GSERM - Ljubljana 2024
Analyzing Panel Data

January 16, 2024

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:

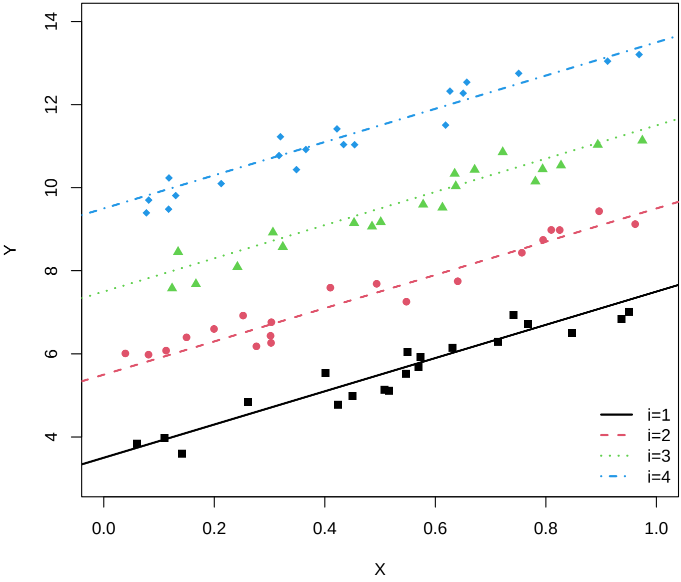
- Assuming $\alpha_i = 0$ (w.l.o.g):

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

- 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>

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

“Fixed” Effects

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, 1960-2022

The World Development Indicators

- Cross-national country-level time series data
- $N = 215$ countries, $T = 73$ years (1960-2022) + 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

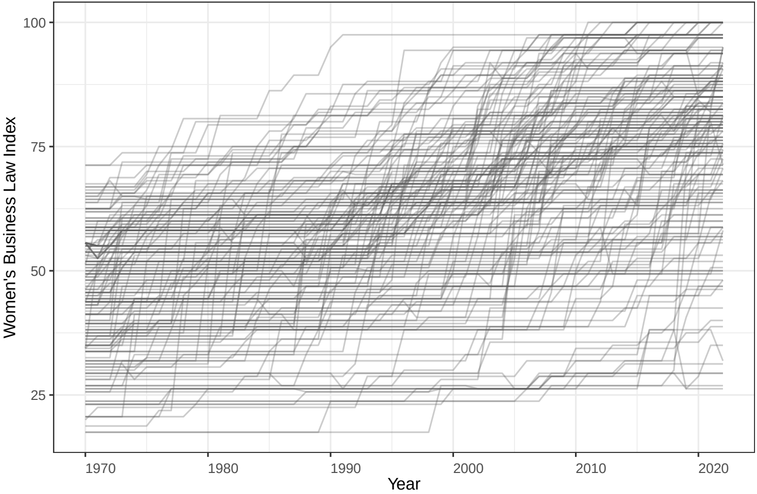
```
> describe(wdi,fast=TRUE,ranges=FALSE,check=TRUE)
                        vars     n            mean              sd             se
ISO3                       1 13545             NaN              NA             NA
Year                       2 13545         1991.00           18.18           0.16
Region                     3 13545             NaN              NA             NA
country                    4 13545             NaN              NA             NA
iso3c                      5 13545             NaN              NA             NA
LandArea                   6 12728       611322.13      1764229.22       15637.77
ArablePercent              7 11375           13.44           13.53           0.13
Population                 8 13300     24919941.35    104042745.30      902165.02
PopGrowth                  9 13084            1.77            1.78           0.02
RuralPopulation           10 13268           48.45           25.74           0.22
UrbanPopulation           11 13268           51.55           25.74           0.22
BirthRatePer1K            12 12937           28.02           13.08           0.12
FertilityRate             13 12779            3.91            2.00           0.02
PrimarySchoolAge          14 10699            6.14            0.62           0.01
LifeExpectancy            15 12766           64.63           11.29           0.10
AgeDepRatioOld            16 13300           10.62            6.93           0.06
CO2Emissions              17  5729            4.29            5.49           0.07
GDP                       18  9843 245055369928.40 1121079127717.87 11299845990.78
GDPPerCapita              19  9843        11874.12        18895.82         190.46
GDPPerCapGrowth           20  9818            1.93            6.17           0.06
Inflation                 21  8502           23.53          327.96           3.56
TotalTrade                22  8548           78.32           54.23           0.59
Exports                   23  8548           36.49           28.94           0.31
Imports                   24  8557           41.84           27.87           0.30
FDIIn                     25  8406            5.50           45.06           0.49
AgriEmployment            26  5764           28.83           23.85           0.31
NetAidReceived            27  8907    473766874.40    900415366.59     9540632.60
MobileCellSubscriptions   28 10057           35.06           51.01           0.51
NaturalResourceRents      29  9211            6.84           11.06           0.12
MilitaryExpenditures      30  7393            2.75            3.21           0.04
GovtExpenditures          31  8197           16.26            8.17           0.09
HIVDeaths                 32  4370         8137.20        24927.88         377.09
WomenBusLawIndex          33  9964           59.61           18.62           0.19
PaidParentalLeave         34  9964            0.11            0.31           0.00
ColdWar                   35 13545            0.48            0.50           0.00
```

