it}'\boldsymbol{\beta} + u_{it}.$$

- Two-way effects model allows intercept to vary over i and t

$$y_{it} = \alpha_i + \gamma_t + \mathbf{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it}.$$

- Individual-specific effects model

$$y_{it} = \alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it},$$

  where α_i may be fixed effect or random effect
  and $\mathbf{x}_{it}$ may include time dummies.

- Mixed model or random coefficients model allows slopes to vary over i

$$y_{it} = \alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta}_i + \varepsilon_{it}.$$

Fixed effects versus random effects model

- Individual-specific effects model:

$$y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + (\alpha_i + \varepsilon_{it}).$$

- Fixed effects (FE):
  - α_i is a random variable possibly correlated with $\mathbf{x}_{it}$
  - so regressor $\mathbf{x}_{it}$ may be endogenous (wrt to α_i but not ε_{it})
    e.g. education (x_{it}) is correlated with time-invariant ability (α_i)
  - pooled OLS, pooled GLS, RE are inconsistent for $\boldsymbol{\beta}$
  - within (FE) and first difference estimators are consistent.
- Random effects (RE) or population-averaged (PA):
  - α_i is purely random (usually iid $(0, \sigma_\alpha^2)$) unrelated to $\mathbf{x}_{it}$
  - so regressor $\mathbf{x}_{it}$ is exogenous
  - all estimators are consistent for $\boldsymbol{\beta}$.
- Fundamental divide: microeconometricians FE versus others RE.

Fixed effects model

- The FE model allows consistent estimation of $\boldsymbol{\beta}$ even if $\mathbf{x}_{it}$ is correlated with the error, without the need to instrument!
- How? It is assumed that $\mathbf{x}_{it}$ is correlated only with the time-invariant component of the error
  - e.g. time-invariant unobserved ability in an earnings - schooling panel regression.
- In principal there is an incidental parameters problem
  - k $\boldsymbol{\beta}'s$ and N $\alpha_i's$ to estimate with $N \rightarrow \infty$
  - in linear models okay as can difference out α_i
  - but this is not possible in most nonlinear models.
- There is also great loss in estimation precision
  - only within variation is used and data may have little within variation
    - ★ extreme case: time-invariant regressor not identified.

Summary of Linear Panel Estimators

- Pooled or population-averaged estimator:
  - OLS of y_{it} on intercept and $\mathbf{x}_{it}$
  - FGLS of y_{it} on intercept and $\mathbf{x}_{it}$ where assume e.g. equicorrelated errors or AR(1) errors.
- Within or fixed effects estimator
  - OLS of $(y_{it} - \bar{y}_i)$ on $(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)$.
- Between estimator
  - OLS of $\bar{y}_i$ on $\bar{\mathbf{x}}_i$
- Random effects estimator
  - OLS of $(y_{it} - \widehat{\lambda}\bar{y}_i)$ on $(1 - \widehat{\lambda})\mu$ and $(\mathbf{x}_{it} - \widehat{\lambda}\bar{\mathbf{x}}_i)$
  - $\widehat{\lambda}$ is consistent for $\lambda = 1 - \sigma_\varepsilon / \sqrt{\sigma_\varepsilon^2 + T\sigma_\alpha^2}$.
- First difference estimator
  - OLS of Δy_{it} on $\Delta \mathbf{x}_{it}$.

Panel Robust Inference

- There are many different panel estimators.
- Many methods assume ε_{it} and α_i (if present) are iid.
  - This yields wrong standard errors if errors are heteroskedastic or if errors are not equicorrelated over time for a given individual.
- Instead, for a short panel can relax assumptions and use cluster-robust inference.
  - This allows heteroskedasticity and general correlation over time for given i.
  - Independence over i is still assumed.
- Implementation in Stata
  - For `xtreg` use option `vce(robust)` does cluster-robust
  - For some other `xt` commands use option `vce(cluster id)`
  - And for some other `xt` commands there is no option, but may be able to do a cluster bootstrap.

