PLSC 504 – Fall 2024 Panel/TSCS Data: Unit Effects (+ Dynamics)

October 9, 2024

Starting Points

- "Longitudinal" $\neq$ "Time Series"
- Terminology:
  - "Unit" / "Units" / "Units of observation" / "Panels" $=$ Things we observe repeatedly
  - "Observations" $=$ Each (one) measurement of a unit
  - "Time points" $=$ When each observation on a unit is made
  - $i \in \{1....N\}$ indexes units
  - $t \in \{1...T\}$ or $\{1...T_i\}$ indexes observations / time points
  - If $T_i = T \; \forall i$ then we have "balanced" panels / units
  - NT $=$ Total number of observations (if balanced)
- Averages:
  - Y_{it} indicates a variable that varies over both units and time,
  - $\bar{Y}_i = \frac{1}{T}\sum_{t=1}^{T} Y_{it}$ $=$ the over-time mean of Y for i,
  - $\bar{Y}_t = \frac{1}{N}\sum_{i=1}^{N} Y_{it}$ $=$ the across-unit mean of Y at t, and
  - $\bar{Y} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} Y_{it}$ $=$ the grand mean of Y.

More Terminology

- $N >> T \rightarrow$ "panel" data
  - NES panel studies ($N = 2000$, $T = 3$)
  - Panel Study of Income Dynamics ($N =$ large, $T \approx 12$)
- $T >> N$ or $T \approx N \rightarrow$ "time-series cross-sectional" ("TSCS") data
- $N = 1 \rightarrow$ "time series" data

Within- and Between-Unit Variation

Define:

$$\bar{Y}_i = \frac{1}{T_i} \sum_{t=1}^{T_i} Y_{it}$$

Then:

$$Y_{it} = \bar{Y}_i + (Y_{it} - \bar{Y}_i).$$

- The *total* variation in Y_{it} can be decomposed into
- The *between-unit* variation in the $\bar{Y}_i$s, and
- The *within-unit* variation around $\bar{Y}_i$ (that is, $Y_{it} - \bar{Y}_i$).

Regression!

Model

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

assumes:

- All the usual OLS assumptions, plus
- $\beta_{0i} = \beta_0 \; \forall \; i$s
- $\beta_{1i} = \beta_1 \; \forall \; i$s

For:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + u_{it}$$

... it's the same.

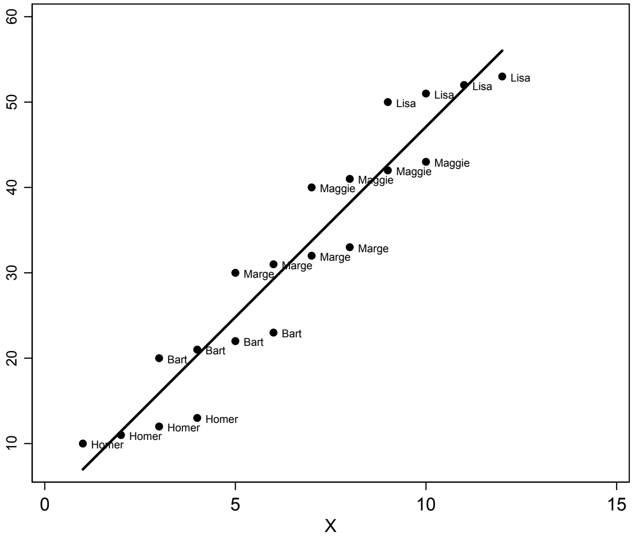
Relax that assumption:

$$Y_{it} = \beta_{0i} + \beta_1 X_{it} + u_{it} \qquad \text{(unit-level)}$$

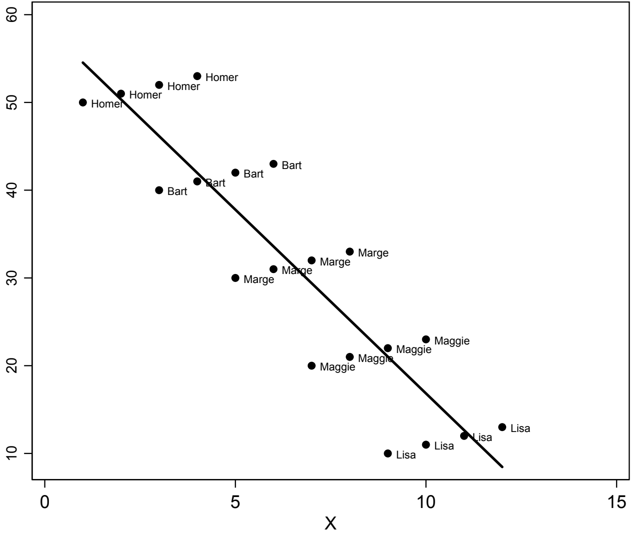
$$Y_{it} = \beta_{0t} + \beta_1 X_{it} + u_{it} \qquad \text{(time-level)}$$

$$Y_{it} = \beta_{0it} + \beta_1 X_{it} + u_{it} \qquad \text{(unit- and time-level)}$$

Varying Intercepts



Varying Intercepts



Varying Slopes (+ Intercepts)

Further relax:

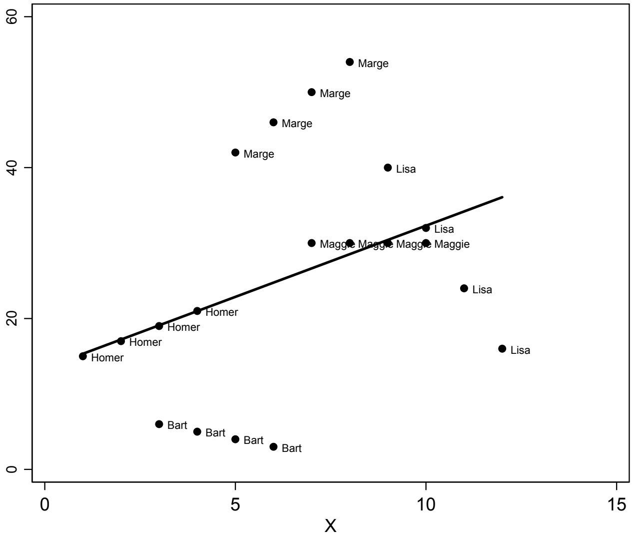
$$Y_{it} = \beta_0 + \beta_{1i} X_{it} + u_{it} \qquad \text{(unit-level slopes)}$$

$$Y_{it} = \beta_{0i} + \beta_{1i} X_{it} + u_{it} \qquad \text{(unit-level slopes and intercepts)}$$

$$Y_{it} = \beta_{0t} + \beta_{1t} X_{it} + u_{it} \qquad \text{(time-level slopes and intercepts)}$$

$$Y_{it} = \beta_{0it} + \beta_{1it} X_{it} + u_{it} \qquad \text{(unit- and time-level slopes and intercepts)}$$

Varying Slopes + Intercepts



The Error

"The usual" assumptions require:

$$u_{it} \sim \text{i.i.d.} N(0, \sigma^2) \; \forall \; i, t$$

This means that:

$$\begin{aligned}
\text{Var}(u_{it}) &= \text{Var}(u_{jt}) \; \forall \; i \neq j \text{ (i.e., no cross-unit heteroscedasticity)} \\
\text{Var}(u_{it}) &= \text{Var}(u_{is}) \; \forall \; t \neq s \text{ (i.e., no temporal heteroscedasticity)} \\
\text{Cov}(u_{it}, u_{js}) &= 0 \; \forall \; i \neq j, \; \forall \; t \neq s \text{ (i.e., no auto- or spatial correlation)}
\end{aligned}$$

Two-Way Variation

Two-way variation:

$$Y_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \gamma V_i + \delta W_t + u_{it}$$

where V_i don't vary over time (within a unit), and W_i don't vary across units (for a given time point).

Note that we can write:

$$\alpha_i = \sum(\gamma V_i)$$

and

$$\eta_t = \sum(\delta W_t).$$

So:

$$\begin{aligned} Y_{it} &= \mathbf{X}_{it}\boldsymbol{\beta} + \gamma V_i + \delta W_t + u_{it} \\ &= \mathbf{X}_{it}\boldsymbol{\beta} + \alpha_i + \eta_t + u_{it} \end{aligned}$$

One- and Two-Way "Unit Effects"

"Two-way" unit effects:

$$Y_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \alpha_i + \eta_t + u_{it}$$

"One-way" effects:

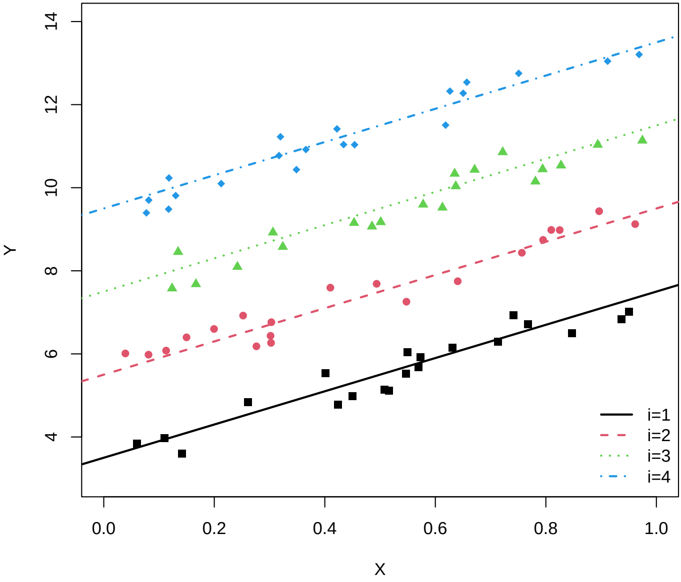
- Assuming $\alpha_i = 0$ (w.l.o.g):

$$Y_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \eta_t + u_{it} \quad \text{(time)}$$

- Assuming $\eta_t = 0$ (w.l.o.g):

$$Y_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \alpha_i + u_{it} \quad \text{(units)}$$

Intuition: One-Way Unit Effects



(One-Way) "Fixed" Effects

"Brute force" model fits:

$$\begin{aligned} Y_{it} &= \mathbf{X}_{it}\boldsymbol{\beta}_{FE} + \alpha_i + u_{it} \\ &= \mathbf{X}_{it}\boldsymbol{\beta}_{FE} + \alpha_1 I(i=1)_i + \alpha_2 I(i=2)_i + ... + u_{it} \end{aligned}$$

Alternatively, decompose:

$$\bar{\mathbf{X}}_i = \frac{\sum_{N_i} \mathbf{X}_{it}}{N_i}$$

and

$$\tilde{\mathbf{X}}_{it} = \mathbf{X}_{it} - \bar{\mathbf{X}}_i.$$

Yields:

$$Y_{it} = \bar{\mathbf{X}}_i\boldsymbol{\beta}_B + \tilde{\mathbf{X}}_{it}\boldsymbol{\beta}_W + \alpha_i + u_{it}$$

<u>But!</u>

$$\text{corr}(\bar{\mathbf{X}}_i\boldsymbol{\beta}_B, \alpha_i) = 1.0$$

Means that:

$$\begin{aligned} Y_{it}^* &= Y_{it} - \bar{Y}_i \\ \mathbf{X}_{it}^* &= \mathbf{X}_{it} - \bar{\mathbf{X}}_i \end{aligned}$$

gives:

$$Y_{it}^* = \mathbf{X}_{it}^* \beta_{FE} + u_{it}.$$

$\rightarrow$ **A “Fixed Effects” Model is actually a “Within-Effects” Model.**

Standard F-test for

$$H_0 : \alpha_i = \alpha_j \forall i \neq j$$

versus

$$H_A : \alpha_i \neq \alpha_j \text{ for some } i \neq j$$

is $\sim F_{N-1, NT-(N-1)}$.

