Econometrics: Multiple Regression and Applications

ECON4004: LAB 4

Duong Trinh

University of Glasgow

February 28, 2024

Intro

- Duong Trinh
  - PhD Student in Economics (Bayesian Microeconometrics)
  - Email: Duong.Trinh@glasgow.ac.uk
- ECON4004-LB01
  - Wednesday 10am -12 pm
  - 5 sessions (7-Feb, 14-Feb, 21-Feb, 28-Feb, 6-March)
  - ST ANDREWS:357
- ECON4004-LB02
  - Wednesday 12-2 pm
  - 5 sessions (7-Feb, 14-Feb, 21-Feb, 28-Feb, 6-March)
  - ST ANDREWS:357

Record Attendance

Plan for LAB 3

- Exercise 1: based on Stock & Watson, E10.1
- Exercise 2: based on Stock & Watson, E10.2

- We will focus on *"Panel Data - Fixed Effects Regressions"*

BRIEF REVIEW

Panel Data - What it looks like...

Panel data is a dataset in which the behavior of entities (i) are observed across time (t).

(X_{it}, Y_{it}),

$i = 1, \ldots, n; t = 1, \ldots, T$

These entities could be states, companies, families, individuals, countries, etc.

Entity	Year	Y	X1	X2	X3	.....
1	1	#	#	#	#	.....
1	2	#	#	#	#	.....
1	3	#	#	#	#	.....
:	:	:	:	:	:	:
2	1	#	#	#	#	.....
2	2	#	#	#	#	.....
2	3	#	#	#	#	.....
:	:	:	:	:	:	:
3	1	#	#	#	#	.....
3	2	#	#	#	#	.....
3	3	#	#	#	#	.....

Panel Data - The Long Form

Preparing data into panel data format:

Entity and Time in rows.
&
Variables in columns.

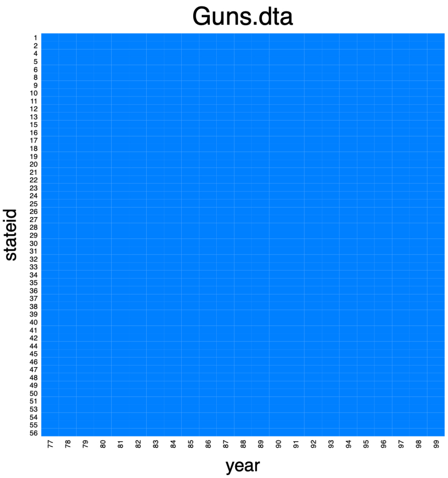
Entity	Year	Y	X1	X2	X3	.....
1	1	#	#	#	#	.....
1	2	#	#	#	#	.....
1	3	#	#	#	#	.....
:	:	:	:	:	:	:
2	1	#	#	#	#	.....
2	2	#	#	#	#	.....
2	3	#	#	#	#	.....
:	:	:	:	:	:	:
3	1	#	#	#	#	.....
3	2	#	#	#	#	.....
3	3	#	#	#	#	.....

Balanced Panel

All entities are observed across all times.

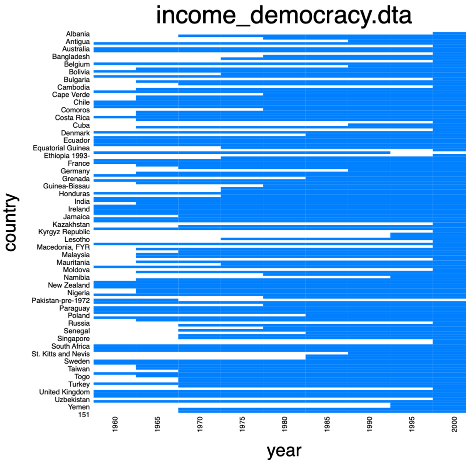
51 states $\times$ 23 years

$= 1173$ observations



Unbalanced Panel

Some entities are not observed in some time periods.



[SN] STATA command for Setting Data as Panel

Once the data is in long form, we need to set it as panel so we can use Stata's panel data `xt` commands.

```
*xtset entityid timeid
//'entityid' and 'timeid' have to be in numeric format
```

```
. use "Guns.dta", clear

. xtset stateid year

Panel variable: stateid (strongly balanced)
 Time variable: year, 77 to 99
         Delta: 1 unit
```

```
. use "income_democracy.dta", clear

. xtset code year

Panel variable: code (unbalanced)
 Time variable: year, 1960 to 2000, but with gaps
         Delta: 1 unit
```

<u>Note</u>: `id` means unique identifiers for entity (`entityid`) or for time period (`timeid`)

[SN] STATA command for Visualizing Panel Data

Once the data is set as panel, we can use a series of `xt` commands to analyze it.

For visualization

```
*xtline varname_of_inteterest
//graphs by entities
```

```
*xtline varname_of_inteterest, overlay legend(off)
//all entities in one graph
```

Fixed Effects Regressions - General Form

For $i = 1, \ldots, n$ and $t = 1, \ldots, T$

$$Y_{it} = \overbrace{\beta_1 X_{1,it} + \beta_2 X_{2,it} + \ldots + \beta_k X_{k,it}}^{\text{k regressor}} + \underbrace{\alpha_i}_{\substack{\text{entity}\\ \text{fixed effects}}} + \underbrace{\lambda_t}_{\substack{\text{time}\\ \text{fixed effects}}} + u_{it}$$

(3) Both E&T Fixed Effects: $Y_{it} = \beta_1 X_{it} + \underbrace{\alpha_i}_{\substack{\text{entity}\\ \text{fixed effects}}} + \underbrace{\lambda_t}_{\substack{\text{time}\\ \text{fixed effects}}} + u_{it}$