Panel Robust Inference (continued)

- For simplicity, consider OLS

$$y_{it} = \mathbf{x}_{it}'\beta + u_{it}.$$

- Stack all years for given i

$$\begin{bmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{i1}' \\ \vdots \\ \mathbf{x}_{iT}' \end{bmatrix} \beta + \begin{bmatrix} u_{i1} \\ \vdots \\ u_{iT} \end{bmatrix}$$

$$\underset{T\times 1}{\mathbf{y}_i} = \underset{T\times k}{\mathbf{X}_i}\underset{k\times 1}{\beta} + \underset{T\times 1}{\mathbf{u}_i}$$

Panel Robust Inference (continued)

- Then stack over all years

$$\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_T \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_T \end{bmatrix} \beta + \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_T \end{bmatrix}$$

$$\underset{NT\times 1}{\mathbf{y}} = \underset{NT\times k}{\mathbf{X}} \underset{k\times 1}{\beta} + \underset{NT\times 1}{\mathbf{u}}$$

- Then three equivalent representations

$$\begin{aligned} \widehat{\beta}_{\text{OLS}} &= [\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'\mathbf{y} \\ &= \left[\sum_{i=1}^{N} \mathbf{X}_i{}'\mathbf{X}_i\right]^{-1} \sum_i \mathbf{X}_i{}'\mathbf{y}_i \\ &= \left[\sum_{i=1}^{N} \sum_{t=1}^{T} \mathbf{x}_{it}\mathbf{x}_{it}'\right]^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbf{x}_{it} y_{it}. \end{aligned}$$

Panel Robust Inference (continued)

- Usual algebra yields

$$\begin{aligned}\widehat{\beta}_{\text{OLS}} &= \beta + [\mathbf{X}'\mathbf{X}]^{-1}\mathbf{X}'\mathbf{u} \\ \text{Var}[\widehat{\beta}] &= [\mathbf{X}'\mathbf{X}]^{-1}\text{Var}[\mathbf{X}'\mathbf{u}][\mathbf{X}'\mathbf{X}]^{-1}\end{aligned}$$

- Old approach: $\text{Var}[\mathbf{X}'\mathbf{u}] = \mathbf{X}'\text{Var}[\mathbf{u}]\mathbf{X}$ and estimate $\text{Var}[\mathbf{u}]$
- New approach (White (1980)): Estimate $k \times k$ $\text{Var}[\mathbf{X}'\mathbf{u}]$
- For panel robust use independence over i :

$$\begin{aligned}\text{Var}[\mathbf{X}'\mathbf{u}] &= \text{Var}\left[\textstyle\sum_{i=1}^{N} \mathbf{X}_i{}'\mathbf{u}_i\right] \\ &= \textstyle\sum_{i=1}^{N} \text{Var}\left[\mathbf{X}_i{}'\mathbf{u}_i\right] \\ &= \textstyle\sum_{i=1}^{N} \text{E}[\mathbf{X}_i{}'\mathbf{u}_i\mathbf{u}_i'\mathbf{X}_i]\end{aligned}$$

- So use

$$\widehat{\text{Var}}[\mathbf{X}'\mathbf{u}] = \mathbf{X}_i{}'\widehat{\mathbf{u}}_i\widehat{\mathbf{u}}_i'\mathbf{X}_i$$

  - where $\widehat{\mathbf{u}}_i = \mathbf{y}_i - \mathbf{X}_i'\widehat{\beta}$.

Panel Robust Inference (continued)

- Panel robust variance estimate that controls for both serial correlation and heteroskedasticity is

$$\widehat{\mathrm{V}}[\widehat{\beta}_{\mathrm{OLS}}] = \left[\sum_{i=1}^{N} \mathbf{X}_i{}'\mathbf{X}_i\right]^{-1} \sum_{i=1}^{N} \mathbf{X}_i{}'\widehat{\mathbf{u}}_i\widehat{\mathbf{u}}_i'\mathbf{X}_i \left[\sum_{i=1}^{N} \mathbf{X}_i{}'\mathbf{X}_i\right]^{-1},$$

  - where $\widehat{\mathbf{u}}_i = \mathbf{y}_i - \mathbf{X}_i'\widehat{\beta}$
  - need $N \rightarrow \infty$ and independence over i.