Running Example Data: WDI, 1960-2023

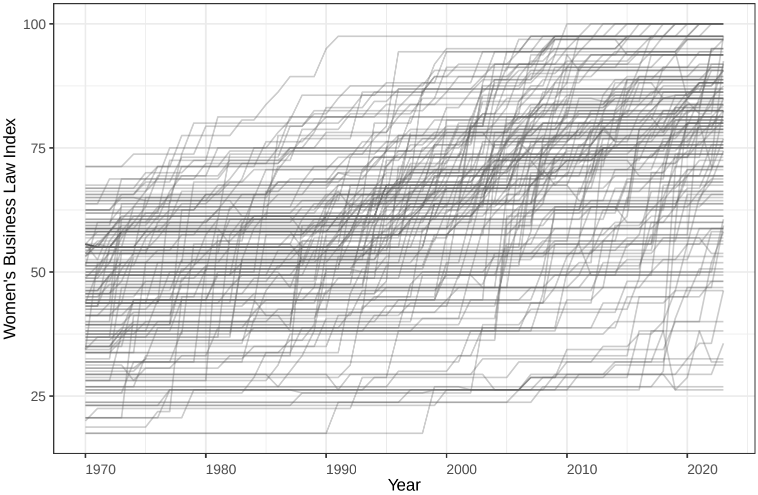
The World Development Indicators

- Cross-national country-level time series data
- $N = 215$ countries, $T = 74$ years (1960-2023) + missingness
- Variables:
  - Geography: land area, arable land
  - Population indicators
  - Demographics: Birth rates, life expectancy, etc.
  - Economics: GDP, inflation, trade, FDI, etc.
  - Governments: expenditures, policies, etc.
- Full descriptions are listed in the Github repo here

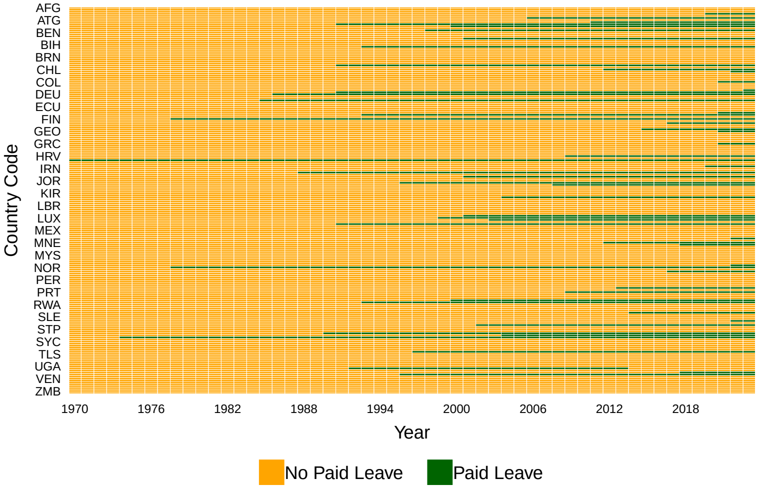
Data Summary

```
> psych::describe(wdi,fast=TRUE,ranges=FALSE,check=TRUE)
                        vars     n            mean          sd  skew kurtosis             se
ISO3                       1 13760             NaN          NA    NA       NA             NA
Year                       2 13760         1991.50 1.847e+01  0.00    -1.20           0.16
Region                     3 13760             NaN          NA    NA       NA             NA
country                    4 13760             NaN          NA    NA       NA             NA
iso3c                      5 13760             NaN          NA    NA       NA             NA
LandArea                   6 11941       605302.93 1.640e+06  5.41    34.35       15006.31
ArablePercent              7 11542           13.35 1.361e+01  1.49     2.05           0.13
Population                 8 13730     25299145.44 1.054e+08  9.72   105.45      899894.67
PopGrowth                  9 13513            1.74 1.780e+00  0.83    22.71           0.02
RuralPopulation           10 13696           48.11 2.574e+01 -0.11    -1.00           0.22
UrbanPopulation           11 13696           51.89 2.574e+01  0.11    -1.00           0.22
BirthRatePer1K            12 13150           27.86 1.310e+01  0.23    -1.24           0.11
FertilityRate             13 12987            3.88 2.000e+00  0.40    -1.21           0.02
PrimarySchoolAge          14 11119            6.13 6.100e-01 -0.04     0.11           0.01
LifeExpectancy            15 12974           64.75 1.128e+01 -0.73    -0.01           0.10
AgeDepRatioOld            16 13730           10.78 7.150e+00  1.76     4.66           0.06
CO2Emissions              17  5920            4.24 5.450e+00  2.75    11.36           0.07
GDP                       18 11016 240364127688.72 1.130e+12 11.46   159.31 10763521700.11
GDPPerCapita              19 11021        12146.44 1.897e+04  3.13    14.36         180.71
GDPPerCapGrowth           20 10942            1.91 6.680e+00  2.58    61.16           0.06
Inflation                 21  8708           23.24 3.241e+02 53.56  3555.07           3.47
TotalTrade                22  8843           78.67 5.383e+01  2.99    17.70           0.57
Exports                   23  8843           36.58 2.884e+01  2.96    16.14           0.31
Imports                   24  8852           42.10 2.756e+01  2.53    13.49           0.29
FDIIn                     25  8861            5.33 4.404e+01 16.01   599.63           0.47
AgriEmployment            26  5951           28.79 2.411e+01  0.64    -0.74           0.31
NetAidReceived            27  9043    506951242.00 9.971e+08  8.32   157.34    10484966.48
MobileCellSubscriptions   28 10212           36.32 5.176e+01  1.29     1.14           0.51
NaturalResourceRents      29  9211            6.85 1.106e+01  2.60     8.04           0.12
MilitaryExpenditures      30  7555            2.72 3.190e+00  9.45   240.84           0.04
GovtExpenditures          31  8475           16.27 7.980e+00  3.71    36.33           0.09
PublicEdExpend            32  4927            4.34 1.950e+00  2.89    39.91           0.03
PublicHealthExpend        33  4098            3.31 2.380e+00  1.34     3.10           0.04
HIVDeaths                 34  4656         6473.06 1.892e+04  5.78    45.97         277.31
WomenBusLawIndex          35 10152           59.85 1.874e+01  0.02    -0.58           0.19
PaidParentalLeave         36 10152            0.11 3.100e-01  2.50     4.27           0.00
ColdWar                   37 13760            0.47 5.000e-01  0.13    -1.98           0.00
```

Visualization (using `panelView`)



Categorical Variable Visualization



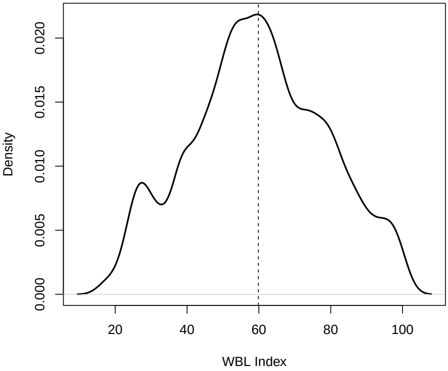
WDI's Women, Business and the Law Index (WBLI)

The basis for a 2021 World Bank report...

- Examines "the laws and regulations that affect women's economic opportunity in 190 economies."
- An index comprising eight indicators "structured around women's interactions with the law as they move through their careers: *Mobility*, *Workplace*, *Pay*, *Marriage*, *Parenthood*, *Entrepreneurship*, *Assets*, and *Pension*."
- The WBL Index:
    - Theoretically ranges from 0 - 100
    - In practice: Lowest values ≈ 20
    - Higher values correspond to higher levels of women's empowerment and greater opportunities and support for women, particularly in business
- "Better performance in the areas measured by the Women, Business and the Law index is associated with a more narrow gender gap in development outcomes, higher female labor force participation, lower vulnerable employment, and greater representation of women in national parliaments."