(1) Entity Fixed Effects: $Y_{it} = \beta_1 X_{it} + \alpha_i + u_{it}$

(2) Time Fixed Effects: $Y_{it} = \beta_1 X_{it} + \lambda_t + u_{it}$

(1) Entity Fixed Effects (Time-invariant)

$$Y_{it} = \beta_1 X_{it} + \overbrace{\alpha_i}^{\text{n entity-specific intercepts}} + u_{it}$$

$$\Downarrow$$

$$Y_{it} = \underbrace{\beta_0}_{\text{an intercept}} + \beta_1 X_{it} + \underbrace{\gamma_2 D2_i + \gamma_3 D3_i + \ldots + \gamma_n Dn_i}_{\text{with n-1 entity binary indicators}} + u_{it}$$

Entity **1**: $\alpha_1 = \beta_0$

Entity **2**: $\alpha_2 = \beta_0 + \gamma_2$

Entity **3**: $\alpha_3 = \beta_0 + \gamma_3$

$\vdots$

Entity **n**: $\alpha_n = \beta_0 + \gamma_n$

Note: This form suggests estimating regression model with $n-1$ binary indicators by OLS, but we can also use *"entity-demeaned" OLS algorithm*.

(2) Time Fixed Effects (Entity-invariant)

$$Y_{it} = \beta_1 X_{it} + \overbrace{\lambda_t}^{\text{T time-specific intercepts}} + u_{it}$$

$$\Downarrow$$

$$Y_{it} = \underbrace{\beta_0}_{\text{an intercept}} + \beta_1 X_{it} + \underbrace{\delta_2 B2_t + \delta_3 B3_t + \ldots + \delta_T BT_t}_{\text{with n-1 time binary indicators}} + u_{it}$$

Time period **1**: $\lambda_1 = \beta_0$

Time period **2**: $\lambda_2 = \beta_0 + \delta_2$

Time period **3**: $\lambda_3 = \beta_0 + \delta_3$

$\vdots$

Time period **T**: $\lambda_T = \beta_0 + \delta_T$

Note: This form suggests estimating regression model with $T - 1$ binary indicators by OLS, but we can also use *"time-demeaned" OLS algorithm*.

(3) Both Entity and Time (Two-way) Fixed Effects

$$Y_{it} = \beta_1 X_{it} + \overset{\substack{\text{entity} \\ \text{fixed effects}}}{\overbrace{\alpha_i}} + \overset{\substack{\text{time} \\ \text{fixed effects}}}{\overbrace{\lambda_t}} + u_{it}$$

$$\Downarrow$$

$$Y_{it} = \underbrace{\beta_0}_{\text{an intercept}} + \beta_1 X_{it} + \underbrace{\gamma_2 D2_i + \gamma_3 D3_i + \ldots + \gamma_n Dn_i}_{\text{with n-1 entity binary indicators}}$$

$$+ \underbrace{\delta_2 B2_t + \delta_3 B3_t + \ldots + \delta_T BT_t}_{\text{with n-1 time binary indicators}} + u_{it}$$

Note: This form suggests estimating regression model with $n-1$ entity binary indicators and $T-1$ time binary indicators by OLS, but we can also combine with *"demeaned" OLS algorithm*.

[SN] STATA syntax for Fixed Effects Regressions

(1) Entity Fixed Effects

```
*xtreg yvar xvar, fe vce(cluster entityid)
```

(2) Time Fixed Effects

```
*regress yvar xvar i.timeid, fe vce(cluster entityid)
```

(3) Both Entity & Time Fixed Effects

```
*xtreg yvar xvar i.timeid, fe vce(cluster entityid)
```

Note: `id` means unique identifiers for entity (`entityid`) or for time period (`timeid`)

Clustered Standard Errors

$$\begin{array}{ccccc} & t=1 & t=2 & \dots & t=T \\ i=1 & u_{11} & u_{12} & \dots & u_{1T} \\ i=2 & u_{21} & u_{22} & \dots & u_{2T} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ i=n & u_{n1} & u_{n2} & \dots & u_{nT} \end{array}$$

- Sampling are *i.i.d* across entities.
- But if the omitted factors comprising the error term u_{it} are serially correlated, then u_{it} is serially correlated aka *autocorrelated* - that is, correlated over time within an entity.

Standard errors need to allow both for this autocorrelation and for potential heteroskedasticity $\rightarrow$ use **clustered standard errors**.

```
*[, vce(cluster entityid)]
//add this option at the end of STATA regressions
```

Excercise 1: based on Stock & Watson, E10.1

Excercise 1: based on Stock & Watson, E10.1

- **Objective:** Analyze effects of *concealed weapons laws* on *violent crimes*.
  - [Proponents:] More people carry concealed weapons, crime will decline because criminals will be deterred from attacking other people.
  - [Opponents:] Crime will increase because of accidental or spontaneous use of the weapons.

- **Dataset:** `Guns.dta`
  - A balanced panel of data from the 50 U.S. states plus the District of Columbia for 23 years (1977-1999).

- **Key variables:**
  - `stateid`: ID number of states (Alabama = 1, Alaska = 2, etc.)
  - `year`: year (1977-1999)
  - `shall`: = 1 if the state has a shall-carry law in effect in that year = 0 otherwise
  - `vio`: violent crime rate.