- Literature:
  - Generalizes White (1980) for heteroskedasticity
  - Liang and Zeger (1986) proposed for biostatistics models
  - Arellano (1989) proposed for fixed effects estimator
  - Rogers (1983) popularized through Stata option `vce(cluster)`.

Panel Robust Inference (continued)

- Now extend to RE, FE ... estimators.
- These are OLS on transformed data $(y_{it}^*, \mathbf{w}_{it}^*)$ where e.g. $y_{it}^* = y_{it} - \bar{y}_i$.

$$\begin{aligned} y_{it}^* &= \mathbf{w}_{it}^{*\prime}\theta + u_{it}^* \\ \mathbf{y}_i^* &= \mathbf{W}_i^*\theta + \mathbf{u}_i^* \end{aligned}$$

- Then similar to OLS of y_{it} on $\mathbf{x}_{it}$, we get

$$\widehat{\mathrm{V}}[\widehat{\theta}_{\mathrm{OLS}}] = \left[\sum\nolimits_{i=1}^N \mathbf{W}_i^{*\prime}\mathbf{W}_i^*\right]^{-1} \sum\nolimits_{i=1}^N \mathbf{W}_i^{*\prime}\widehat{\mathbf{u}}_i^*\widehat{\mathbf{u}}_i^{*\prime}\mathbf{W}_i^* \left[\sum\nolimits_{i=1}^N \mathbf{W}_i^{*\prime}\mathbf{W}_i^*\right]^{-1},$$

  - where $\widehat{\mathbf{u}}_i^* = \mathbf{y}_i^* - \mathbf{W}_i^*\widehat{\theta}$
  - need $N \rightarrow \infty$ and independence over i.
- An equivalent expression is
  - $\widehat{\mathrm{V}}[\widehat{\theta}_{\mathrm{OLS}}] =$ $\left[\sum_i \sum_t \mathbf{w}_{it}^*\mathbf{w}_{it}^{*\prime}\right]^{-1} \sum_i \sum_{t=1}^T \sum_{s=1}^T \mathbf{w}_{it}^*\mathbf{w}_{is}^{*\prime} u_{it}^* u_{is}^* \left[\sum_i \sum_t \mathbf{w}_{it}^*\mathbf{w}_{it}^{*\prime}\right]^{-1}.$

5. Fixed Effects Estimator

- Mean-differencing eliminates α_i

$$\begin{aligned} & y_{it} = \alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it} \\ \Rightarrow \quad & \bar{y}_i = \alpha_i + \bar{\mathbf{x}}_i'\boldsymbol{\beta} + \bar{\varepsilon}_i \\ \Rightarrow \quad & (y_{it} - \bar{y}_i) = (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)'\boldsymbol{\beta} + (\varepsilon_{it} - \bar{\varepsilon}_i) \end{aligned}$$

- The within or fixed effects estimator is OLS of $(y_{it} - \bar{y}_i)$ on $(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)$
  - Efficiency loss as use only within variation
  - Coefficient of any time-invariant regressor is not identified $(x_{it} = \overline{x}_i)$
  - Use cluster-robust standard errors
  - Stata command `xtreg, fe`

- Within or FE estimates:

```
. * Within or FE estimator with cluster-robust standard errors
. xtreg lwage exp exp2 wks ed, fe vce(robust)

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(3,594)           =   1059.72
corr(u_i, Xb)  = -0.9107                        Prob > F           =    0.0000

                                    (Std. Err. adjusted for 595 clusters in id)
------------------------------------------------------------------------------
             |               Robust
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .1137879   .0040289    28.24   0.000     .1058753    .1217004
        exp2 |  -.0004244   .0000822    -5.16   0.000    -.0005858   -.0002629
         wks |   .0008359   .0008697     0.96   0.337    -.0008721    .0025439
          ed |  (dropped)
       _cons |   4.596396   .0600887    76.49   0.000     4.478384    4.714408
-------------+----------------------------------------------------------------
     sigma_u |  1.0362039
     sigma_e |  .15220316
         rho |  .97888036   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

- Variable ed is not identified as time-invariant regressor in this dataset.

Fixed Effects Estimator (continued)

- Several ways to compute FE estimator aside from `xtreg, fe`.
- Least squares dummy variables:
  - $d_{ji,t}$ for $j = 1, ..., N$ are N dummies equal to 1 if $i = j$
  - Estimate directly using `regress` or use areg

$$y_{it} = \sum_{j=1}^{N} \alpha_i d_{ji,t} + \mathbf{x}_{it}' \boldsymbol{\beta} + \varepsilon_{it}$$