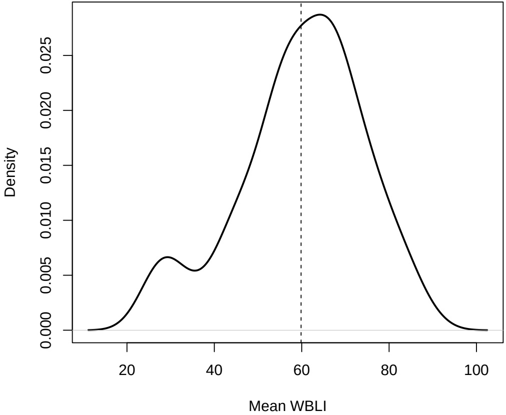
WBLI: Total Variation

```
> WDI<-pdata.frame(wdi)
> WBLI<-WDI$WomenBusLawIndex
> class(WBLI)
[1] "pseries" "numeric"

> describe(WBLI,na.rm=TRUE) # all variation

   vars     n  mean    sd median trimmed   mad  min max range skew kurtosis   se
X1    1 10152 59.85 18.74  59.38   59.86 19.46 17.5 100  82.5 0.02    -0.58 0.19
```



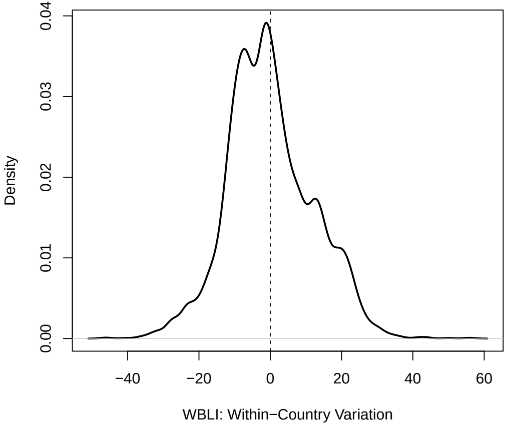
WBLI: "Between" Variation

```
> describe(plm::between(WBLI,effect="individual",na.rm=TRUE)) # "between" variation

   vars   n  mean    sd median trimmed   mad   min   max range skew kurtosis   se
X1    1 188 59.85 14.48  61.04   60.73 13.08 23.67 89.94 66.27 -0.5    -0.12 1.06
```



WBLI: "Within" Variation

```
> describe(Within(WBLI,na.rm=TRUE)) # "within" variation

   vars     n mean    sd median trimmed   mad    min   max range skew kurtosis   se
X1    1 10152    0 11.94  -1.25   -0.32 11.14 -46.06 55.88 101.9 0.26     0.35 0.12
```



A Regression Model

Regression model:

$$\begin{aligned} \text{WBLI}_{it} &= \beta_0 + \beta_1 \text{Population Growth}_{it} + \beta_2 \text{Urban Population}_{it} + \beta_3 \text{Fertility Rate}_{it} + \\ &\quad \beta_4 \ln(\text{GDP Per Capita})_{it} + \beta_5 \text{Natural Resource Rents}_{it} + \beta_6 \text{Cold War}_t + u_{it} \end{aligned}$$

Descriptive Statistics:

	vars	n	mean	sd	median	min	max	range	skew	kurtosis	se
WomenBusLawIndex	1	8257	60.39	19.07	60.62	17.50	100.00	82.50	-0.03	-0.69	0.21
PopGrowth	2	8257	1.67	1.64	1.66	-27.72	19.36	47.08	0.09	33.15	0.02
UrbanPopulation	3	8257	51.94	24.02	51.81	2.85	100.00	97.16	0.07	-1.04	0.26
FertilityRate	4	8257	3.63	1.90	3.10	0.77	8.61	7.83	0.51	-1.03	0.02
NaturalResourceRents	5	8257	7.25	11.17	2.47	0.00	88.59	88.59	2.52	7.52	0.12
ColdWar	6	8257	0.31	0.46	0.00	0.00	1.00	1.00	0.83	-1.32	0.01
lnGDPPerCap	7	8257	8.32	1.45	8.22	4.92	11.68	6.76	0.12	-0.91	0.02

Regression: Pooled OLS

```
> OLS<-plm(WomenBusLawIndex~PopGrowth+UrbanPopulation+FertilityRate+
+             log(GDPPerCapita)+NaturalResourceRents+ColdWar,
+           data=WDI,model="pooling")

> summary(OLS)
Pooling Model

Call:
plm(formula = WomenBusLawIndex ~ PopGrowth + UrbanPopulation +
    FertilityRate + log(GDPPerCapita) + NaturalResourceRents +
    ColdWar, data = WDI, model = "pooling")

Unbalanced Panel: n = 187, T = 1-52, N = 8257

Residuals:
   Min. 1st Qu.  Median 3rd Qu.    Max.
 -82.38   -8.53    1.09    9.30   45.84

Coefficients:
                     Estimate Std. Error t-value Pr(>|t|)
(Intercept)           60.7592     1.6694   36.40  < 2e-16 ***
PopGrowth             -1.8639     0.1202  -15.51  < 2e-16 ***
UrbanPopulation       -0.0691     0.0105   -6.59  4.5e-11 ***
FertilityRate         -2.8208     0.1556  -18.13  < 2e-16 ***
log(GDPPerCapita)      2.7018     0.1931   13.99  < 2e-16 ***
NaturalResourceRents  -0.3661     0.0151  -24.30  < 2e-16 ***
ColdWar              -10.5054     0.3675  -28.59  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    3000000
Residual Sum of Squares: 1500000
R-Squared:      0.501
Adj. R-Squared: 0.501
F-statistic: 1382.99 on 6 and 8250 DF, p-value: <2e-16
```

“Fixed” (Within) Effects

```
> FE<-plm(WomenBusLawIndex~PopGrowth+UrbanPopulation+FertilityRate+
+           log(GDPPerCapita)+NaturalResourceRents+ColdWar,data=WDI,
+           effect="individual",model="within")

> summary(FE)
Oneway (individual) effect Within Model

Call:
plm(formula = WomenBusLawIndex ~ PopGrowth + UrbanPopulation +
    FertilityRate + log(GDPPerCapita) + NaturalResourceRents +
    ColdWar, data = WDI, effect = "individual", model = "within")

Unbalanced Panel: n = 187, T = 1-52, N = 8257

Residuals:
   Min. 1st Qu.  Median 3rd Qu.    Max.
-33.789  -5.124  -0.489   4.960  53.525

Coefficients:
                     Estimate Std. Error t-value Pr(>|t|)
PopGrowth             -0.1493     0.0838   -1.78    0.075 .
UrbanPopulation        0.3092     0.0197   15.73  < 2e-16 ***
FertilityRate         -1.7285     0.1551  -11.14  < 2e-16 ***
log(GDPPerCapita)      8.9404     0.2980   30.00  < 2e-16 ***
NaturalResourceRents   0.0725     0.0167    4.35 0.000014 ***
ColdWar               -6.9613     0.2908  -23.94  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1070000
Residual Sum of Squares: 505000
R-Squared:      0.528
Adj. R-Squared: 0.516
F-statistic: 1501.91 on 6 and 8064 DF, p-value: <2e-16
```

A Nicer Table

Table: Models of WBLI

	OLS	FE
Population Growth	-1.864^{***}	-0.149^{*}
	(0.120)	(0.084)
Urban Population	-0.069^{***}	0.309^{***}
	(0.010)	(0.020)
Fertility Rate	-2.821^{***}	-1.728^{***}
	(0.156)	(0.155)
ln(GDP Per Capita)	2.702^{***}	8.940^{***}
	(0.193)	(0.298)
Natural Resource Rents	-0.366^{***}	0.072^{***}
	(0.015)	(0.017)
Cold War	-10.500^{***}	-6.961^{***}
	(0.367)	(0.291)
Constant	60.760^{***}	
	(1.669)	
Observations	8,257	8,257
R^2	0.501	0.528
Adjusted R^2	0.501	0.516
F Statistic	$1{,}383.000^{***}$ (df = 6; 8250)	$1{,}502.000^{***}$ (df = 6; 8064)

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Time-Period Fixed Effects

The model is:

$$Y_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \eta_t + u_{it}$$

which is estimated via:

$$\begin{aligned} Y_{it}^{**} &= Y_{it} - \bar{Y}_t \\ \mathbf{X}_{it}^{**} &= \mathbf{X}_{it} - \bar{\mathbf{X}}_t \end{aligned}$$

$$Y_{it}^{**} = \boldsymbol{\beta}_{FE}\mathbf{X}_{it}^{**} + u_{it}.$$

Comparison: Unit vs. Time Fixed Effects

Table: FE Models of WBLI (Units vs. Time)

	FE.Units	FE.Time
Population Growth	-0.149^{*}	-2.079^{***}
	(0.084)	(0.114)
Urban Population	0.309^{***}	-0.069^{***}
	(0.020)	(0.010)
Fertility Rate	-1.728^{***}	-1.856^{***}
	(0.155)	(0.151)
ln(GDP Per Capita)	8.940^{***}	3.107^{***}
	(0.298)	(0.183)
Natural Resource Rents	0.072^{***}	-0.401^{***}
	(0.017)	(0.014)
Cold War	-6.961^{***}	
	(0.291)	
Observations	8,257	8,257
R^2	0.528	0.406
Adjusted R^2	0.516	0.402
F Statistic	$1{,}502.000^{***}$ (df = 6; 8064)	$1{,}119.000^{***}$ (df = 5; 8200)

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

The specification:

$$Y_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \alpha_i + u_{it}$$

...suggests that we can use an F-test to examine the hypothesis:

$$H_0 : \alpha_i = 0 \; \forall \; i$$

(and a similar test for $\eta_t = 0$ in the time-centered case).

Arguably better are Lagrange multiplier-based tests:

- Breusch-Pagan (1980)
- King and Wu (1997)
- See (e.g.) Croissant and Millo (2018, §4.1) for details

FE (Country) Model Tests

```
> pFtest(FE,OLS)

 F test for individual effects

data:  WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +  ...
F = 85, df1 = 186, df2 = 8064, p-value <2e-16
alternative hypothesis: significant effects


> plmtest(FE,effect=c("individual"),type=c("bp"))

 Lagrange Multiplier Test - (Breusch-Pagan)

data:  WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +  ...
chisq = 55694, df = 1, p-value <2e-16
alternative hypothesis: significant effects


> plmtest(FE,effect=c("individual"),type=c("kw"))

 Lagrange Multiplier Test - (King and Wu)

data:  WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +  ...
normal = 236, p-value <2e-16
alternative hypothesis: significant effects
```

Fixed Effects: Interpretation

FE models are *subject-specific* models, that rely entirely on *within-unit* variability to estimate variable effects.

- This means that:

$$\hat{\beta}_k = \frac{\partial E(Y|\hat{\alpha})}{\partial X_k}$$

- That is, $\hat{\beta}_k$ is *the expected change in E(Y) associated with a one-unit increase in observation i's value of* X_k
- Key: *within-unit* changes in $\mathbf{X}$ are associated with *within-unit* expected changes in Y.
- In a linear model, the value of $\hat{\alpha}$ doesn't affect the value of that partial derivative...

Fixed Effects: Interpretation

Mummolo and Peterson (2018) note that:

> *"...because the within-unit variation is always smaller (or at least, no larger) than the overall variation in the independent variable, researchers should use within-unit variation to motivate counter-factuals when discussing the substantive impact of a treatment" (2018, 829).*

Significance:

- Predictors $\mathbf{X}$ in FE models typically have both cross-sectional and temporal variation
- FE models only consider *within-unit* variation in $\mathbf{X}$ and Y
- As a result, the degree of actual/observed (within-unit) variation in predictors is almost always less – and sometimes significantly less – than if cross-sectional variation were also considered

Interpretation Example: Urban Population

`UrbanPopulation` – All Variation:

```
> with(WDI, sd(UrbanPopulation,na.rm=TRUE)) # all variation
[1] 25.74
```

`UrbanPopulation` – "Within" Variation:

```
> WDI<-ddply(WDI, .(ISO3), mutate,
+            UPMean = mean(UrbanPopulation,na.rm=TRUE))
> WDI$UPWithin<-with(WDI, UrbanPopulation-UPMean)
>
> with(WDI, sd(UPWithin,na.rm=TRUE)) # "within" variation
[1] 9.13
```

"While the overall variation in the independent variable may be large, the within-unit variation used to estimate β may be much smaller" (M & P 2018, 830).