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

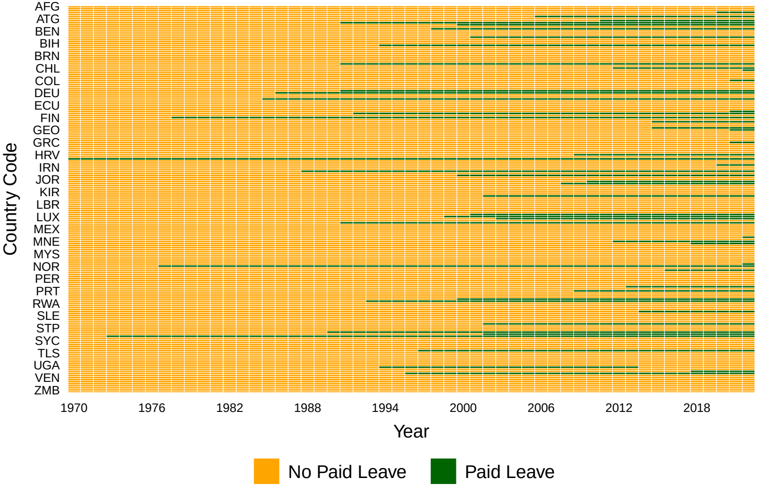
The basis for a 2021 World Bank report...

- Examines "the laws and regulations that affect women's economic opportunity in 190 economies" from 1970-2022.
- 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 $\approx$ 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."

Visualization (using `panelView`)

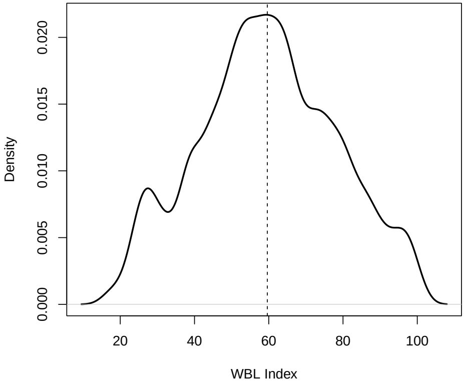


Categorical Variable Visualization



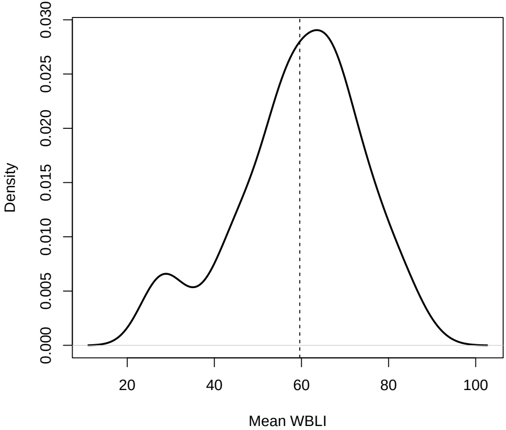
WBLI: Total Variation

```
> 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 9964 59.6 18.6   58.8    59.6 19.5 17.5 100  82.5 0.02    -0.57 0.19
```