```
* FE model fitted as LSDV using areg
areg lwage exp exp2 wks ed, absorb(id) vce(cluster id)

* FE model fitted using LSDV using regress
set matsize 800
quietly xi:  regress lwage exp exp2 wks ed i.id,
vce(cluster id)
estimates table, keep(exp exp2 wks ed _cons) b se b(%12.7f)
```

Mundlak/Chamberlain Model

- Mundlak and Chamberlain suppose the fixed effects

$$\alpha_i = \bar{\mathbf{x}}_i' \boldsymbol{\pi} + error.$$

- So OLS regress y_{it} on intercept $\mathbf{x}_{it}$ and $\bar{\mathbf{x}}_i$
- Yields same $\boldsymbol{\beta}$ estimate as the FE estimator.

```
* FE model fitted by add mean of x as a regressor
global xlist exp exp2 wks ed
sort id
foreach x of varlist $xlist {
   by id:  egen mean`x' = mean(`x')
   }
regress lwage exp exp2 wks ed mean*, vce(robust)
```

6. Between Estimator

- OLS of $\bar{y}_i$ on intercept and $\bar{\mathbf{x}}_i$
  - `xtreg, be` has no heteroskedastic robust option but can bootstrap.

```
. *******  BETWEEN ESTIMATOR
.
. * Between estimator with default standard errors
. xtreg lwage exp exp2 wks ed, 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
------------------------------------------------------------------------------
```

7. Random Effects Estimator

- Random effects estimator is FGLS estimator for the RE model

$$\begin{aligned} y_{it} &= \alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it} \\ \alpha_i &\sim \text{i.i.d.}[\alpha, \sigma_\alpha^2] \\ \varepsilon_{it} &\sim \text{i.i.d.}[0, \sigma_\varepsilon^2] \end{aligned}$$

- The RE model implies equicorrelated (or exchangeable) errors
  - $\text{Var}[\alpha_i + \varepsilon_{it}] = \text{Var}[\alpha_i] + \text{Var}[\varepsilon_{it}] = \sigma_\alpha^2 + \sigma_\varepsilon^2$
  - For $s \neq t$, $\text{Cov}[\alpha_i + \varepsilon_{it}, \alpha_i + \varepsilon_{is}] = \text{Cov}[\alpha_i, \alpha_i] = \sigma_\alpha^2$
  - So $\text{Corr}[\alpha_i + \varepsilon_{it}, \alpha_i + \varepsilon_{is}] = \sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_\varepsilon^2)$ for all $s \neq t$.
- FGLS can be shown to equal OLS in the transformed model

$$(y_{it} - \widehat{\theta}_i \bar{y}_i) = (\mathbf{x}_{it} - \widehat{\theta}_i \bar{\mathbf{x}}_i)'\boldsymbol{\beta} + error,$$

where $\widehat{\theta}_i$ is a consistent estimate of $\theta_i = 1 - \sqrt{\sigma_\varepsilon^2/(T_i\sigma_\alpha^2 + \sigma_\varepsilon^2)}$.

- Random effects estimates:

```
. * Random effects estimator with cluster-robust standard errors
. xtreg lwage exp exp2 wks ed, re vce(robust) 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)       =   1598.50
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000
theta          = .82280511

                                  (Std. Err. adjusted for 595 clusters in id)
------------------------------------------------------------------------------
             |               Robust
       lwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .0888609   .0039992    22.22   0.000     .0810227    .0966992
        exp2 |  -.0007726   .0000896    -8.62   0.000    -.0009481    -.000597
         wks |   .0009658   .0009259     1.04   0.297     -.000849    .0027806
          ed |   .1117099   .0083954    13.31   0.000     .0952552    .1281647
       _cons |   3.829366   .1333931    28.71   0.000     3.567921    4.090812
-------------+----------------------------------------------------------------
     sigma_u |  .31951859
     sigma_e |  .15220316
         rho |  .81505521   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

- Option theta gives $\widehat{\theta} = 0.82 = 1 - \sqrt{0.152^2/(7 \times 0.319^2 + 0.152^2)}$.