Variable Description

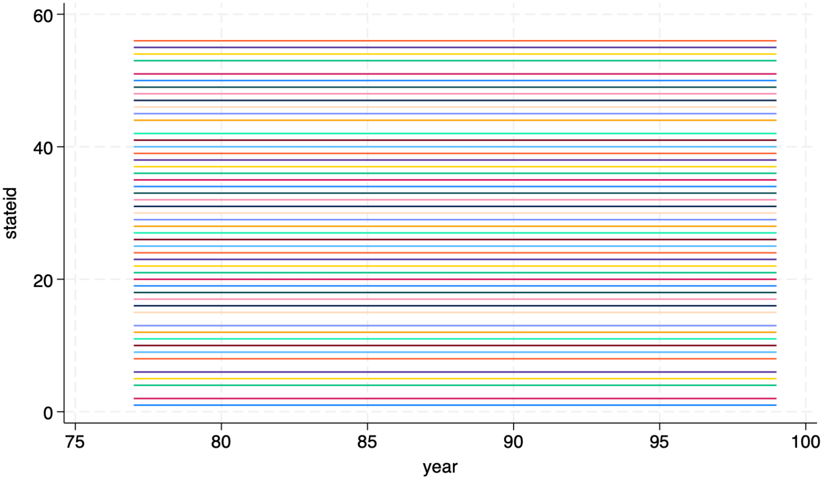
Variable	Definition
vio	violent crime rate (incidents per 100,000 members of the population)
rob	robbery rate (incidents per 100,000)
mur	murder rate (incidents per 100,000)
shall	= 1 if the state has a shall-carry law in effect in that year = 0 otherwise
incarc rate	incarceration rate in the state in the previous year (sentenced prisoners per 100,000 residents; value for the previous year)
density	population per square mile of land area, divided by 1000
avginc	real per capita personal income in the state, in thousands of dollars
pop	state population, in millions of people
pm1029	percent of state population that is male, ages 10 to 29
pw1064	percent of state population that is white, ages 10 to 64
pb1064	percent of state population that is black, ages 10 to 64
stateid	ID number of states (Alabama = 1, Alaska = 2, etc.)
year	Year (1977-1999)

Source: Ayres, I. and Donohue, J.J., 2002. Shooting down the more guns, less crime hypothesis.

```
. xtset stateid year

Panel variable: stateid (strongly balanced)
 Time variable: year, 77 to 99
         Delta: 1 unit

. xtline stateid, overlay legend(off)
```



Questions

(a) Estimate a linear regression model of:

- (M1) `ln(vio)` against `shall`.
- (M2) `ln(vio)` against `shall`, `incarc_rate`, `density`, `avginc`, `pop`, `pb1064`, `pw1064` and `pm1029`.

i. Interpret the coefficient on shall in (M2).
   Is this estimate large or small in a real-world sense?

ii. Does adding the control variables in (M2) change the estimated effect of a shall-issue law in (M1) as measured by statistical significance? As measured by the real-world significance of the estimated coefficient?

iii. Suggest a variable that varies across states but plausibly varies little or not at all over time and that could cause omitted variable bias in (M2).

Questions

(b) Do the results change when you add *state fixed effects*?
If so, which set of regression results is more credible, and why?

(c) Do the results change when you add *time fixed effects*?
If so, which set of regression results is more credible, and why?

(d) Repeat the analysis using `ln(rob)` and `ln(mur)` in place of `ln(vio)`.

(a-i,ii) Estimate linear regression models (M1) and (M2)

From OLS estimation results for (M1) (»stata) and (M2) (»stata)

- The coefficient is -0.368, which suggests that shall-issue laws reduce violent crime by 36%. This is a large effect.
- The coefficient in (1) is -0.443; in (2) it is -0.368. Both are highly statistically significant. Adding the control variables results in a small drop in the coefficient.

(a-iii) Suggest a variable that varies across states but plausibly varies little or not at all over time and that could cause omitted variable bias in (M2).

There are several examples:

- Residents' attitudes towards guns and crime, these are typically slow to change.
- Quality of police and other crime-prevention programs.
- For historical reasons, cities can have very different crime rates.
- Geographic features.
- etc.

$\Longrightarrow$ These factors constitute an *unobserved state effect/ state fixed effect/ unobserved state heterogeneity*. If they and the variable of interest (`shall`) are correlated, omitting such variables results in OVB (**heterogeneity bias**).

(b) Do the results change when you add *state fixed effects*?

Good news: These time-invariant factors can be effectively captured by the state fixed effect a_i! (»review)

(b) Do the results change when you add *state fixed effects*?

```
*xtreg yvar xvar, fe vce(cluster entityid)
```

```
. xtreg lvio shall $basevars, fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-squared:                                      Obs per group:
     Within  = 0.2178                                         min =         23
     Between = 0.0033                                         avg =       23.0
     Overall = 0.0001                                         max =         23

                                                F(8, 50)          =      34.10
corr(u_i, Xb) = -0.3687                         Prob > F          =     0.0000

                                 (Std. err. adjusted for 51 clusters in stateid)
```

lvio	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
shall	-.0461415	.0417616	-1.10	0.275	-.1300223	.0377392
incarc_rate	-.000071	.0002504	-0.28	0.778	-.0005739	.0004318
density	-.1722901	.1376129	-1.25	0.216	-.4486936	.1041135
avginc	-.0092037	.0129649	-0.71	0.481	-.0352445	.016837
pop	.0115247	.014224	0.81	0.422	-.0170452	.0400945
pb1064	.1042804	.0326849	3.19	0.002	.0386308	.1699301
pw1064	.0408611	.0134585	3.04	0.004	.0138289	.0678932
pm1029	-.0502725	.0206949	-2.43	0.019	-.0918394	-.0087057
_cons	3.866017	.7701057	5.02	0.000	2.319214	5.412819

(b) Do the results change when you add *state fixed effects*?

- In (M3) the coefficient on shall falls to -0.046, a large reduction in the coefficient from (M2). The estimate is not statistically significantly different from zero.
- Evidently there was important omitted variable bias leading to a spurious effect in (M2).
- The constant reported in `xtreg, fe` results implies the estimated average fixed effect.