Pros and Cons of “Fixed” Effects

Pros:

- Mitigates (Some) Specification Bias
- Simple + Intuitive
- Widely Used/Understood

Cons (see e.g. Collischon and Eberl 2020):

- Can't Estimate $\boldsymbol{\beta}_B$
- Slowly-Changing **X**s
- (In)Efficiency / Inconsistency (Incidental Parameters)
- Cannot Control for (e.g.) Time-Varying Heterogeneity
- Sensitivity to Measurement Error

"Between" Effects

From the equation above:

$$Y_{it} = \bar{\mathbf{X}}_i \boldsymbol{\beta}_B + \tilde{\mathbf{X}}_{it} \boldsymbol{\beta}_W + \alpha_i + u_{it}$$

...we can derive a "Between Effects" model:

$$\bar{Y}_i = \bar{\mathbf{X}}_i \boldsymbol{\beta}_B + u_{it}$$

This model:

- is essentially cross-sectional,
- is based on N observations,
- considers *only* between-unit (average) differences
- Interpretation:

  $\hat{\beta}_B$ is the expected difference in Y between two units whose values on $\bar{X}$ differ by a value of 1.0.

“Between” Effects

```
> BE<-plm(WomenBusLawIndex~PopGrowth+UrbanPopulation+FertilityRate+
+           log(GDPPerCapita)+NaturalResourceRents+ColdWar,data=WDI,
+         effect="individual",model="between")

> summary(BE)
Oneway (individual) effect Between Model

Call:
plm(formula = WomenBusLawIndex ~ PopGrowth + UrbanPopulation +
    FertilityRate + log(GDPPerCapita) + NaturalResourceRents +
    ColdWar, data = WDI, effect = "individual", model = "between")

Unbalanced Panel: n = 187, T = 1-52, N = 8257
Observations used in estimation: 187

Residuals:
   Min. 1st Qu.  Median 3rd Qu.    Max.
-29.952  -6.630   0.921   8.034  21.403

Coefficients:
                     Estimate Std. Error t-value    Pr(>|t|)
(Intercept)           61.7112    10.5532    5.85 0.000000023 ***
PopGrowth             -5.2836     1.1363   -4.65 0.000006407 ***
UrbanPopulation       -0.0491     0.0552   -0.89     0.37513
FertilityRate         -1.0921     1.1845   -0.92     0.35776
log(GDPPerCapita)      2.2886     1.1533    1.98     0.04873 *
NaturalResourceRents  -0.3237     0.0910   -3.56     0.00048 ***
ColdWar               -8.4268     5.3216   -1.58     0.11506
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    44000
Residual Sum of Squares: 18400
R-Squared:      0.582
Adj. R-Squared: 0.568
F-statistic: 41.7756 on 6 and 180 DF, p-value: <2e-16
```

A Nicer Table (Again)

Table: Models of WBLI

	OLS	FE	BE
Population Growth	-1.864^{***}	-0.149^{*}	-5.284^{***}
	(0.120)	(0.084)	(1.136)
Urban Population	-0.069^{***}	0.309^{***}	-0.049
	(0.010)	(0.020)	(0.055)
Fertility Rate	-2.821^{***}	-1.728^{***}	-1.092
	(0.156)	(0.155)	(1.184)
ln(GDP Per Capita)	2.702^{***}	8.940^{***}	2.289^{**}
	(0.193)	(0.298)	(1.153)
Natural Resource Rents	-0.366^{***}	0.072^{***}	-0.324^{***}
	(0.015)	(0.017)	(0.091)
Cold War	-10.500^{***}	-6.961^{***}	-8.427
	(0.367)	(0.291)	(5.322)
Constant	60.760^{***}		61.710^{***}
	(1.669)		(10.550)
Observations	8,257	8,257	187
R^2	0.501	0.528	0.582
Adjusted R^2	0.501	0.516	0.568
F Statistic	$1{,}383.000^{***}$ (df = 6; 8250)	$1{,}502.000^{***}$ (df = 6; 8064)	41.780^{***} (df = 6; 180)

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Model:

$$Y_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + u_{it}$$

with:

$$u_{it} = \alpha_i + \lambda_t + \eta_{it}$$

and

$$\begin{aligned}
E(\alpha_i) = E(\lambda_t) = E(\eta_{it}) &= 0, \\
E(\alpha_i\lambda_t) = E(\alpha_i\eta_{it}) = E(\lambda_t\eta_{it}) &= 0, \\
E(\alpha_i\alpha_j) &= \sigma^2_\alpha \text{ if } i = j, \ 0 \text{ otherwise}, \\
E(\lambda_t\lambda_s) &= \sigma^2_\lambda \text{ if } t = s, \ 0 \text{ otherwise}, \\
E(\eta_{it}\eta_{js}) &= \sigma^2_\eta \text{ if } i = j, \ t = s, \ 0 \text{ otherwise}, \\
E(\alpha_i\mathbf{X}_{it}) = E(\lambda_t\mathbf{X}_{it}) = E(\eta_{it}\mathbf{X}_{it}) &= 0.
\end{aligned}$$

If those assumptions are met, we can consider the "two-way variance components" model where:

$$\begin{aligned} Var(u_{it}) &= Var(Y_{it}|\mathbf{X}_{it}) \\ &= \sigma^2_\alpha + \sigma^2_\lambda + \sigma^2_\eta \end{aligned}$$

If we assume $\lambda_t = 0$, then we get a model like:

$$Y_{it} = \mathbf{X}_{it}\beta + \alpha_i + \eta_{it}$$

with total error variance:

$$\sigma^2_u = \sigma^2_\alpha + \sigma^2_\eta.$$

“Random” Effects: Estimation

The model above will violate the standard OLS assumptions of uncorrelated errors, because the (compound) “errors” u_{it} within each unit share a common component α_i.

Consider the within-i variance-covariance matrix of the errors $\mathbf{u}$:

$$
\begin{aligned}
E(\mathbf{u}_i\mathbf{u}_i') \equiv \mathbf{\Sigma}_i &= \sigma_\eta^2 \mathbf{I}_T + \sigma_\alpha^2 \mathbf{i}\mathbf{i}' \\
&= \begin{pmatrix} \sigma_\eta^2 + \sigma_\alpha^2 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\eta^2 + \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \cdots & \sigma_\eta^2 + \sigma_\alpha^2 \end{pmatrix}
\end{aligned}
$$

Assuming conditional independence across units, we then have:

$$
\mathrm{Var}(\mathbf{u}) \equiv \mathbf{\Omega} = \begin{pmatrix} \mathbf{\Sigma}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{\Sigma}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{\Sigma}_N \end{pmatrix}
$$

"Random" Effects: Estimation

We can then show that:

$$\mathbf{\Sigma}^{-1/2} = \frac{1}{\sigma_\eta}\left[\mathbf{I}_T - \left(\frac{\theta}{T}\mathbf{i}\mathbf{i}'\right)\right]$$

where

$$\theta = 1 - \sqrt{\frac{\sigma_\eta^2}{T\sigma_\alpha^2 + \sigma_\eta^2}}$$

is an unknown quantity to be estimated.

Starting with an estimate of $\hat{\theta}$, calculate:

$$\begin{aligned} Y_{it}^* &= Y_{it} - \hat{\theta}\bar{Y}_i \\ X_{it}^* &= X_{it} - \hat{\theta}\bar{X}_i, \end{aligned}$$

then estimate:

$$Y_{it}^* = (1-\hat{\theta})\alpha + X_{it}^*\beta_{RE} + [(1-\hat{\theta})\alpha_i + (\eta_{it} - \hat{\theta}\bar{\eta}_i)]$$

and iterate between the two processes until convergence.

"Random" Effects: An Alternative View

Random Effects

```
> RE<-plm(WomenBusLawIndex~PopGrowth+UrbanPopulation+FertilityRate+
+           log(GDPPerCapita)+NaturalResourceRents+ColdWar,data=WDI,
+         effect="individual",model="random")

> summary(RE)
Oneway (individual) effect Random Effect Model
   (Swamy-Arora's transformation)

Call:
plm(formula = WomenBusLawIndex ~ PopGrowth + UrbanPopulation +
    FertilityRate + log(GDPPerCapita) + NaturalResourceRents +
    ColdWar, data = WDI, effect = "individual", model = "random")

Unbalanced Panel: n = 187, T = 1-52, N = 8257

Effects:
                var std.dev share
idiosyncratic 62.61    7.91   0.4
individual    95.18    9.76   0.6
theta:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.370   0.876   0.888   0.880   0.888   0.888

Residuals:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  -34.0    -5.4    -0.5     0.0     5.6    44.7

Coefficients:
                     Estimate Std. Error z-value Pr(>|z|)
(Intercept)           -0.1918     2.5251   -0.08   0.9395
PopGrowth             -0.1916     0.0856   -2.24   0.0253 *
UrbanPopulation        0.1907     0.0185   10.31   <2e-16 ***
FertilityRate         -2.0719     0.1550  -13.37   <2e-16 ***
log(GDPPerCapita)      7.3542     0.2824   26.04   <2e-16 ***
NaturalResourceRents   0.0449     0.0167    2.68   0.0073 **
ColdWar               -8.2289     0.2854  -28.83   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    1110000
Residual Sum of Squares: 543000
R-Squared:      0.511
Adj. R-Squared: 0.511
Chisq: 8408.9 on 6 DF, p-value: <2e-16
```

A Nicer Table (Yet Again)

Table: Models of WBLI

	OLS	FE	BE	RE
Population Growth	-1.864^{***}	-0.149^{*}	-5.284^{***}	-0.192^{**}
	(0.120)	(0.084)	(1.136)	(0.086)
Urban Population	-0.069^{***}	0.309^{***}	-0.049	0.191^{***}
	(0.010)	(0.020)	(0.055)	(0.019)
Fertility Rate	-2.821^{***}	-1.728^{***}	-1.092	-2.072^{***}
	(0.156)	(0.155)	(1.184)	(0.155)
ln(GDP Per Capita)	2.702^{***}	8.940^{***}	2.289^{**}	7.354^{***}
	(0.193)	(0.298)	(1.153)	(0.282)
Natural Resource Rents	-0.366^{***}	0.072^{***}	-0.324^{***}	0.045^{***}
	(0.015)	(0.017)	(0.091)	(0.017)
Cold War	-10.500^{***}	-6.961^{***}	-8.427	-8.229^{***}
	(0.367)	(0.291)	(5.322)	(0.285)
Constant	60.760^{***}		61.710^{***}	-0.192
	(1.669)		(10.550)	(2.525)
Observations	8,257	8,257	187	8,257
R^2	0.501	0.528	0.582	0.511
Adjusted R^2	0.501	0.516	0.568	0.511
F Statistic	$1{,}383.000^{***}$ (df = 6; 8250)	$1{,}502.000^{***}$ (df = 6; 8064)	41.780^{***} (df = 6; 180)	$8{,}409.000^{***}$

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Intuition:

- RE models require that $\text{Cov}(X_{it}, \alpha_i) = 0$.
- FE models do not.