8. Fixed versus Random Effects Estimators

- RE has advantages: estimates all parameters & may be more efficient.
  - But RE is inconsistent if fixed effects present.
- Use Hausman test to discriminate between FE and RE.
  - This tests difference between FE and RE estimates is statistically significantly different from zero.
- Do not use `hausman` command – it requires that RE estimator is fully efficient (see next slide).
- Instead do one of the following
  - 1. Do a panel bootstrap of the Hausman test.
  - 2. Do the Wooldridge (2002) robust version of Hausman test.
    - ★ Test $H_0 : \gamma = \mathbf{0}$ in the auxiliary OLS regression

$$(y_{it} - \widehat{\theta}\bar{y}_i) = (1 - \widehat{\theta})\alpha + (\mathbf{x}_{it} - \widehat{\theta}\bar{\mathbf{x}}_i)'\boldsymbol{\beta} + (\mathbf{x}_{1it} - \bar{\mathbf{x}}_{1i})'\boldsymbol{\gamma} + v_{it},$$

      where $\mathbf{x}_{1i} \subset \mathbf{x}_i$ denotes time-varying regressors only.
    - ★ Use cluster-robust standard errors for this test.
    - ★ Stata add-on `xtoverid` after `xtreg,re` does this.

Hausman Test Theory

- Hausman test compares to estimators $\widehat{\boldsymbol{\theta}}$ and $\widetilde{\boldsymbol{\theta}}$
  - Test $H_0 : \text{plim}(\widehat{\boldsymbol{\theta}} - \widetilde{\boldsymbol{\theta}}) = 0$ against $H_a : \text{plim}(\widehat{\boldsymbol{\theta}} - \widetilde{\boldsymbol{\theta}}) \neq 0$.
  - e.g. OLS versus 2SLS with possible endogenous regressor
  - e.g. RE versus FE with possible fixed effect.
- Under H_0, as usual $(\widehat{\boldsymbol{\theta}} - \widetilde{\boldsymbol{\theta}}) \overset{a}{\sim} N[\mathbf{0}, \mathsf{V}[\widehat{\boldsymbol{\theta}} - \widetilde{\boldsymbol{\theta}}])$.
- So form χ^2 statistic: $H = (\widehat{\boldsymbol{\theta}} - \widetilde{\boldsymbol{\theta}})' \left[\mathsf{V}[\widehat{\boldsymbol{\theta}} - \widetilde{\boldsymbol{\theta}}] \right]^{-1} (\widehat{\boldsymbol{\theta}} - \widetilde{\boldsymbol{\theta}})$
  - reject H_0 if $H > \chi^2$ critical value.
- Problem: To implement we need estimate of $\mathsf{V}[\widehat{\boldsymbol{\theta}} - \widetilde{\boldsymbol{\theta}}]$.
- Hausman (1978) assumed $\widehat{\boldsymbol{\theta}}$ is fully efficient under H_0
  - then $\text{Cov}[\widehat{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\theta}}] = \text{Var}[\widehat{\boldsymbol{\theta}}]$
  - implying $\mathsf{V}[\widehat{\boldsymbol{\theta}} - \widetilde{\boldsymbol{\theta}}] = \mathsf{V}[\widehat{\boldsymbol{\theta}}] + \mathsf{V}[\widetilde{\boldsymbol{\theta}}] - 2 \times \mathsf{V}[\widehat{\boldsymbol{\theta}}] = \mathsf{V}[\widehat{\boldsymbol{\theta}}] - \mathsf{V}[\widetilde{\boldsymbol{\theta}}]$.
  - but we rarely have $\widehat{\boldsymbol{\theta}}$ fully efficient.