Like experiments $\rightarrow$ Let's compare:

```
reg lvio shall i.stateid
```

```
xtreg lvio shall, fe
```

(c) Do the results change when you add *time fixed effects*?

```
*xtreg yvar xvar i.timeid, fe vce(cluster entityid)
```

```
. xtreg lvio shall $basevars i.year, fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-squared:                                      Obs per group:
     Within  = 0.4180                                         min =         23
     Between = 0.0419                                         avg =       23.0
     Overall = 0.0009                                         max =         23

                                                F(30, 50)         =      56.86
corr(u_i, Xb) = -0.2929                         Prob > F          =     0.0000

                                 (Std. err. adjusted for 51 clusters in stateid)
------------------------------------------------------------------------------
             |               Robust
        lvio | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
       shall |  -.0279935   .0407168    -0.69   0.495    -.1097757    .0537886
 incarc_rate |    .000076   .0002079     0.37   0.716    -.0003416    .0004935
     density |   -.091555   .1238622    -0.74   0.463    -.3403396    .1572296
      avginc |   .0009587   .0164931     0.06   0.954    -.0321688    .0340861
         pop |  -.0047544   .0152294    -0.31   0.756    -.0353436    .0258347
      pb1064 |   .0291862   .0495407     0.59   0.558    -.0703192    .1286916
      pw1064 |   .0092501   .0237564     0.39   0.699    -.0384659    .0569662
      pm1029 |   .0733254   .0524733     1.40   0.168    -.0320704    .1787211
             |
        year |
         78  |   .0585261   .0161556     3.62   0.001     .0260767    .0909755
         79  |   .1639486   .0244579     6.70   0.000     .1148233    .2130738
```

(c) Do the results change when you add *time fixed effects*?

```
. quiet xtreg lvio shall $basevars i.year, fe vce(cluster stateid)

. _ms_extract_varlist i.year

. test `r(varlist)'

 ( 1)  77b.year = 0
 ( 2)  78.year = 0
 ( 3)  79.year = 0
 ( 4)  80.year = 0
 ( 5)  81.year = 0
 ( 6)  82.year = 0
 ( 7)  83.year = 0
 ( 8)  84.year = 0
 ( 9)  85.year = 0
 (10)  86.year = 0
 (11)  87.year = 0
 (12)  88.year = 0
 (13)  89.year = 0
 (14)  90.year = 0
 (15)  91.year = 0
 (16)  92.year = 0
 (17)  93.year = 0
 (18)  94.year = 0
 (19)  95.year = 0
 (20)  96.year = 0
 (21)  97.year = 0
 (22)  98.year = 0
 (23)  99.year = 0
       Constraint 1 dropped

       F( 22,    50) =   21.62
            Prob > F =    0.0000
```

(c) Do the results change when you add *time fixed effects*?

- The coefficient falls further to -0.028. The coefficient is insignificantly different from zero.
- The time effects are jointly statistically significant, so this regression seems better specified than (M3).

Table of Results - Exercise 1

Regressor	Models			
	(M1)	**(M2)**	**(M3)**	**(M4)**
shall	-0.443*** (0.157) [-0.76, -0.13]	-0.368*** (0.114) [-0.60, -0.14]	-0.0461 (0.042) [-0.13, 0.04]	-0.0280 (0.041) [-0.11, 0.05]
Controls	No	Yes	Yes	Yes
State effects	No	No	Yes	Yes
Time effects	No	No	No	Yes

Notes: Clustered standard errors shown in parentheses and 95% confidence intervals are shown in brackets; ***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.

Excercise 2: based on Stock & Watson, E10.2

Picture the Scenario

- **Objective:** Do citizens demand more democracy and political freedom as their incomes grow? That is, is democracy a normal good?

- **Dataset:** `Income-Democracy.dta`
  - a panel data set from 195 countries for the years $1960, 1965, \ldots, 2000$.

- **Key variables:** For each country in each year
  - `Dem_ind`: an index of political freedom/democracy.
  - `Log_GDPPC`: per capita income.
  - various demographic controls.

Variable Description

Variable Name	Description
country	country name
year	year
dem_ind	index of democracy
log_gdppc	logarithm of real GDP per capita
log_pop	logarithm of population
age_1	fraction of the population age 0-14
age_2	fraction of the population age 15-29
age_3	fraction of the population age 30-44
age_4	fraction of the population age 45-59
age_5	fraction of the population age 60 and older
educ	average years of education for adults (25 years and older)
age_median	median age
code	country code

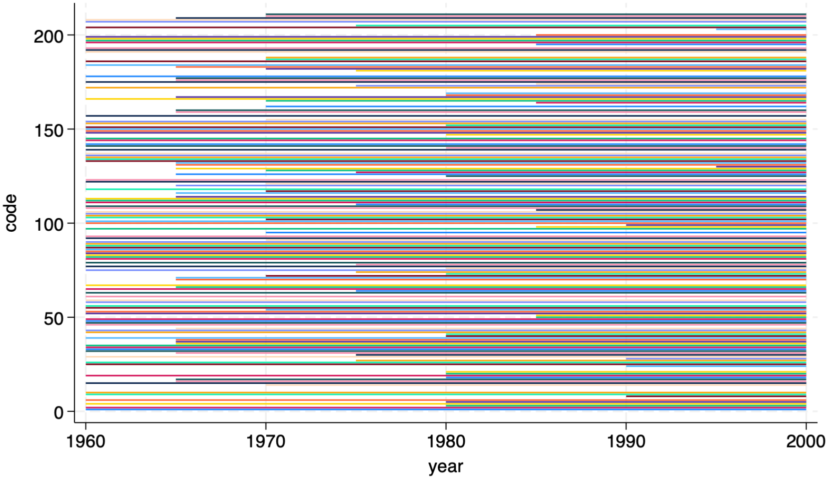
Source: Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared. 2008. "Income and Democracy." American Economic Review, 98 (3): 808-42.