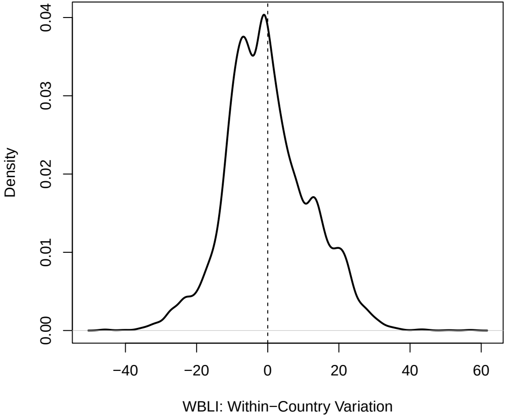
WBLI: "Between" Variation

```
> 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 59.6 14.5   60.8    60.5 13.2 23.5 90.2  66.7 -0.49    -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 9964    0 11.8  -1.08   -0.32 10.7 -45.7 56.9   103 0.25     0.39 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	min	max	range	se
WomenBusLawIndex	1	8100	60.69	18.95	17.50	100.00	82.50	0.21
PopGrowth	2	8100	1.65	1.54	-16.88	19.36	36.24	0.02
UrbanPopulation	3	8100	51.56	23.82	2.85	100.00	97.16	0.26
FertilityRate	4	8100	3.61	1.90	0.77	8.61	7.83	0.02
NaturalResourceRents	5	8100	7.04	10.77	0.00	88.59	88.59	0.12
ColdWar	6	8100	0.30	0.46	0.00	1.00	1.00	0.01
lnGDPPerCap	7	8100	8.29	1.44	5.04	11.64	6.60	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 = 8100

Residuals:
   Min. 1st Qu.  Median 3rd Qu.    Max.
 -50.32   -8.50    1.05    9.20   43.83

Coefficients:
                     Estimate Std. Error t-value    Pr(>|t|)
(Intercept)           60.4325     1.6861    35.8     < 2e-16 ***
PopGrowth             -2.3630     0.1306   -18.1     < 2e-16 ***
UrbanPopulation       -0.0587     0.0105    -5.6 0.000000022 ***
FertilityRate         -2.5215     0.1592   -15.8     < 2e-16 ***
log(GDPPerCapita)      2.6533     0.1936    13.7     < 2e-16 ***
NaturalResourceRents  -0.3398     0.0155   -21.9     < 2e-16 ***
ColdWar              -10.9584     0.3715   -29.5     < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    2910000
Residual Sum of Squares: 1450000
R-Squared:      0.501
Adj. R-Squared: 0.501
F-statistic: 1354.19 on 6 and 8093 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 = 8100

Residuals:
   Min. 1st Qu.  Median 3rd Qu.    Max.
-33.588  -5.063  -0.437   4.833  52.886

Coefficients:
                     Estimate Std. Error t-value Pr(>|t|)
PopGrowth             -0.0891     0.0955   -0.93  0.35085
UrbanPopulation        0.3062     0.0198   15.46  < 2e-16 ***
FertilityRate         -2.0328     0.1620  -12.55  < 2e-16 ***
log(GDPPerCapita)      8.7230     0.2998   29.10  < 2e-16 ***
NaturalResourceRents   0.0647     0.0172    3.76  0.00017 ***
ColdWar               -6.8691     0.2959  -23.22  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1050000
Residual Sum of Squares: 492000
R-Squared:      0.532
Adj. R-Squared: 0.521
F-statistic: 1500.96 on 6 and 7907 DF, p-value: <2e-16
```

A Nicer Table

Table: Models of WBLI

	OLS	FE
Population Growth	-2.360^{***}	-0.089
	(0.131)	(0.095)
Urban Population	-0.059^{***}	0.306^{***}
	(0.010)	(0.020)
Fertility Rate	-2.520^{***}	-2.030^{***}
	(0.159)	(0.162)
ln(GDP Per Capita)	2.650^{***}	8.720^{***}
	(0.194)	(0.300)
Natural Resource Rents	-0.340^{***}	0.065^{***}
	(0.015)	(0.017)
Cold War	-11.000^{***}	-6.870^{***}
	(0.372)	(0.296)
Constant	60.400^{***}	
	(1.690)	
Observations	8,100	8,100
R^2	0.501	0.532
Adjusted R^2	0.501	0.521
F Statistic	$1{,}354.000^{***}$ (df = 6; 8093)	$1{,}501.000^{***}$ (df = 6; 7907)