Hausman Test Wrong

```
. * Wrong Hausman test assuming RE estimator is fully efficient under null hypothesi
. hausman FE_def RE_def, sigmamore

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     FE_def       RE_def        Difference          S.E.
-------------+----------------------------------------------------------------
         exp |    .1137879     .0888609        .0249269        .0012778
        exp2 |   -.0004244    -.0007726        .0003482        .0000285
         wks |    .0008359     .0009658       -.0001299        .0001108
------------------------------------------------------------------------------
                           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)
                          =     1513.02
                Prob>chi2 =      0.0000
```

Hausman Test Correct

- Following is manual

```
. * Correct Robust Hausman test using method of Wooldridge (2002)
. global xlist exp exp2 wks

. foreach x of varlist $xlist {
  2.     by id: egen mean`x' = mean(`x')
  3.     }

. quietly regress lwage exp exp2 wks meanexp meanexp2 meanwks, vce(cluster id)

. test meanexp meanexp2 meanwks

 ( 1)  meanexp = 0
 ( 2)  meanexp2 = 0
 ( 3)  meanwks = 0

       F(  3,   594) =  630.59
            Prob > F =    0.0000
```

- Get exactly same result using simpler

```
quietly regress lwage exp exp2 wks ///
meanexp meanexp2 meanwks, vce(cluster id)
test meanexp meanexp2 meanwks
```

- Can also use Stata add-on `xtoverid` after `xtreg,re`

9. First Difference Estimator

- First-differencing eliminates α_i

$$\begin{array}{rrcl} & y_{it} & = & \alpha_i + \mathbf{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it} \\ \Rightarrow & y_{i,t-1} & = & \alpha_i + \mathbf{x}_{i,t-1}'\boldsymbol{\beta} + \varepsilon_{i,t-1} \\ \Rightarrow & (y_{it} - y_{i,t-1}) & = & (\mathbf{x}_{it} - \mathbf{x}_{i,t-1})'\boldsymbol{\beta} + (\varepsilon_{it} - \varepsilon_{i,t-1}) \end{array}$$

- First differences estimator
  - OLS regression of Δy_{it} on $\Delta \mathbf{x}_{it}$, i.e. use first differences.
  - Coefficient of any time-invariant regressor is not identified $(x_{it} = x_{i,t-1})$.
- Not used much for basic FE model
  - FE estimator is fully efficient if ε_{it} is iid $(0, \sigma_\varepsilon^2)$
  - FD estimator is fully efficient if $\varepsilon_{it} = \varepsilon_{it-1} + v_{it}$ where v_{it} is iid $(0, \sigma_v^2)$
  - FE=FD if $T = 2$ as then $y_{i2} - \bar{y}_i = y_{i2} - \frac{y_{i1}+y_{i2}}{2} = (y_{i2} - y_{i1})/2$.

First Difference Estimator (continued)

- No direct Stata command.
- Can `regress D.(lwage $xlist), vce(cluster id)`
- More comparable to FE is regress with noconstant
  `regress D.(lwage $xlist), noconstant vce(cluster id)`
  - FE is also noconstant, but then adds back in $\bar{y}$.

- First differences estimator

```
. regress d.lwage d.exp d.exp2 d.wks d.ed, noconstant vce(cluster id)
note: _delete omitted because of collinearity

Linear regression                               Number of obs =    3570
                                                F(  3,    594) = 1035.19
                                                Prob > F      =  0.0000
                                                R-squared     =  0.2209
                                                Root MSE      =  .18156

                                 (Std. Err. adjusted for 595 clusters in id)
```

D.lwage	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
exp D1.	.1170654	.0040974	28.57	0.000	.1090182	.1251126
exp2 D1.	-.0005321	.0000808	-6.58	0.000	-.0006908	-.0003734
wks D1.	-.0002683	.0011783	-0.23	0.820	-.0025824	.0020459
ed D1.	(omitted)					

10.Estimator comparison

```
. * Compare various estimators (with cluster-robust se's)
. global xlist exp exp2 wks ed
. quietly regress lwage $xlist, vce(cluster id)
. estimates store OLS
. quietly xtgee lwage exp exp2 wks ed, corr(ar 2)
   vce(robust)
. estimates store PFGLS
. quietly xtreg lwage $xlist, be
. estimates store BE
. quietly xtreg lwage $xlist, re vce(robust)
. estimates store RE
. quietly xtreg lwage $xlist, fe vce(robust)
. estimates store FE
. estimates table OLS PFGLS BE RE FE, b(%9.4f) se stats(N)
```