Notes: The income and demographic variable are lagged five years.

```
. xtset code year

Panel variable: code (unbalanced)
 Time variable: year, 1960 to 2000, but with gaps
         Delta: 1 unit

. xtline code, overlay legend(off)
```



Questions

(a) Is the data set a balanced panel? Explain.

(b) The logarithm of per capita income is labeled `Log_GDPPC`. Regress `Dem_ind` on `Log_GDPPC`. Use standard errors that are clustered by country.

i. How large is the estimated coefficient on `Log_GDPPC`?
Is the coefficient statistically significant?

ii. If per capita income in a country increases by 20%, by how much is `Dem_ind` predicted to increase?
What is a 95% confidence interval for the prediction?
Is the predicted increase in `Dem_ind` large or small? (Explain what you mean by large or small.)

iii. Why is it important to use clustered standard errors for the regression?
Do the results change if you do not use clustered standard errors?

Questions

(c) Estimate the regression in (b), allowing for country fixed effects. How do your answers to (b-i) change?

(d) Estimate the regression in (b), allowing for time and country fixed effects. How do your answers to (b-i) change?

(e) There are additional demographic controls in the data set. Should these variables be included in the regression? If so, how do the results change when they are included?

(a) Is the data set a balanced panel? Explain.

- The dataset is unbalanced because data are available over different years for different countries. For example, data on `Dem_ind` are available
  - for Andorra during 1970, 1995, and 2000;
  - for Afghanistan during $1960, 1965, \ldots, 2000$.

(b) Regress `Dem_ind` on `Log_GDPPC` using standard errors clustered by country.

```
. reg dem_ind log_gdppc, vce(cluster code) //using clustered standard errors

Linear regression                               Number of obs     =        958
                                                F(1, 149)         =     396.40
                                                Prob > F          =     0.0000
                                                R-squared         =     0.4385
                                                Root MSE          =      .2719

                                 (Std. err. adjusted for 150 clusters in code)
```

dem_ind	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
log_gdppc	.2356731	.011837	19.91	0.000	.212283	.2590632
_cons	-1.354828	.1004215	-13.49	0.000	-1.553262	-1.156394

(b-i) How large is the estimated coefficient on `Log_GDPPC`?

The coefficient is 0.236 with a standard error of 0.012. The 95% confidence interval is 0.212 to 0.259. The coefficient is statistically significant.

(b-ii) If per capita income in a country increases by 20%, by how much is `Dem_ind` predicted to increase?

Hint: Review Regression models with functional forms involving logarithms.

- A 20% increase in GDP per capita implies that log_gdp increases by approximately 0.20, so that `Dem_ind` is predicted to increase by approximately $0.20 \times 0.236 = 0.0472$, or about $1/8$ of the standard deviation in the dataset.
- The 95% confidence for the effect is (approximately) 0.20×0.212 to $0.20 \times 0.259 = 0.0472$ or 0.0424 to 0.0518.

(b-iii) Why is it important to use clustered standard errors for the regression?

```
. reg dem_ind log_gdppc, vce(cluster code) //using clustered standard errors

Linear regression                               Number of obs     =        958
                                                F(1, 149)         =     396.40
                                                Prob > F          =     0.0000
                                                R-squared         =     0.4385
                                                Root MSE          =      .2719

                                  (Std. err. adjusted for 150 clusters in code)
------------------------------------------------------------------------------
             |               Robust
     dem_ind | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
   log_gdppc |   .2356731    .011837    19.91   0.000      .212283    .2590632
       _cons |  -1.354828   .1004215   -13.49   0.000    -1.553262   -1.156394
------------------------------------------------------------------------------
```

```
. reg dem_ind log_gdppc, robust //without clustering

Linear regression                               Number of obs     =        958
                                                F(1, 956)         =    1100.57
                                                Prob > F          =     0.0000
                                                R-squared         =     0.4385
                                                Root MSE          =      .2719

------------------------------------------------------------------------------
             |               Robust
     dem_ind | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
   log_gdppc |   .2356731    .007104    33.17   0.000      .221732    .2496143
       _cons |  -1.354828   .0607549   -22.30   0.000    -1.474056     -1.2356
------------------------------------------------------------------------------
```

(b-iii) Why is it important to use clustered standard errors for the regression?

- Clustered standard errors are needed because of country-specific omitted factors in the regressions.
- The clustered standard error for `Dem_ind` is 0.012; the unclustered standard error is smaller (0.007) because it ignores the positive within-country autocorrelation of the errors.

(c) Estimate the regression in (b), allowing for country fixed effects.

Estimation result (»stata)

The estimated coefficient falls by a factor of 3, to 0.083 with a standard error of 0.032. The estimated effect, while significantly smaller is still statistically significant at the 1% significance level.

(d) Estimate the regression in (b), allowing for time and country fixed effects.

Estimation and test results (»stata)

The estimated coefficient falls further to 0.054, approximately 1/5 of the value that omits time and country fixed effects. The estimate is not statistically significant at the 10% level.

(e) Include additional demographic controls in the regression

Estimation and test results (»stata)

When age, population, and education are included, the estimated coefficient on `log_gdppc` falls further to 0.025 with a standard error of 0.054. Jointly, these variables are not statistically significant in the regression, although the age variables are significant at the 10% level.

Table of Results - Exercise 2

Regressor	**Models**				
	(M1)	**(M2)**	**(M3)**	**(M4)**	**(M5)**
log_GDPPC	0.236*** (0.012) [0.212, 0.259]	0.235*** (0.012) [0.211, 0.259]	0.083*** (0.031) [0.021, 0.146]	0.054 (0.042) [-0.030, 0.137]	0.025 (0.054) [-0.057, 0.120]
Controls	No	No	No	No	Yes
Country effects	No	No	Yes	Yes	Yes
Time effects	No	Yes	No	Yes	Yes
F-statistics and p-values testing exclusion of groups of variables					
Time effects		9.31 (0.000)		5.73 (0.00)	4.61 (0.000)
Age variables					2.12 (0.08)
Age, educ, pop variables					1.44 (0.21)

Notes: Clustered standard errors shown in parentheses and 95% confidence intervals are shown in brackets; ***$p < 0.01$,**$p < 0.05$,*$p < 0.1$.