This means that:

	Reality	
Model	$\text{Cov}(X_{it}, \alpha_i) = 0$	$\text{Cov}(X_{it}, \alpha_i) \neq 0$
Fixed Effects	Consistent, Inefficient	Consistent, Efficient
Random Effects	Consistent, Efficient	Inconsistent

The Hausman Test

Hausman test (FE vs. RE):

$$\hat{W} = (\hat{\boldsymbol{\beta}}_{\mathbf{FE}} - \hat{\boldsymbol{\beta}}_{\mathbf{RE}})'(\hat{\mathbf{V}}_{\mathbf{FE}} - \hat{\mathbf{V}}_{\mathbf{RE}})^{-1}(\hat{\boldsymbol{\beta}}_{\mathbf{FE}} - \hat{\boldsymbol{\beta}}_{\mathbf{RE}})$$

$$W \sim \chi^2_k$$

Null: The RE model is consistent ($\text{Cov}(X_{it}, \alpha_i) = 0$).

Issues:

- Asymptotic
- No guarantee $(\hat{\mathbf{V}}_{\mathbf{FE}} - \hat{\mathbf{V}}_{\mathbf{RE}})^{-1}$ is positive definite
- A general specification test...

Hausman Test Results

Hausman test (FE vs. RE):

```
> phtest(FE, RE)

 Hausman Test

data:  WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +  ...
chisq = 2396, df = 6, p-value <2e-16

alternative hypothesis: one model is inconsistent
```

Practical “Fixed” vs. “Random” Effects

Factors to consider:

- “Panel” vs. “TSCS” Data
- Nature of the Data-Generating Processes
- Importance of Non-Time-Varying Covariate Effects
- Dimension/Location of the Variance in the Data (N vs. T)

Connections: Hierarchical Linear Models

Begin by considering a two-level "nested" data structure, with:

$$\begin{aligned} i &\in \{1, 2, ...N\} \text{ indexing first-level units, and} \\ j &\in \{1, 2, ...J\} \text{ indexing second-level groups.} \end{aligned}$$

A general two-level HLM is an equation of the form:

$$Y_{ij} = \beta_{0j} + \mathbf{X}_{ij}\boldsymbol{\beta}_j + u_{ij} \tag{1}$$

where β_{0j} is a "constant" term, $\mathbf{X}_{ij}$ is a $NJ \times K$ matrix of K covariates, $\boldsymbol{\beta}_j$ is a $K \times 1$ vector of parameters, and $u_{ij} \sim$ i.i.d. $N(0, \sigma_u^2)$ is the usual random-disturbance assumption.

Each of the $K+1$ "level-one" parameters is then allowed to vary across Q "level-two" variables $\mathbf{Z}_j$, so that:

$$\beta_{0j} = \gamma_{00} + \mathbf{Z}_j \gamma_0 + \varepsilon_{0j} \tag{2}$$

for the "intercept" and

$$\beta_{kj} = \gamma_{k0} + \mathbf{Z}_j \gamma_k + \varepsilon_{kj} \tag{3}$$

for the "slopes" of $\mathbf{X}$. The εs are typically assumed to be distributed multivariate Normal, with parameters for the variances and covariances selected by the analyst. Substitution of (2) and (3) into (1) yields:

$$Y_{ij} = \gamma_{00} + \mathbf{Z}_j \gamma_0 + \mathbf{X}_{ij} \gamma_{k0} + \mathbf{X}_{ij} \mathbf{Z}_j \gamma_k + \mathbf{X}_{ij} \varepsilon_{kj} + \varepsilon_{0j} + u_{ij} \tag{4}$$

The form is essentially a model with "saturated" interaction effects across the various levels, as well as "errors" which are multivariate Normal.

HLM Details

Model Assumptions

- Linearity / Additivity
- Normality of *u*s
- Homoscedasticity
- Residual Independence:
  - $\text{Cov}(\varepsilon_{\cdot j}, u_{ij}) = 0$
  - $\text{Cov}(u_{ij}, u_{\ell j}) = 0$

Model Fitting

- MLE
- "Restricted" MLE ("RMLE")
- Choosing:
  - MLE is biased in small samples, especially for estimating variances
  - RMLE is not, but prevents use of LR tests when the models do not have identical fixed effects
  - In general: RMLE is better with small sample sizes, but MLE is fine in larger ones

HLMs: Attributes

Note that if we specify:

$$\beta_{0j} = \gamma_{00} + \varepsilon_{0j}$$

and

$$\beta_{kj} = \gamma_{k0}$$

we get:

$$Y_{ij} = \gamma_{00} + \mathbf{X}_{ij}\gamma_{k0} + \varepsilon_{0j} + u_{ij}$$

which is the RE model above (formally, a "one-level random-intercept" HLM).

In addition:

- HLMs can be expanded to 3- and 4- and higher-level models
- One can include cross-level interactions...

HLMs are widely used in education, psychology, ecology, etc. (less so in economics, political science). Also, there are many, many excellent books, websites, etc. that address HLMs

Random Effects Remix (using `lmer`)

```
> library(lme4)
> AltRE<-lmer(WomenBusLawIndex~PopGrowth+UrbanPopulation+FertilityRate+
+               log(GDPPerCapita)+NaturalResourceRents+ColdWar+(1|ISO3),
+             data=WDI)

> summary(AltRE)
Linear mixed model fit by REML ['lmerMod']
Formula: WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +
    log(GDPPerCapita) + NaturalResourceRents + ColdWar + (1 |      ISO3)
   Data: WDI

REML criterion at convergence: 58628

Scaled residuals:
   Min     1Q Median     3Q    Max
-4.273 -0.655 -0.070  0.635  6.662

Random effects:
 Groups   Name        Variance Std.Dev.
 ISO3     (Intercept) 342.7    18.51
 Residual              62.7     7.92
Number of obs: 8257, groups:  ISO3, 187

Fixed effects:
                     Estimate Std. Error t value
(Intercept)          -13.9553     2.8768   -4.85
PopGrowth             -0.1607     0.0838   -1.92
UrbanPopulation        0.2679     0.0191   13.99
FertilityRate         -1.8416     0.1540  -11.96
log(GDPPerCapita)      8.3801     0.2908   28.82
NaturalResourceRents   0.0644     0.0166    3.89
ColdWar               -7.4173     0.2867  -25.87

Correlation of Fixed Effects:
            (Intr) PpGrwt UrbnPp FrtltR l(GDPP NtrlRR
PopGrowth    0.057
UrbanPopltn -0.191  0.010
FertilityRt -0.395 -0.242  0.428
lg(GDPPrCp) -0.758 -0.069 -0.288  0.101
NtrlRsrcRnt -0.003 -0.095 -0.025 -0.112 -0.008
ColdWar     -0.126  0.024  0.218 -0.429  0.121  0.058
```

Q: Are They The Same? [A: More Or Less]

Table: RE and HLM Models of WBLI

	RE	AltRE
Population Growth	-0.192^{**}	-0.161^{*}
	(0.086)	(0.084)
Urban Population	0.191^{***}	0.268^{***}
	(0.019)	(0.019)
Fertility Rate	-2.072^{***}	-1.842^{***}
	(0.155)	(0.154)
ln(GDP Per Capita)	7.354^{***}	8.380^{***}
	(0.282)	(0.291)
Natural Resource Rents	0.045^{***}	0.064^{***}
	(0.017)	(0.017)
Cold War	-8.229^{***}	-7.417^{***}
	(0.285)	(0.287)
Constant	-0.192	-13.960^{***}
	(2.525)	(2.877)
Observations	8,257	8,257
R^2	0.511	
Adjusted R^2	0.511	
Log Likelihood		-29,314.000
Akaike Inf. Crit.		58,646.000
Bayesian Inf. Crit.		58,710.000
F Statistic	$8,409.000^{***}$	

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

For more about why they're a bit different, see here.

HLM with Country-Level Random βs for `ColdWar`

```
> HLM1<-lmer(WomenBusLawIndex~PopGrowth+UrbanPopulation+FertilityRate+
+              log(GDPPerCapita)+NaturalResourceRents+ColdWar+(ColdWar|ISO3),
+            data=WDI,control=lmerControl(optimizer="bobyqa"))

> summary(HLM1)
Linear mixed model fit by REML ['lmerMod']
Formula: WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +
    log(GDPPerCapita) + NaturalResourceRents + ColdWar + (ColdWar |      ISO3)
   Data: WDI
Control: lmerControl(optimizer = "bobyqa")

REML criterion at convergence: 55110

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 ISO3     (Intercept) 575.8    24.00
          ColdWar     142.0    11.92    -0.20
 Residual              37.5     6.12
Number of obs: 8257, groups:  ISO3, 187

Fixed effects:
                     Estimate Std. Error t value
(Intercept)          -29.2427     3.2694   -8.94
PopGrowth             -0.3020     0.0674   -4.48
UrbanPopulation        0.3361     0.0223   15.09
FertilityRate         -3.9182     0.1673  -23.42
log(GDPPerCapita)     10.5996     0.3198   33.14
NaturalResourceRents   0.0119     0.0144    0.82
ColdWar               -2.3983     1.0063   -2.38

Correlation of Fixed Effects:
            (Intr) PpGrwt UrbnPp FrtltR l(GDPP NtrlRR
PopGrowth    0.063
UrbanPopltn -0.149 -0.003
FertilityRt -0.479 -0.169  0.484
lg(GDPPrCp) -0.707 -0.076 -0.370  0.174
NtrlRsrcRnt  0.021 -0.076  0.058 -0.047 -0.078
ColdWar     -0.097 -0.004  0.050 -0.115  0.002  0.010
```