Variable	OLS	PFGLS	BE	RE	FE
exp	0.0447	0.0719	0.0382	0.0889	0.1138
	0.0054	0.0040	0.0057	0.0040	0.0040
exp2	-0.0007	-0.0009	-0.0006	-0.0008	-0.0004
	0.0001	0.0001	0.0001	0.0001	0.0001
wks	0.0058	0.0003	0.0131	0.0010	0.0008
	0.0019	0.0011	0.0041	0.0009	0.0009
ed	0.0760	0.0905	0.0738	0.1117	(omitted)
	0.0052	0.0060	0.0049	0.0084	
_cons	4.9080	4.5264	4.6830	3.8294	4.5964
	0.1400	0.1057	0.2101	0.1334	0.0601
N	4165	4165	4165	4165	4165

legend: b/se

- Coefficients vary considerably across OLS, RE, FE and RE estimators.
  - FE and RE similar as $\widehat{\theta} = 0.82 \simeq 1$.
- Not shown is that even for FE and RE cluster-robust changes se's.
- Coefficient of `ed` not identified for FE as time-invariant regressor!

Standard Errors Comparison

- Compares default to panel-robust standard errors for RE and FE.

Variable	RE_def	RE	FE_def	FE
exp	0.0889	0.0889	0.1138	0.1138
	0.0028	0.0040	0.0025	0.0040
exp2	-0.0008	-0.0008	-0.0004	-0.0004
	0.0001	0.0001	0.0001	0.0001
wks	0.0010	0.0010	0.0008	0.0008
	0.0007	0.0009	0.0006	0.0009
ed	0.1117	0.1117	(omitted)	(omitted)
	0.0061	0.0084		
_cons	3.8294	3.8294	4.5964	4.5964
	0.0936	0.1334	0.0389	0.0601
N	4165	4165	4165	4165

legend: b/se

11. Panel Bootstrap

- Do pairs bootstrap where resample $(y, \mathbf{x})$ over individuals i rather than observations (i, t)
- Do B iterations of this step. On the b^{th} iteration:
  - form a sample of G clusters $\{(\mathbf{y}_1^*, \mathbf{X}_1^*), ..., (\mathbf{y}_G^*, \mathbf{X}_G^*)\}$ by resampling with replacement G times from the original sample
  - obtain estimate $\widehat{\beta}_g$.
- Then $\widehat{\mathrm{V}}[\widehat{\beta}] = \frac{1}{B-1}\sum_{b=1}^{B}(\widehat{\beta}_b - \overline{\widehat{\beta}})(\widehat{\beta}_b - \overline{\widehat{\beta}})'$, $\quad \overline{\widehat{\beta}} = \frac{1}{B}\sum_{b=1}^{B}\widehat{\beta}_b$.
- In Stata use `cluster(id)` option of `bootstrap`
  - need to replace `xtset id t` with simply `xtset id`
  - and if FE then also add `idcluster(newid)` option
- This pairs panel (or clustered) bootstrap
  - yields essentially the same results as usual cluster-robust standard errors
  - is a bootstrap without asymptotic refinement.

Panel Bootstrap OLS Estimator

```
. * OLS with cluster(id) bootstrap
. bootstrap _b, reps(400) seed(10101) cluster(id) nodots: ///
>    regress lwage exp exp2 wks ed

Linear regression                               Number of obs      =      4165
                                                Replications       =       400
                                                Wald chi2(4)       =    305.95
                                                Prob > chi2        =    0.0000
                                                R-squared          =    0.2836
                                                Adj R-squared      =    0.2829
                                                Root MSE           =    0.3908

                                   (Replications based on 595 clusters in id)
```

lwage	Observed Coef.	Bootstrap Std. Err.	z	P>|z|	Normal-based [95% Conf.	Interval]
exp	.044675	.0051437	8.69	0.000	.0345936	.0547565
exp2	-.0007156	.0001192	-6.00	0.000	-.0009493	-.000482
wks	.005827	.0019416	3.00	0.003	.0020216	.0096324
ed	.0760407	.0049605	15.33	0.000	.0663182	.0857632
_cons	4.907961	.1382027	35.51	0.000	4.637088	5.178833