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

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	-0.089	-2.600^{***}
	(0.095)	(0.123)
Urban Population	0.306^{***}	-0.060^{***}
	(0.020)	(0.010)
Fertility Rate	-2.030^{***}	-1.560^{***}
	(0.162)	(0.154)
ln(GDP Per Capita)	8.720^{***}	3.020^{***}
	(0.300)	(0.183)
Natural Resource Rents	0.065^{***}	-0.378^{***}
	(0.017)	(0.015)
Cold War	-6.870^{***}	
	(0.296)	
Observations	8,100	8,100
R^2	0.532	0.403
Adjusted R^2	0.521	0.399
F Statistic	$1,501.000^{***}$ (df = 6; 7907)	$1,088.000^{***}$ (df = 5; 8043)

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

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 = 83, df1 = 186, df2 = 7907, 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 = 51501, 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 = 227, p-value <2e-16
alternative hypothesis: significant effects
```

Same For Time Effects

```
> pFtest(FE.Time,OLS)

F test for time effects

data:  WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +  ...
F = 23, df1 = 50, df2 = 8043, p-value <2e-16
alternative hypothesis: significant effects

> plmtest(FE.Time,effect=c("time"),type=c("bp"))

Lagrange Multiplier Test - time effects (Breusch-Pagan)

data:  WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +  ...
chisq = 9131, df = 1, p-value <2e-16
alternative hypothesis: significant effects

> plmtest(FE.Time,effect=c("time"),type=c("kw"))

Lagrange Multiplier Test - time effects (King and Wu)

data:  WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +  ...
normal = 96, 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 **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] 25.7
```

`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] 8.96
```

> *"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

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 = 8100
Observations used in estimation: 187

Residuals:
   Min. 1st Qu.  Median 3rd Qu.    Max.
-30.892  -6.245   0.909   8.222  22.558

Coefficients:
                     Estimate Std. Error t-value    Pr(>|t|)
(Intercept)           57.5438    10.4874    5.49 0.000000137 ***
PopGrowth             -5.8141     1.0423   -5.58 0.000000088 ***
UrbanPopulation       -0.0420     0.0553   -0.76     0.44838
FertilityRate         -0.3862     1.1364   -0.34     0.73439
log(GDPPerCapita)      2.6886     1.1516    2.33     0.02066 *
NaturalResourceRents  -0.3323     0.0913   -3.64     0.00036 ***
ColdWar              -11.9779     5.1223   -2.34     0.02047 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    45000
Residual Sum of Squares: 18700
R-Squared:      0.584
Adj. R-Squared: 0.57
F-statistic: 42.1563 on 6 and 180 DF, p-value: <2e-16
```

A Nicer Table (Again)

Table: Models of WBLI

	OLS	FE	BE
Population Growth	-2.360^{***}	-0.089	-5.810^{***}
	(0.131)	(0.095)	(1.040)
Urban Population	-0.059^{***}	0.306^{***}	-0.042
	(0.010)	(0.020)	(0.055)
Fertility Rate	-2.520^{***}	-2.030^{***}	-0.386
	(0.159)	(0.162)	(1.140)
ln(GDP Per Capita)	2.650^{***}	8.720^{***}	2.690^{**}
	(0.194)	(0.300)	(1.150)
Natural Resource Rents	-0.340^{***}	0.065^{***}	-0.332^{***}
	(0.015)	(0.017)	(0.091)
Cold War	-11.000^{***}	-6.870^{***}	-12.000^{**}
	(0.372)	(0.296)	(5.120)
Constant	60.400^{***}		57.500^{***}
	(1.690)		(10.500)
Observations	8,100	8,100	187
R^2	0.501	0.532	0.584
Adjusted R^2	0.501	0.521	0.570
F Statistic	$1{,}354.000^{***}$ (df = 6; 8093)	$1{,}501.000^{***}$ (df = 6; 7907)	42.200^{***} (df = 6; 180)

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

“Random” Effects

Model:

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

with:

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

and

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

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:

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

"Random" Effects: Estimation

We can then show that:

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

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 = 8100

Effects:
                var std.dev share
idiosyncratic 62.19    7.89   0.4
individual    94.85    9.74   0.6
theta:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  0.371   0.876   0.888   0.880   0.888   0.888

Residuals:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  -33.8    -5.4    -0.5     0.0     5.7    44.0

Coefficients:
                     Estimate Std. Error z-value Pr(>|z|)
(Intercept)            3.1144     2.5564    1.22    0.223
PopGrowth             -0.1835     0.0974   -1.88    0.060 .
UrbanPopulation        0.1881     0.0186   10.09   <2e-16 ***
FertilityRate         -2.3476     0.1618  -14.51   <2e-16 ***
log(GDPPerCapita)      7.1058     0.2839   25.03   <2e-16 ***
NaturalResourceRents   0.0342     0.0172    1.99    0.047 *
ColdWar               -8.1682     0.2908  -28.08   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    1100000
Residual Sum of Squares: 531000
R-Squared:      0.516
Adj. R-Squared: 0.516
Chisq: 8389.06 on 6 DF, p-value: <2e-16
```