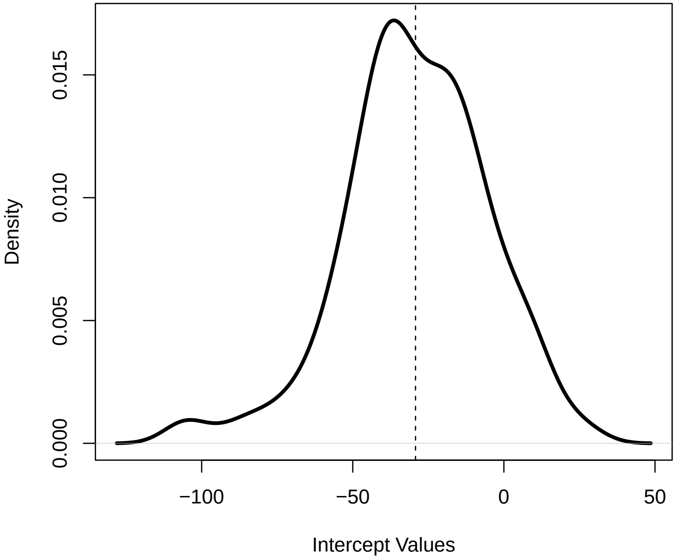
Testing

```
> anova(AltRE,HLM1)

refitting model(s) with ML (instead of REML)

Data: WDI
Models:
AltRE: WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +
       log(GDPPerCapita) + NaturalResourceRents + ColdWar + (1 | ISO3)
HLM1: WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +
       log(GDPPerCapita) + NaturalResourceRents + ColdWar + (ColdWar | ISO3)

      npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
AltRE    9 58629 58692 -29305    58611
HLM1    11 55118 55195 -27548    55096  3515  2     <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


> VarCorr(HLM1)
 Groups   Name        Std.Dev. Corr
 ISO3     (Intercept) 24.00
          ColdWar     11.92    -0.20
 Residual              6.12
```

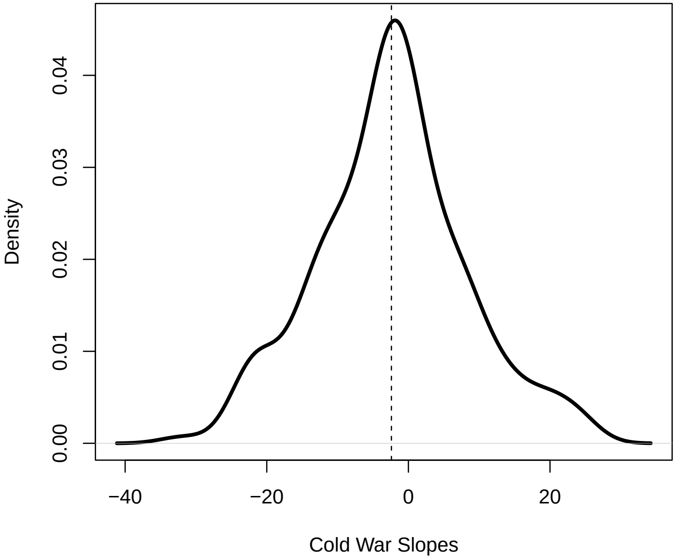
Random Coefficients

```
> Bs<-data.frame(coef(HLM1)[1])

> head(Bs)
    ISO3..Intercept. ISO3.PopGrowth ISO3.UrbanPopulation ISO3.FertilityRate
AFG           -18.16         -0.302               0.3361             -3.918
AGO           -11.74         -0.302               0.3361             -3.918
ALB           -12.48         -0.302               0.3361             -3.918
ARE          -102.89         -0.302               0.3361             -3.918
ARG           -50.45         -0.302               0.3361             -3.918
ARM           -24.22         -0.302               0.3361             -3.918
    ISO3.log.GDPPerCapita. ISO3.NaturalResourceRents ISO3.ColdWar
AFG                   10.6                   0.01186       -3.501
AGO                   10.6                   0.01186      -13.877
ALB                   10.6                   0.01186       -6.422
ARE                   10.6                   0.01186       -3.097
ARG                   10.6                   0.01186      -22.810
ARM                   10.6                   0.01186       -2.898

> mean(Bs$ISO3..Intercept.)
[1] -29.24

> mean(Bs$ISO3.ColdWar)
[1] -2.398
```

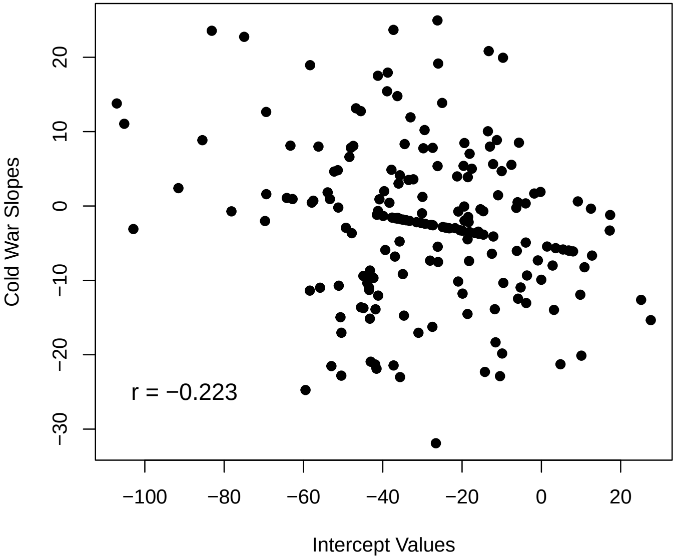
Random Intercepts (Plotted)



Random Slopes for `ColdWar` (Plotted)



Scatterplot: Random Intercepts and Slopes



Separating Within and Between Effects

Recall that we can decompose the variance of our regression model as:

$$Y_{it} = \bar{\mathbf{X}}_i \boldsymbol{\beta}_B + (\mathbf{X}_{it} - \bar{\mathbf{X}}_i)\boldsymbol{\beta}_W + u_{it} \tag{5}$$

This raises the possibility of fitting the model in (5) directly...

- Simple to estimate (via OLS or other)
- Relatively easy interpretation
- Easy to test $\hat{\boldsymbol{\beta}}_B = \hat{\boldsymbol{\beta}}_W$

Example data: Separate effects for within- and between-country *Natural Resource Rents*...

Combining Within- and Between-Effects

Table: BE + WE Model of WBLI

	WEBE.OLS
Population Growth	-1.728^{***}
	(0.117)
Urban Population	-0.049^{***}
	(0.010)
Fertility Rate	-2.465^{***}
	(0.153)
ln(GDP Per Capita)	2.683^{***}
	(0.188)
Within-Country Nat. Resource Rents	0.109^{***}
	(0.027)
Between-Country Nat. Resource Rents	-0.545^{***}
	(0.017)
Cold War	-11.170^{***}
	(0.360)
Constant	59.860^{***}
	(1.629)
Observations	8,257
R^2	0.526
Adjusted R^2	0.525
Residual Std. Error	13.140 (df = 8249)
F Statistic	$1{,}307.000^{***}$ (df = 7; 8249)

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Our original decomposition considered "two-way" effects:

$$Y_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \alpha_i + \eta_t + u_{it}$$

This implies that we can use (e.g.) an F-test to examine the hypothesis:

$$H_0 : \alpha_i = \eta_t = 0 \; \forall \; i, \; t$$

...that is, whether adding the (two-way) effects improves the model's fit.

We can also consider the partial hypotheses:

$$H_0 : \alpha_i = 0 \; \forall \; i$$

and

$$H_0 : \eta_t = 0 \; \forall \; t$$

separately.

Two-Way Effects: Good & Bad

The Good:

- Addresses both time-invariant, unit-level heterogeneity and cross-sectionally-invariant, time-point-specific heterogeneity
- May be "fixed" or "random"...
- Two-way FE is equivalent to differences-in-differences when $X \in \{0, 1\}$ and $T = 2$ (more on that later)

The (Potentially) Bad:

- Ability to control for unobserved heterogeneity depends on (probably invalid) functional form assumptions (Imai and Kim 2020)
- FE *requires* predictors that vary both within and between both units and time points
- RE requires the (additional) assumption that $\text{Cov}(\mathbf{X}_{it}, \eta_t) = Cov(\alpha_i, \eta_t) = 0$
- Two-way effects models ask a *lot* of your data (effectively fits $N + T + k$ parameters using NT observations)

Example: Two-Way Fixed Effects

```
> TwoWayFE<-plm(WomenBusLawIndex~PopGrowth+UrbanPopulation+FertilityRate+
+                 log(GDPPerCapita)+NaturalResourceRents+ColdWar,data=WDI,
+               effect="twoway",model="within")

> summary(TwoWayFE)
Twoways effects Within Model

Call:
plm(formula = WomenBusLawIndex ~ PopGrowth + UrbanPopulation +
    FertilityRate + log(GDPPerCapita) + NaturalResourceRents +
    ColdWar, data = WDI, effect = "twoway", model = "within")

Unbalanced Panel: n = 187, T = 1-52, N = 8257

Residuals:
   Min. 1st Qu.  Median 3rd Qu.    Max.
-32.030  -4.082   0.243   4.168  43.647

Coefficients:
                     Estimate Std. Error t-value Pr(>|t|)
PopGrowth             -0.2653     0.0702   -3.78  0.00016 ***
UrbanPopulation        0.0305     0.0173    1.77  0.07733 .
FertilityRate          1.3623     0.1418    9.61  < 2e-16 ***
log(GDPPerCapita)      2.0709     0.2756    7.51  6.4e-14 ***
NaturalResourceRents   0.0354     0.0145    2.43  0.01494 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    357000
Residual Sum of Squares: 349000
R-Squared:      0.0211
Adj. R-Squared: -0.00846
F-statistic: 34.5516 on 5 and 8014 DF, p-value: <2e-16
```

```
> TwoWayFE.BF<-lm(WomenBusLawIndex~PopGrowth+UrbanPopulation+
+                   FertilityRate+log(GDPPerCapita)+NaturalResourceRents+
+                   factor(ISO3)+factor(Year),data=WDI)

> summary(TwoWayFE.BF)

Call:
lm(formula = WomenBusLawIndex ~ PopGrowth + UrbanPopulation +
    FertilityRate + log(GDPPerCapita) + NaturalResourceRents +
    factor(ISO3) + factor(Year), data = WDI)

Residuals:
   Min     1Q Median     3Q    Max
-32.03  -4.08   0.24   4.17  43.65

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)          -17.0883     2.4379   -7.01  2.6e-12 ***
PopGrowth             -0.2653     0.0702   -3.78  0.00016 ***
UrbanPopulation        0.0305     0.0173    1.77  0.07733 .
FertilityRate          1.3623     0.1418    9.61  < 2e-16 ***
log(GDPPerCapita)      2.0709     0.2756    7.51  6.4e-14 ***
NaturalResourceRents   0.0354     0.0145    2.43  0.01494 *
factor(ISO3)AGO       28.4399     1.9617   14.50  < 2e-16 ***
factor(ISO3)ALB       51.9612     1.9960   26.03  < 2e-16 ***
factor(ISO3)ARE       -4.5542     2.4751   -1.84  0.06581 .
.
.
.
factor(Year)1977       4.2020     0.8814    4.77  1.9e-06 ***
factor(Year)1978       5.1089     0.8819    5.79  7.2e-09 ***
 [ reached getOption("max.print") -- omitted 43 rows ]
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.6 on 8014 degrees of freedom
  (5503 observations deleted due to missingness)
Multiple R-squared:  0.884, Adjusted R-squared:  0.88
F-statistic:  252 on 242 and 8014 DF,  p-value: <2e-16
```