STATA CODES & RESULTS

Exercise 1(a-i) (»back(1a))

```
. regress lvio shall, vce(cluster stateid)

Linear regression                               Number of obs     =      1,173
                                                F(1, 50)          =       7.96
                                                Prob > F          =     0.0068
                                                R-squared         =     0.0866
                                                Root MSE          =     .61735

                              (Std. err. adjusted for 51 clusters in stateid)
```

lvio	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
shall	-.4429646	.1570184	-2.82	0.007	-.7583452	-.1275839
_cons	6.134919	.0790269	77.63	0.000	5.976189	6.293649

Exercise 1(a-ii) (»back(1a))

```
. global basevars "incarc_rate density avginc pop pb1064 pw1064 pm1029"

. reg lvio shall $basevars, vce(cluster stateid)

Linear regression                               Number of obs     =      1,173
                                                F(8, 50)          =      62.13
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5643
                                                Root MSE          =     .42769

                               (Std. err. adjusted for 51 clusters in stateid)
```

lvio	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
shall	-.3683869	.113937	-3.23	0.002	-.5972361	-.1395378
incarc_rate	.0016126	.0005999	2.69	0.010	.0004076	.0028177
density	.0266885	.0414909	0.64	0.523	-.0566485	.1100255
avginc	.0012051	.0240808	0.05	0.960	-.0471626	.0495728
pop	.0427098	.011729	3.64	0.001	.0191515	.0662681
pb1064	.0808526	.0713875	1.13	0.263	-.0625334	.2242386
pw1064	.0312005	.03409	0.92	0.364	-.0372713	.0996723
pm1029	.0088709	.0340964	0.26	0.796	-.0596137	.0773554
_cons	2.981738	2.166513	1.38	0.175	-1.369831	7.333307

Exercise 1(b-i)

```
*xtreg yvar xvar, fe vce(cluster entityid)
```

```
. xtreg lvio shall $basevars, fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-squared:                                      Obs per group:
     Within  = 0.2178                                         min =         23
     Between = 0.0033                                         avg =       23.0
     Overall = 0.0001                                         max =         23

                                                F(8, 50)          =      34.10
corr(u_i, Xb) = -0.3687                         Prob > F          =     0.0000

                                 (Std. err. adjusted for 51 clusters in stateid)
```

lvio	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
shall	-.0461415	.0417616	-1.10	0.275	-.1300223	.0377392
incarc_rate	-.000071	.0002504	-0.28	0.778	-.0005739	.0004318
density	-.1722901	.1376129	-1.25	0.216	-.4486936	.1041135
avginc	-.0092037	.0129649	-0.71	0.481	-.0352445	.016837
pop	.0115247	.014224	0.81	0.422	-.0170452	.0400945
pb1064	.1042804	.0326849	3.19	0.002	.0386308	.1699301
pw1064	.0408611	.0134585	3.04	0.004	.0138289	.0678932
pm1029	-.0502725	.0206949	-2.43	0.019	-.0918394	-.0087057
_cons	3.866017	.7701057	5.02	0.000	2.319214	5.412819

Exercise 1(b-ii-I)

```
*xtreg yvar xvar i.timeid, fe vce(cluster entityid)
```

```
. xtreg lvio shall $basevars i.year, fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51
```

Alternative,

```
. quiet tabulate year, generate(yr)

. xtreg lvio shall $basevars yr*, fe vce(cluster stateid)
note: yr23 omitted because of collinearity.

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51
```

Alternative,

```
. quiet tabulate year, generate(yr)

. global yr_vars "yr2  yr3  yr4  yr5  yr6  yr7  yr8  yr9 yr10 yr11 yr12 yr13 yr14 yr15 yr16 yr17 yr18 yr19 yr20 yr21 yr22 yr23"

. xtreg lvio shall $basevars $yr_vars, fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51
```

Exercise 1(b-ii-II)

```
. quiet tabulate year, generate(yr)

. global yr_vars "yr2  yr3  yr4  yr5  yr6  yr7  yr8  yr9 yr10 yr11 yr12 yr13 yr14 yr15 yr16 yr17 yr18 yr19 yr20 yr21 yr22 yr23"

. xtreg lvio shall $basevars $yr_vars, fe vce(cluster stateid)

Fixed-effects (within) regression               Number of obs     =      1,173
Group variable: stateid                         Number of groups  =         51

R-squared:                                      Obs per group:
     Within  = 0.4180                                         min =         23
     Between = 0.0419                                         avg =       23.0
     Overall = 0.0009                                         max =         23

                                                F(30, 50)         =      56.86
corr(u_i, Xb) = -0.2929                         Prob > F          =     0.0000

                                 (Std. err. adjusted for 51 clusters in stateid)
--------------------------------------------------------------------------------
               |               Robust
          lvio | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
---------------+----------------------------------------------------------------
         shall |  -.0279935   .0407168    -0.69   0.495    -.1097757    .0537886
   incarc_rate |    .000076   .0002079     0.37   0.716    -.0003416    .0004935
       density |   -.091555   .1238622    -0.74   0.463    -.3403396    .1572296
        avginc |   .0009587   .0164931     0.06   0.954    -.0321688    .0340861
           pop |  -.0047544   .0152294    -0.31   0.756    -.0353436    .0258347
        pb1064 |   .0291862   .0495407     0.59   0.558    -.0703192    .1286916
        pw1064 |   .0092501   .0237564     0.39   0.699    -.0384659    .0569662
        pm1029 |   .0733254   .0524733     1.40   0.168    -.0320704    .1787211
```