A Nicer Table (Yet Again)

Table: Models of WBLI

	OLS	FE	BE	RE
Population Growth	-2.360^{***}	-0.089	-5.810^{***}	-0.184^{*}
	(0.131)	(0.095)	(1.040)	(0.097)
Urban Population	-0.059^{***}	0.306^{***}	-0.042	0.188^{***}
	(0.010)	(0.020)	(0.055)	(0.019)
Fertility Rate	-2.520^{***}	-2.030^{***}	-0.386	-2.350^{***}
	(0.159)	(0.162)	(1.140)	(0.162)
ln(GDP Per Capita)	2.650^{***}	8.720^{***}	2.690^{**}	7.110^{***}
	(0.194)	(0.300)	(1.150)	(0.284)
Natural Resource Rents	-0.340^{***}	0.065^{***}	-0.332^{***}	0.034^{**}
	(0.015)	(0.017)	(0.091)	(0.017)
Cold War	-11.000^{***}	-6.870^{***}	-12.000^{**}	-8.170^{***}
	(0.372)	(0.296)	(5.120)	(0.291)
Constant	60.400^{***}		57.500^{***}	3.110
	(1.690)		(10.500)	(2.560)
Observations	8,100	8,100	187	8,100
R^2	0.501	0.532	0.584	0.516
Adjusted R^2	0.501	0.521	0.570	0.516
F Statistic	$1{,}354.000^{***}$ (df = 6; 8093)	$1{,}501.000^{***}$ (df = 6; 7907)	42.200^{***} (df = 6; 180)	$8{,}389.000^{***}$

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

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)  # ugh...

Hausman Test

data:  WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +  ...
chisq = 10835, 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:

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

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 us
- Homoscedasticity
- Residual Independence:
  - $\text{Cov}(\varepsilon_{\cdot j}, u_{ij}) = 0$
  - $\text{Cov}(u_{ij}, u_{i\ell}) = 0$

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`)

```
> 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: 57477

Scaled residuals:
   Min     1Q Median     3Q    Max
-4.262 -0.644 -0.066  0.624  6.595

Random effects:
 Groups   Name        Variance Std.Dev.
 ISO3     (Intercept) 347.3    18.64
 Residual              62.3     7.89
Number of obs: 8100, groups:  ISO3, 187

Fixed effects:
                     Estimate Std. Error t value
(Intercept)          -10.9001     2.9106   -3.74
PopGrowth             -0.1166     0.0954   -1.22
UrbanPopulation        0.2655     0.0193   13.76
FertilityRate         -2.1365     0.1608  -13.29
log(GDPPerCapita)      8.1586     0.2925   27.90
NaturalResourceRents   0.0558     0.0171    3.27
ColdWar               -7.3293     0.2918  -25.12

Correlation of Fixed Effects:
            (Intr) PpGrwt UrbnPp FrtltR l(GDPP NtrlRR
PopGrowth    0.041
UrbanPopltn -0.199  0.004
FertilityRt -0.402 -0.272  0.433
lg(GDPPrCp) -0.757 -0.049 -0.282  0.105
NtrlRsrcRnt -0.018 -0.080 -0.021 -0.099  0.005
ColdWar     -0.114  0.043  0.214 -0.432  0.112  0.062
```