Example: Two-Way Random Effects

```
> TwoWayRE<-plm(WomenBusLawIndex~PopGrowth+UrbanPopulation+FertilityRate+log(GDPPerCapita)+NaturalResourceRents+ColdWar,data=WDI,
+               effect="twoway",model="random")

> summary(TwoWayRE)
Twoways effects Random Effect Model
   (Swamy-Arora's transformation)

Call:
plm(formula = WomenBusLawIndex ~ PopGrowth + UrbanPopulation +
    FertilityRate + log(GDPPerCapita) + NaturalResourceRents +
    ColdWar, data = WDI, effect = "twoway", model = "random")

Unbalanced Panel: n = 187, T = 1-52, N = 8257

Effects:
                 var std.dev share
idiosyncratic 43.555   6.600  0.31
individual    95.615   9.778  0.69
time           0.381   0.617  0.00
theta:
        Min. 1st Qu. Median   Mean 3rd Qu.   Max.
id    0.4406  0.8964 0.9068 0.9002  0.9068 0.9068
time  0.2846  0.3336 0.3737 0.3563  0.3790 0.3821
total 0.2772  0.3331 0.3730 0.3555  0.3783 0.3814

Coefficients:
                      Estimate Std. Error z-value Pr(>|z|)
(Intercept)           15.17580    0.33458    45.4   <2e-16 ***
PopGrowth             -0.23222    0.01059   -21.9   <2e-16 ***
UrbanPopulation        0.13538    0.00237    57.1   <2e-16 ***
FertilityRate         -0.65175    0.02003   -32.5   <2e-16 ***
log(GDPPerCapita)      5.34377    0.03672   145.5   <2e-16 ***
NaturalResourceRents   0.03851    0.00212    18.1   <2e-16 ***
ColdWar              -11.73699    0.04687  -250.4   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    3000000
Residual Sum of Squares: 2200000
R-Squared:      0.296
Adj. R-Squared: 0.295
Chisq: 238275 on 6 DF, p-value: <2e-16
```

A Prettier Table

Table: Models of WBLI

	OLS	FE	BE	RE	TwoWayFE	TwoWayRE
Population Growth	−1.864***	−0.149*	−5.284***	−0.192**	−0.265***	−0.232***
	(0.120)	(0.084)	(1.136)	(0.086)	(0.070)	(0.011)
Urban Population	−0.069***	0.309***	−0.049	0.191***	0.031*	0.135***
	(0.010)	(0.020)	(0.055)	(0.019)	(0.017)	(0.002)
Fertility Rate	−2.821***	−1.728***	−1.092	−2.072***	1.362***	−0.652***
	(0.156)	(0.155)	(1.184)	(0.155)	(0.142)	(0.020)
ln(GDP Per Capita)	2.702***	8.940***	2.289**	7.354***	2.071***	5.344***
	(0.193)	(0.298)	(1.153)	(0.282)	(0.276)	(0.037)
Natural Resource Rents	−0.366***	0.072***	−0.324***	0.045***	0.035**	0.039***
	(0.015)	(0.017)	(0.091)	(0.017)	(0.015)	(0.002)
Cold War	−10.500***	−6.961***	−8.427	−8.229***		−11.740***
	(0.367)	(0.291)	(5.322)	(0.285)		(0.047)
Constant	60.760***		61.710***	−0.192		15.180***
	(1.669)		(10.550)	(2.525)		(0.335)
Observations	8,257	8,257	187	8,257	8,257	8,257
R^2	0.501	0.528	0.582	0.511	0.021	0.296
Adjusted R^2	0.501	0.516	0.568	0.511	-0.008	0.295

*p<0.1; **p<0.05; ***p<0.01

Other Variations: FEIS

“Fixed Effects Individual Slope” models

- Cite: Bruederl, Josef, and Volker Ludwig. 2015. “Fixed-Effects Panel Regression.” In *The Sage Handbook of Regression Analysis and Causal Inference*, Eds. Henning Best and Christof Wolf. Los Angeles: Sage, pp. 327-357.
- FE + unit-level slopes for (some / all) predictor variables
- Equivalent to including $N-1$ interactions between a predictor $\mathbf{X}$ and each of the α_is
- Also can test for homogeneity of estimated slopes (Hausman-like test)
- See the `feisr` R package, and its accompanying vignette, or `xtfeis` in Stata

Unit Effects Models: Software

R :

- the `plm` package; `plm` command
    - Fits one- and two-way FE, BE, RE models
    - Also fits first difference (FD) and instrumental variable (IV) models
- the `fixest` package; fast/scalable FE estimation for OLS and GLMs
- the `lme4` package; command is `lmer`
- the `nlme` package; command `lme`
- the `Paneldata` package

Stata : `xtreg`

- option `re` (the default) = random effects
- option `fe` = fixed (within) effects
- option `be` = between-effects
- Stata package `fect` = two-way models

Dynamics

Issues with Unit Roots in Panel Data

In general, in panel / TSCS data:

- Short series + Asymptotic tests → "borrow strength"
- Requires uniform unit roots across cross-sectional units
- Many tests require balanced panels...
- Various alternatives:
  - Maddala and Wu (1999)
  - Hadri (2000)
  - Levin, Lin and Chu (2002)
  - Im, Pesaran, and Shin (2003)
- What to do?
  - Difference the data...
  - Error-correction models

Panel Unit Root Tests: R

```
[data wrangling...]

> purtest(WBLI.W,exo="trend",test="levinlin",pmax=2)

 Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts and Trend)

data:  WBLI.W
z = -2.5, p-value = 0.007
alternative hypothesis: stationarity

> purtest(WBLI.W,exo="trend",test="hadri",pmax=2)

 Hadri Test (ex. var.: Individual Intercepts and Trend) (Heterosked.
 Consistent)

data:  WBLI.W
z = 200, p-value <2e-16
alternative hypothesis: at least one series has a unit root

> purtest(WBLI.W,exo="trend",test="madwu",pmax=2)

 Maddala-Wu Unit-Root Test (ex. var.: Individual Intercepts and Trend)

data:  WBLI.W
chisq = 331, df = 376, p-value = 1
alternative hypothesis: stationarity

> purtest(WBLI.W,exo="trend",test="ips",pmax=2)

 Im-Pesaran-Shin Unit-Root Test (ex. var.: Individual Intercepts and Trend)

data:  WBLI.W
Wtbar = 2.9, p-value = 1
alternative hypothesis: stationarity
```

A Better Table

Table: Panel Unit Root Tests: WBRI

	Test	Alternative	Statistic	Estimate	P-Value
1	Levin-Lin-Chu	stationarity	z	-2.476	0.0066
2	Hadri	at least one series has a unit root	z	199.634	0
3	Maddala-Wu	stationarity	chisq	330.698	0.9553
4	Im-Pesaran-Shin	stationarity	Wtbar	2.851	0.9978

Note: All assume individual intercepts and trends.

Consider a model like this:

$$Y_{it} = \phi Y_{it-1} + \mathbf{X}_{it}\boldsymbol{\beta}_{LDV} + \epsilon_{it}$$

If ϵ_{it} is perfect...

- $\hat{\beta}_{LDV}$ is biased (but consistent),
- O(bias) $= \frac{-1+3\boldsymbol{\beta}_{LDV}}{T}$

If ϵ_{it} is autocorrelated...

- $\hat{\beta}_{LDV}$ is biased and <u>inconsistent</u>
- IV is one (bad) option...

Lagged Ys and GLS-ARMA

Can rewrite:

$$\begin{aligned} Y_{it} &= \mathbf{X}_{it}\boldsymbol{\beta}_{AR} + u_{it} \\ u_{it} &= \phi u_{it-1} + \eta_{it} \end{aligned}$$

as

$$\begin{aligned} Y_{it} &= \mathbf{X}_{it}\boldsymbol{\beta}_{AR} + \phi u_{it-1} + \eta_{it} \\ &= \mathbf{X}_{it}\boldsymbol{\beta}_{AR} + \phi(Y_{it-1} - \mathbf{X}_{it-1}\boldsymbol{\beta}_{AR}) + \eta_{it} \\ &= \phi Y_{it-1} + \mathbf{X}_{it}\boldsymbol{\beta}_{AR} + \mathbf{X}_{it-1}\psi + \eta_{it} \end{aligned}$$

where $\psi = \phi\beta_{AR}$ and $\psi = 0$ (by assumption).

Lagged Ys and World Domination

In:

$$Y_{it} = \phi Y_{it-1} + \mathbf{X}_{it}\boldsymbol{\beta}_{LDV} + \epsilon_{it}$$

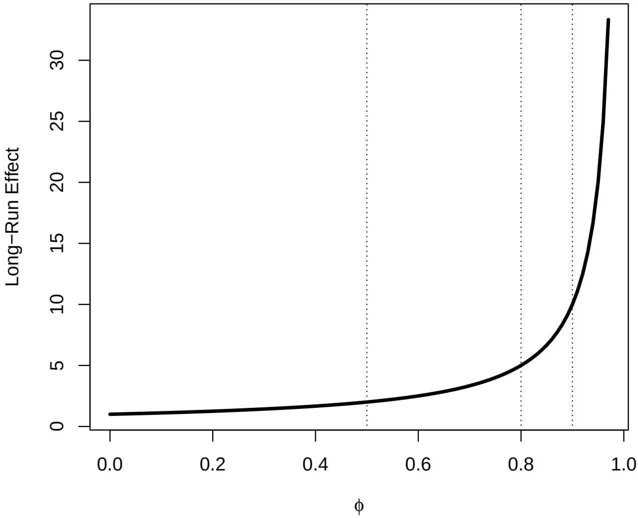
Achen: Bias "deflates" $\hat{\boldsymbol{\beta}}_{LDV}$ relative to $\hat{\phi}$, "suppress" the effects of **X**...