- Note: the following gives exactly the same
  - `xtreg lwage exp exp2 wks ed, ///`
    `pa corr(ind) vce(boot, reps(400) seed(10101))`

Panel Bootstrap RE Estimator

```
. * RE with cluster(id) bootstrap
. bootstrap _b, reps(400) seed(10101) cluster(id) nodots: ///
>     xtreg lwage exp exp2 wks ed, re

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)       =    767.06
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

                                  (Replications based on 595 clusters in id)
------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
       lwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .0888609   .0043459    20.45   0.000     .0803431    .0973788
        exp2 |  -.0007726   .0000939    -8.23   0.000    -.0009567   -.0005885
         wks |   .0009658   .0008768     1.10   0.271    -.0007527    .0026842
          ed |   .1117099    .007582    14.73   0.000     .0968495    .1265704
       _cons |   3.829366   .1330233    28.79   0.000     3.568645    4.090087
-------------+----------------------------------------------------------------
     sigma_u |  .31951859
     sigma_e |  .15220316
         rho |  .81505521   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

- Same as `xtreg lwage exp exp2 wks ed, /// re vce(boot, reps(400) seed(10101) nodots)`

Panel Bootstrap FE Estimator - add idcluster()

```
. * FE with cluster(id) bootstrap - add idcluster
. bootstrap _b, reps(400) seed(10101) cluster(id) idcluster(newid) ///
>     nodots: xtreg lwage exp exp2 wks ed, fe

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

                                                Wald chi2(3)       =   2787.50
corr(u_i, Xb)  = -0.9107                        Prob > chi2        =    0.0000

                                   (Replications based on 595 clusters in id)
------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
       lwage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         exp |   .1137879   .0041335    27.53   0.000     .1056863    .1218894
        exp2 |  -.0004244   .0000827    -5.13   0.000    -.0005865   -.0002622
         wks |   .0008359   .0008459     0.99   0.323    -.0008221    .0024938
          ed |  (omitted)
       _cons |   4.596396   .0711119    64.64   0.000      4.45702    4.735773
-------------+----------------------------------------------------------------
     sigma_u |  1.0362039
     sigma_e |  .15220316
         rho |  .97888036   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

- Same as `xtreg lwage exp exp2 wks ed, ///`
  `fe vce(boot, reps(400) seed(10101) nodots)`

Panel Jackknife

- An alternative re-sampling scheme is a leave-one-cluster-out jackknife.
- Let $\widehat{\boldsymbol{\beta}}_{-i}$ denote the estimator of $\boldsymbol{\beta}$ when the i^{th} cluster (here i^{th} if N individuals) is deleted

$$\widehat{\mathrm{V}}_{\text{jack;boot}}[\widehat{\boldsymbol{\beta}}] = \frac{N-1}{N} \sum_{i=1}^{N} (\widehat{\boldsymbol{\beta}}_{-i} - \overline{\widehat{\boldsymbol{\beta}}})(\widehat{\boldsymbol{\beta}}_{-i} - \overline{\widehat{\boldsymbol{\beta}}})',$$

where $\overline{\widehat{\boldsymbol{\beta}}} = \frac{1}{N} \sum_{i=1}^{N} \widehat{\boldsymbol{\beta}}_{-i}$.

- For Stata `xt` commands this is option `vce(jackknife)`

12. Summary of Stata Commands

- Linear panel estimators for short panels with exogenous regressors

Panel summary	`xtset; xtdescribe; xtsum; xtdata; xtline; xttab; xttran`
Pooled OLS	`regress`
Feasible GLS	`xtreg, pa` `xtgee, family(gaussian)`
Random effects	`xtreg, re; xtregar, re`
Fixed effects	`xtreg, fe; xtregar, fe`
Random slopes	`xtmixed; quadchk; xtrc` (later)
First differences	`regress` (with differenced data)