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

Table: RE and HLM Models of WBLI

	RE	AltRE
Population Growth	-0.184^{*}	-0.117
	(0.097)	(0.095)
Urban Population	0.188^{***}	0.266^{***}
	(0.019)	(0.019)
Fertility Rate	-2.350^{***}	-2.140^{***}
	(0.162)	(0.161)
ln(GDP Per Capita)	7.110^{***}	8.160^{***}
	(0.284)	(0.292)
Natural Resource Rents	0.034^{**}	0.056^{***}
	(0.017)	(0.017)
Cold War	-8.170^{***}	-7.330^{***}
	(0.291)	(0.292)
Constant	3.110	-10.900^{***}
	(2.560)	(2.910)
Observations	8,100	8,100
R^2	0.516	
Adjusted R^2	0.516	
Log Likelihood		-28,739.000
Akaike Inf. Crit.		57,495.000
Bayesian Inf. Crit.		57,558.000
F Statistic	$8,389.000^{***}$	

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

For more, 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: 54105

Scaled residuals:
   Min     1Q Median     3Q    Max
-4.454 -0.533  0.006  0.536  7.995

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 ISO3     (Intercept) 583.7    24.16
          ColdWar     141.0    11.88    -0.18
 Residual              37.6     6.13
Number of obs: 8100, groups:  ISO3, 187

Fixed effects:
                      Estimate Std. Error t value
(Intercept)          -28.03144    3.32446   -8.43
PopGrowth             -0.25074    0.07768   -3.23
UrbanPopulation        0.33793    0.02247   15.04
FertilityRate         -4.07247    0.17329  -23.50
log(GDPPerCapita)     10.51642    0.32562   32.30
NaturalResourceRents   0.00695    0.01490    0.47
ColdWar               -2.41568    1.02418   -2.36

Correlation of Fixed Effects:
            (Intr) PpGrwt UrbnPp FrtltR l(GDPP NtrlRR
PopGrowth    0.051
UrbanPopltn -0.147 -0.015
FertilityRt -0.485 -0.194  0.480
lg(GDPPrCp) -0.712 -0.056 -0.369  0.184
NtrlRsrcRnt  0.011 -0.059  0.058 -0.040 -0.068
ColdWar     -0.083 -0.001  0.052 -0.114 -0.001  0.010
```

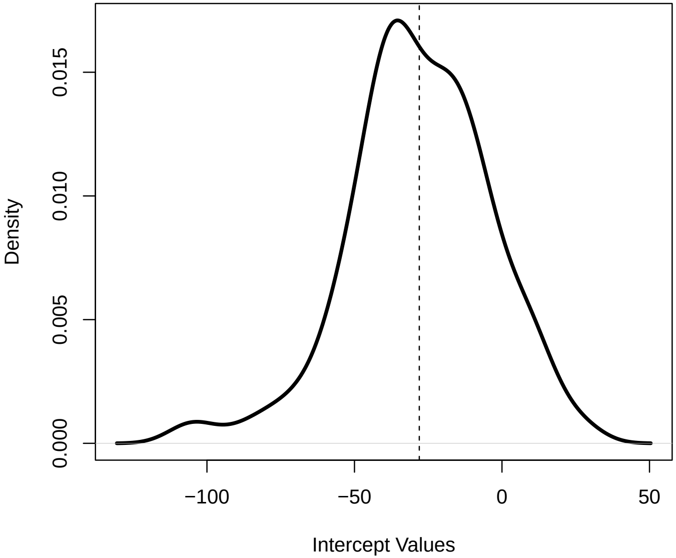
```
> anova(AltRE,HLM1)

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 57478 57541 -28730    57460
HLM1    11 54113 54190 -27046    54091  3369  2     <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> VarCorr(HLM1)