Keele & Kelly (2006):

- Contingent on ϵs having autocorrelation
- Key: In LDV, *long-run impact of a unit change in X is*:

$$\hat{\beta}_{LR} = \frac{\hat{\beta}_{LDV}}{1 - \hat{\phi}}$$

Long-Run Impact for $\hat{\beta} = 1$



Lagged Ys and Unit Effects

Consider:

$$Y_{it} = \phi Y_{it-1} + \mathbf{X}_{it}\boldsymbol{\beta} + \alpha_i + u_{it}.$$

If we omit the unit effects, we have:

$$Y_{it} = \phi Y_{it-1} + \mathbf{X}_{it}\boldsymbol{\beta} + u^*_{it}$$

with

$$u^*_{it} = \alpha_i + u_{it}$$

Lagging yields:

$$Y_{it-1} = \phi Y_{it-2} + \mathbf{X}_{it-1}\boldsymbol{\beta} + \alpha_i + u_{it-1}$$

which means

$$\text{Cov}(Y_{it-1}, u^*_{it}) \neq 0. \rightarrow \text{ bias in } \hat{\phi}, \hat{\boldsymbol{\beta}}$$

“Nickell” Bias

Omitting fixed effects in a model with Y_{it-1} yields bias in $\hat{\phi}$ that is:

- toward zero when $\phi > 0$,
- increasing in ϕ.

Including unit effects still yields bias in $\hat{\phi}$ of $O(\frac{1}{T})$, and bias in $\hat{\boldsymbol{\beta}}$.

Solutions:

- Difference/GMM estimation
- Bias correction approaches

First Difference Estimation

$$
\begin{aligned}
Y_{it} - Y_{it-1} &= \phi(Y_{it-1} - Y_{it-2}) + (\mathbf{X}_{it} - \mathbf{X}_{it-1})\boldsymbol{\beta} + (\alpha_i - \alpha_i) + (u_{it} - u_{it-1}) \\
\Delta Y_{it} &= \phi \Delta Y_{it-1} + \Delta \mathbf{X}_{it}\boldsymbol{\beta} + \Delta u_{it}
\end{aligned}
$$

Anderson/Hsiao: If $\nexists$ autocorrelation, then use ΔY_{it-2} or Y_{it-2} as instruments for ΔY_{it-1}...

- Consistent in theory,
- in practice, the former is preferred, and both have issues if ϕ is high;
- both are inefficient.

A&H $\rightarrow$ A&B

Arellano & Bond (also Wawro): Use *all* lags of Y_{it} and $\mathbf{X}_{it}$ from $t-2$ and before.

- "Good" estimates, better as $T \rightarrow \infty$,
- Easy to handle higher-order lags of Y,
- Easy software (`plm` in R , `xtabond` in Stata).
- Model *is* fixed effects...
- $\mathbf{Z}_i$ has $T-p-1$ rows, $\sum_{i=p}^{T-2} i$ columns $\rightarrow$ difficulty of estimation declines in p, <u>grows in T</u>.

Bias-Correction Models

Kiviet (1995, 1999; Bun and Kiviet 2003; Bruno 2005a,b): Derive the bias in $\hat{\phi}$ and $\hat{\beta}$, then correct it...

- More accurate than the instrumental-variables/GMM estimators of A&H/A&B...
- ...especially when T is small; but not as T gets reasonably large ($T \approx 20$)

Some Dynamic Models

	Lagged Y	First Difference	FE	Lagged Y + FE
Intercept	2.170*	0.637*		
	(0.314)	(0.038)		
Lagged WBLI	0.987*			0.954*
	(0.002)			(0.003)
Population Growth	−0.044*	−0.003	−0.149	−0.048
	(0.021)	(0.028)	(0.084)	(0.025)
Urban Population	0.003	−0.027	0.309*	0.013*
	(0.002)	(0.060)	(0.020)	(0.006)
Fertility Rate	−0.086*	−0.616	−1.728*	−0.250*
	(0.028)	(0.327)	(0.155)	(0.047)
ln(GDP Per Capita)	−0.041	0.690	8.940*	0.174
	(0.034)	(0.427)	(0.298)	(0.094)
Natural Resource Rents	−0.009*	0.025*	0.072*	−0.001
	(0.003)	(0.007)	(0.017)	(0.005)
Cold War	−0.263*	−0.007	−6.961*	−0.360*
	(0.067)	(0.197)	(0.291)	(0.089)
R^2	0.985	0.003	0.528	0.958
Adj. R^2	0.985	0.002	0.516	0.957
Num. obs.	8148	8070	8257	8148

$^*p < 0.05$

Trends!

What if Y is *trending* over time?

- First Question: Why?
  - Organic growth (e.g., populations)
  - Temporary / short-term factors
  - Covariates...
- Second question: Should we care?
  (A: Yes, usually... $\rightarrow$ "spurious regressions")
- Third question: What to do?
  - Ignore it...
  - Include a counter / trend term...

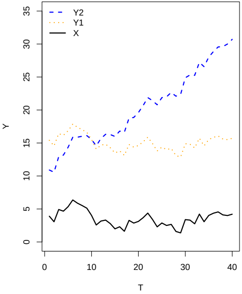
In general, adding a trend term will *decrease* the magnitudes of $\hat{\beta}$...

Trends Matter, Illustrated

Data generating processes:

$$Y_{1t} = 10 + (1 \times X_t) + u_t$$

$$Y_{2t} = 5 + (1 \times X_t) + (0.5 \times T) + u_t$$



	Y_1	Y_2	
		No Trend	Trend
X	0.921***	−0.382	0.874***
	(0.245)	(0.786)	(0.255)
T			0.482***
			(0.026)
Constant	10.300***	20.200***	5.860***
	(0.917)	(2.950)	(1.200)
Observations	40	40	40
R^2	0.272	0.006	0.905
Adjusted R^2	0.253	−0.020	0.900
Residual Std. Error	1.800 (df = 38)	5.790 (df = 38)	1.810 (df = 37)
Note:			*p<0.1; **p<0.05; ***p<0.01

Trends Matter, Part II

Table: FE Models of WBLI

	FE	FE.trend	FE.intx
Population Growth	-0.149^{*}	-0.270^{***}	-0.254^{***}
	(0.084)	(0.070)	(0.070)
Urban Population	0.309^{***}	0.015	0.031^{*}
	(0.020)	(0.017)	(0.017)
Fertility Rate	-1.728^{***}	1.393^{***}	1.350^{***}
	(0.155)	(0.140)	(0.140)
ln(GDP Per Capita)	8.940^{***}	2.384^{***}	2.161^{***}
	(0.298)	(0.273)	(0.274)
Natural Resource Rents	0.072^{***}	0.044^{***}	0.045^{***}
	(0.017)	(0.014)	(0.014)
Cold War	-6.961^{***}	1.725^{***}	7.858^{***}
	(0.291)	(0.284)	(0.914)
Trend (1950=0)		0.737^{***}	0.761^{***}
		(0.012)	(0.013)
Cold War x Trend			-0.179^{***}
			(0.025)
Observations	8,257	8,257	8,257
R^2	0.528	0.671	0.673
Adjusted R^2	0.516	0.663	0.665
F Statistic	$1{,}502.000^{***}$ (df = 6; 8064)	$2{,}344.000^{***}$ (df = 7; 8063)	$2{,}070.000^{***}$ (df = 8; 8062)

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Another Approach: Orthogonalization

Note: We're rarely substantively interested in the fixed effects $\hat{\alpha}$...

- $\rightarrow$ reparameterize the αs so that they are *information-orthogonal* to the other parameters in the model (including the βs and ϕ)
- Key idea: Transform the αs so that (for example):

$$\mathsf{E}\left(\frac{\partial^2 L_i}{\partial\alpha\partial\beta}\right) = 0$$

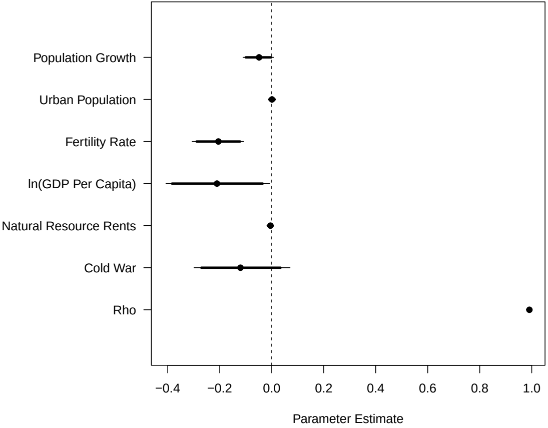
- Can do this via imposition of priors, in a Bayesian framework...
- **In general**, this approach is less assumption-laden and more efficient than the IV/GMM-based approaches discussed above.
- Provides consistent-in-N estimates for T as low as 2...

References:

- Lancaster, T. 2002. "Orthogonal Parameters and Panel Data." *Review of Economic Studies* 69:647-666.
- Pickup et al. (2017) [the "orthogonalized panel model" ("OPM")]

FE + Dynamics Using Orthogonalization

```
> library(OrthoPanels)
> set.seed(7222009)
> OPM.fit <- opm(WomenBusLawIndex~PopGrowth+UrbanPopulation+FertilityRate+
               lnGDPPerCap+NaturalResourceRents+ColdWar,data=WDI,
               index=c("ISO3","Year"),n.samp=1000)
```



OPM Results: Short- and Long-Run Effects

For $\hat{\phi} \approx 0.99$:

Parameter	Short-Run	Long-Run
Population Growth	-0.0486	-5.2938
Urban Population	0.0010	0.1076
Fertility Rate	-0.2051	-23.0112
ln(GDP Per Capita)	-0.2106	-23.4552
Natural Resource Rents	-0.0049	-0.5348
Cold War	-0.1202	-13.2339

Dynamic Models: Software

R :

- the `plm` package (`purtest` for unit roots; `plm` for first-difference models)
- the `panelAR` package (GLS-ARMA models)
- the `gls` package (GLS)
- the `pgmm` package (A&B)
- the `dynpanel` package (A&H, A&B)

Stata :

- `xtgls` (GLS)
- `xtpcse` (PCSEs)
- `xtabond` / `xtdpd` (A&H A&B dynamic models)

Final Thoughts: Dynamic Panel Models

- N vs. T...
- Are dynamics nuisance or substance?
- What problem(s) do you *really* care about?