Exercise 1(b-iii)

```
. quiet xtreg lvio shall $basevars $yr_vars, fe vce(cluster stateid)

. test $yr_vars

 ( 1)  yr2 = 0
 ( 2)  yr3 = 0
 ( 3)  yr4 = 0
 ( 4)  yr5 = 0
 ( 5)  yr6 = 0
 ( 6)  yr7 = 0
 ( 7)  yr8 = 0
 ( 8)  yr9 = 0
 ( 9)  yr10 = 0
 (10)  yr11 = 0
 (11)  yr12 = 0
 (12)  yr13 = 0
 (13)  yr14 = 0
 (14)  yr15 = 0
 (15)  yr16 = 0
 (16)  yr17 = 0
 (17)  yr18 = 0
 (18)  yr19 = 0
 (19)  yr20 = 0
 (20)  yr21 = 0
 (21)  yr22 = 0
 (22)  yr23 = 0

       F( 22,    50) =   21.62
            Prob > F =    0.0000
```

Exercise 2(b-I)

```
. reg dem_ind log_gdppc, vce(cluster code) //using clustered standard errors

Linear regression                               Number of obs     =        958
                                                F(1, 149)         =     396.40
                                                Prob > F          =     0.0000
                                                R-squared         =     0.4385
                                                Root MSE          =      .2719

                                   (Std. err. adjusted for 150 clusters in code)
```

dem_ind	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
log_gdppc	.2356731	.011837	19.91	0.000	.212283	.2590632
_cons	-1.354828	.1004215	-13.49	0.000	-1.553262	-1.156394

Exercise 2(b-II)

```
. reg dem_ind log_gdppc, robust //without clustering

Linear regression                               Number of obs     =        958
                                                F(1, 956)         =    1100.57
                                                Prob > F          =     0.0000
                                                R-squared         =     0.4385
                                                Root MSE          =      .2719

------------------------------------------------------------------------------
             |               Robust
     dem_ind | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
   log_gdppc |   .2356731    .007104    33.17   0.000      .221732    .2496143
       _cons |  -1.354828   .0607549   -22.30   0.000    -1.474056     -1.2356
------------------------------------------------------------------------------
```

```
. //time fixed effects only
. reg dem_ind log_gdppc y1965 y1970 y1975 y1980 y1985 y1990 y1995 y2000, vce(cluster code)

Linear regression                               Number of obs     =        958
                                                F(9, 149)         =      56.80
                                                Prob > F          =     0.0000
                                                R-squared         =     0.4767
                                                Root MSE          =     .26357

                                    (Std. err. adjusted for 150 clusters in code)
```

dem_ind	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
log_gdppc	.2351066	.0122358	19.21	0.000	.2109285	.2592847
y1965	-.0756674	.0205124	-3.69	0.000	-.1162003	-.0351346
y1970	-.2387064	.0343905	-6.94	0.000	-.3066625	-.1707504
y1975	-.2801793	.0341178	-8.21	0.000	-.3475965	-.2127621
y1980	-.2445214	.0331917	-7.37	0.000	-.3101086	-.1789341
y1985	-.2415764	.035402	-6.82	0.000	-.3115311	-.1716216
y1990	-.2064564	.0328683	-6.28	0.000	-.2714045	-.1415083
y1995	-.1720611	.0351667	-4.89	0.000	-.2415511	-.1025712
y2000	-.1687362	.0344812	-4.89	0.000	-.2368716	-.1006009
_cons	-1.156693	.1062762	-10.88	0.000	-1.366696	-.9466898

```
. quiet reg dem_ind log_gdppc y1965 y1970 y1975 y1980 y1985 y1990 y1995 y2000, vce(cluster code)

. test y1965 y1970 y1975 y1980 y1985 y1990 y1995 y2000

 ( 1)  y1965 = 0
 ( 2)  y1970 = 0
 ( 3)  y1975 = 0
 ( 4)  y1980 = 0
 ( 5)  y1985 = 0
 ( 6)  y1990 = 0
 ( 7)  y1995 = 0
 ( 8)  y2000 = 0

       F(  8,   149) =    9.31
            Prob > F =    0.0000
```

Exercise 2(c) (»back(2c))

```
. //country fixed effects only
. xtreg dem_ind log_gdppc, fe vce(cluster code)

Fixed-effects (within) regression               Number of obs     =        958
Group variable: code                            Number of groups  =        150

R-squared:                                      Obs per group:
     Within  = 0.0197                                         min =          1
     Between = 0.5365                                         avg =        6.4
     Overall = 0.4385                                         max =          9

                                                F(1, 149)         =       7.06
corr(u_i, Xb) = 0.6173                          Prob > F          =     0.0088

                                 (Std. err. adjusted for 150 clusters in code)
------------------------------------------------------------------------------
             |               Robust
     dem_ind | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
   log_gdppc |    .083741   .0315258     2.66   0.009     .0214456    .1460364
       _cons |   -.115316    .257198    -0.45   0.655    -.6235425    .3929106
-------------+----------------------------------------------------------------
     sigma_u |  .26651952
     sigma_e |  .20351058
         rho |  .63168655   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Exercise 2(d-I) (»back(2d))

```
. // both country and time fixed effects
. xtreg dem_ind log_gdppc y1965 y1970 y1975 y1980 y1985 y1990 y1995 y2000, fe vce(cluster code)

Fixed-effects (within) regression               Number of obs     =        958
Group variable: code                            Number of groups  =        150

R-squared:                                      Obs per group:
     Within  = 0.1182                                         min =          1
     Between = 0.3832                                         avg =        6.4
     Overall = 0.3190                                         max =          9