 Groups   Name        Std.Dev. Corr
 ISO3     (Intercept) 24.16
          ColdWar     11.88    -0.18
 Residual              6.13
```

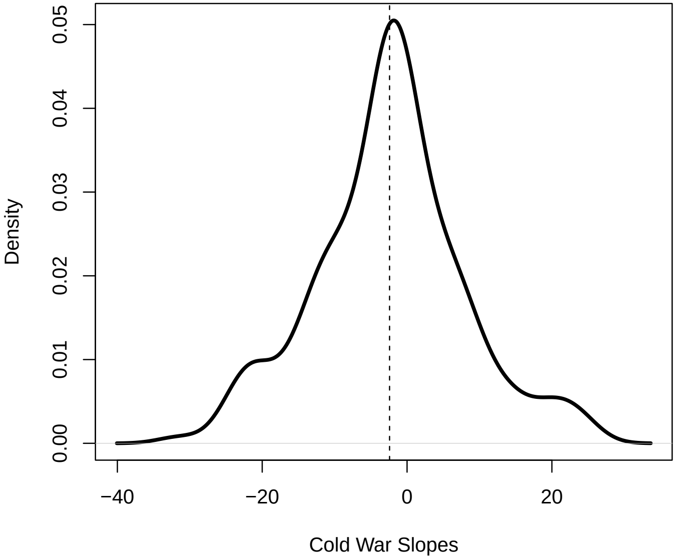
Random Coefficients

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

> head(Bs)
    ISO3..Intercept. ISO3.PopGrowth ISO3.UrbanPopulation ISO3.FertilityRate
AFG           -17.42         -0.251                0.338              -4.07
AGO            -9.73         -0.251                0.338              -4.07
ALB           -11.40         -0.251                0.338              -4.07
ARE          -101.67         -0.251                0.338              -4.07
ARG           -49.50         -0.251                0.338              -4.07
ARM           -23.39         -0.251                0.338              -4.07
    ISO3.log.GDPPerCapita. ISO3.NaturalResourceRents ISO3.ColdWar
AFG                   10.5                   0.00695        -3.36
AGO                   10.5                   0.00695       -13.82
ALB                   10.5                   0.00695        -6.53
ARE                   10.5                   0.00695        -2.60
ARG                   10.5                   0.00695       -22.73
ARM                   10.5                   0.00695        -2.83

> mean(Bs$ISO3..Intercept.)
[1] -28

> mean(Bs$ISO3.ColdWar)
[1] -2.42
```

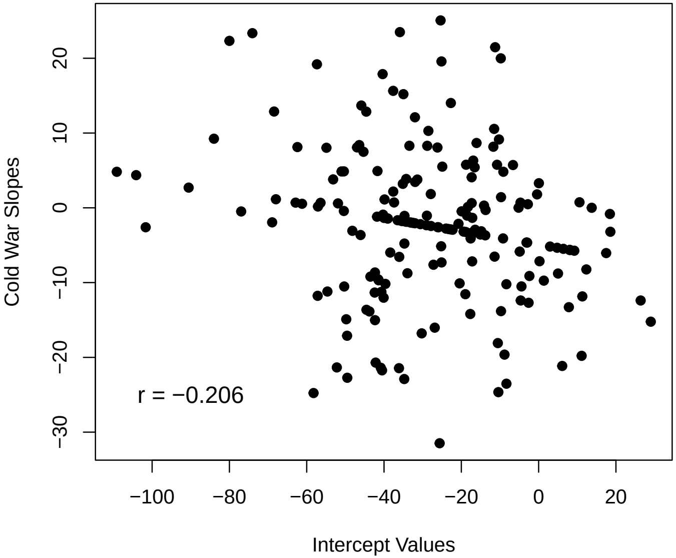
Random Intercepts (Plotted)



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:

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

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	-2.150^{***}
	(0.128)
Urban Population	-0.044^{***}
	(0.010)
Fertility Rate	-2.270^{***}
	(0.157)
ln(GDP Per Capita)	2.620^{***}
	(0.190)
Within-Country Nat. Resource Rents	0.094^{***}
	(0.028)
Between-Country Nat. Resource Rents	-0.507^{***}
	(0.018)
Cold War	-11.500^{***}
	(0.365)
Constant	60.000^{***}
	(1.650)
Observations	8,100
R^2	0.521
Adjusted R^2	0.520
Residual Std. Error	13.100 (df = 8092)
F Statistic	$1{,}256.000^{***}$ (df = 7; 8092)