                                                F(9, 149)         =       5.65
corr(u_i, Xb) = 0.4393                          Prob > F          =     0.0000

                                 (Std. err. adjusted for 150 clusters in code)
------------------------------------------------------------------------------
             |               Robust
     dem_ind | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
   log_gdppc |   .0535878    .042432     1.26   0.209    -.0302585     .137434
       y1965 |   .0002347   .0209199     0.01   0.991    -.0411033    .0415727
       y1970 |  -.1268076   .0340453    -3.72   0.000    -.1940816   -.0595337
       y1975 |  -.1477264   .0370153    -3.99   0.000    -.2208692   -.0745836
       y1980 |   -.097822   .0355399    -2.75   0.007    -.1680494   -.0275947
       y1985 |  -.0871025   .0391062    -2.23   0.027    -.1643769    -.009828
       y1990 |  -.0421216   .0353035    -1.19   0.235    -.1118818    .0276385
       y1995 |   .0095646   .0426094     0.22   0.823    -.0746322    .0937613
       y2000 |   .0323636   .0432037     0.75   0.455    -.0530075    .1177348
       _cons |   .1802954    .327202     0.55   0.582    -.4662601    .8268508
-------------+----------------------------------------------------------------
     sigma_u |  .28355993
     sigma_e |  .19397224
         rho |  .68122712   (fraction of variance due to u_i)
```

Exercise 2(d-II) (»back(2d))

```
. quiet xtreg dem_ind log_gdppc y1965 y1970 y1975 y1980 y1985 y1990 y1995 y2000, fe vce(cluster code)

. test y1965 y1970 y1975 y1980 y1985 y1990 y1995 y2000

 ( 1)  y1965 = 0
 ( 2)  y1970 = 0
 ( 3)  y1975 = 0
 ( 4)  y1980 = 0
 ( 5)  y1985 = 0
 ( 6)  y1990 = 0
 ( 7)  y1995 = 0
 ( 8)  y2000 = 0

       F(  8,   149) =    5.73
            Prob > F =    0.0000
```

Exercise 2(e-I) (»back(2e))

```
. xtreg dem_ind log_gdppc log_pop educ age_2 age_3 age_4 age_5 y1965 y1970 y1975 y1980 y1985 y1990 y1995 y2000, fe vce(cluster code)
note: y2000 omitted because of collinearity.

Fixed-effects (within) regression               Number of obs     =        680
Group variable: code                            Number of groups  =         96

R-squared:                                      Obs per group:
     Within  = 0.1327                                         min =          1
     Between = 0.0165                                         avg =        7.1
     Overall = 0.0359                                         max =          8

                                                F(14, 95)         =       3.22
corr(u_i, Xb) = -0.2385                         Prob > F          =     0.0004

                                  (Std. err. adjusted for 96 clusters in code)
------------------------------------------------------------------------------
             |               Robust
     dem_ind | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
   log_gdppc |   .0252013   .0539626     0.47   0.642    -.0819281    .1323306
     log_pop |  -.0692295   .1245471    -0.56   0.580    -.3164868    .1780278
        educ |  -.0004013   .0232475    -0.02   0.986    -.0465534    .0457509
       age_2 |  -.5255157    .609811    -0.86   0.391    -1.736144    .6851122
       age_3 |  -2.481235   .8941413    -2.77   0.007     -4.25633   -.7061401
       age_4 |   .2978116   1.297809     0.23   0.819    -2.278666    2.874289
       age_5 |   .6054059   1.296464     0.47   0.642      -1.9684    3.179212
       y1965 |  -.1582882    .119523    -1.32   0.189    -.3955713    .0789949
       y1970 |  -.2599235   .1058436    -2.46   0.016    -.4700497   -.0497974
       y1975 |  -.2815462   .0923211    -3.05   0.003    -.4648267   -.0982656
       y1980 |  -.2052376   .0784034    -2.62   0.010     -.360888   -.0495871
       y1985 |  -.1537799   .0596167    -2.58   0.011     -.272134   -.0354257
       y1990 |  -.1008704   .0443634    -2.27   0.025    -.1889428   -.0127979
       y1995 |  -.0397771   .0220018    -1.81   0.074    -.0834561    .0039019
       y2000 |          0  (omitted)
       _cons |   1.649546   1.406281     1.17   0.244    -1.142275    4.441368
-------------+----------------------------------------------------------------
     sigma_u |  .30898253
     sigma_e |  .19831249
         rho |  .70824619   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Exercise 2(e-II) (»back(2e))

```
. quiet xtreg dem_ind log_gdppc log_pop educ age_2 age_3 age_4 age_5 y1965 y1970 y1975 y1980 y1985 y1990 y1995 y2000, fe vce(cluster code)

. test y1965 y1970 y1975 y1980 y1985 y1990 y1995 y2000

 ( 1)  y1965 = 0
 ( 2)  y1970 = 0
 ( 3)  y1975 = 0
 ( 4)  y1980 = 0
 ( 5)  y1985 = 0
 ( 6)  y1990 = 0
 ( 7)  y1995 = 0
 ( 8)  o.y2000 = 0
       Constraint 8 dropped

       F(  7,    95) =    4.61
            Prob > F =    0.0002


. test age_2 age_3 age_4 age_5

 ( 1)  age_2 = 0
 ( 2)  age_3 = 0
 ( 3)  age_4 = 0
 ( 4)  age_5 = 0

       F(  4,    95) =    2.12
            Prob > F =    0.0837


. test age_2 age_3 age_4 age_5 educ log_pop

 ( 1)  age_2 = 0
 ( 2)  age_3 = 0
 ( 3)  age_4 = 0
 ( 4)  age_5 = 0
 ( 5)  educ = 0
 ( 6)  log_pop = 0

       F(  6,    95) =    1.44
            Prob > F =    0.2070
```