*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 $\text{Cov}(\mathbf{X}_{it}, \eta_t) = \text{Cov}(\alpha_i, \eta_t) = 0$
- 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 = 8100

Residuals:
   Min. 1st Qu.  Median 3rd Qu.    Max.
-32.011  -4.038   0.268   4.151  43.093

Coefficients:
                     Estimate Std. Error t-value      Pr(>|t|)
PopGrowth             -0.2496     0.0799   -3.13        0.0018 **
UrbanPopulation        0.0363     0.0173    2.10        0.0359 *
FertilityRate          1.2291     0.1477    8.32       < 2e-16 ***
log(GDPPerCapita)      1.7450     0.2764    6.31 0.00000000029 ***
NaturalResourceRents   0.0258     0.0149    1.73        0.0838 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    343000
Residual Sum of Squares: 337000
R-Squared:      0.0155
Adj. R-Squared: -0.0148
F-statistic: 24.7792 on 5 and 7857 DF, p-value: <2e-16
```

Two-Way Effects: Testing

```
> # Two-way effects:

> pFtest(WomenBusLawIndex~PopGrowth+UrbanPopulation+FertilityRate+
+        log(GDPPerCapita)+NaturalResourceRents+ColdWar,data=WDI,
+        effect="twoway",model="within")

F test for twoways effects

data:  WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +  ...
F = 110, df1 = 236, df2 = 7857, p-value <2e-16
alternative hypothesis: significant effects

> plmtest(TwoWayFE,c("twoways"),type=("kw"))

Lagrange Multiplier Test - two-ways effects (King and Wu)

data:  WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +  ...
normal = 191, p-value <2e-16
alternative hypothesis: significant effects


> # One-way effects in the two-way model:

> plmtest(TwoWayFE,c("individual"),type=("kw"))

Lagrange Multiplier Test - (King and Wu)

data:  WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +  ...
normal = 227, p-value <2e-16
alternative hypothesis: significant effects

> plmtest(TwoWayFE,c("time"),type=("kw"))

Lagrange Multiplier Test - time effects (King and Wu)

data:  WomenBusLawIndex ~ PopGrowth + UrbanPopulation + FertilityRate +  ...
normal = 96, p-value <2e-16
alternative hypothesis: significant effects
```

Two-Way Fixed Effects via `lm`

```
> 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
-32.01  -4.04   0.27   4.15  43.09

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)          -14.7211     2.4598   -5.98  2.3e-09 ***
PopGrowth             -0.2496     0.0799   -3.13  0.00178 **
UrbanPopulation        0.0363     0.0173    2.10  0.03594 *
FertilityRate          1.2291     0.1477    8.32  < 2e-16 ***
log(GDPPerCapita)      1.7450     0.2764    6.31  2.9e-10 ***
NaturalResourceRents   0.0258     0.0149    1.73  0.08380 .
factor(ISO3)AGO       29.5316     1.9500   15.14  < 2e-16 ***
factor(ISO3)ALB       52.2820     1.9857   26.33  < 2e-16 ***
factor(ISO3)ARE       -3.2022     2.4748   -1.29  0.19572
.
.
.
factor(Year)1977       4.2415     0.8950    4.74  2.2e-06 ***
factor(Year)1978       4.9507     0.8955    5.53  3.3e-08 ***
 [ reached getOption("max.print") -- omitted 43 rows ]
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.55 on 7857 degrees of freedom
  (5445 observations deleted due to missingness)
Multiple R-squared:  0.884,Adjusted R-squared:  0.881
F-statistic:  248 on 242 and 7857 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 = 8100

Effects:
                 var std.dev share
idiosyncratic 42.925   6.552  0.31
individual    95.295   9.762  0.69
time           0.429   0.655  0.00
theta:
       Min. 1st Qu. Median  Mean 3rd Qu.  Max.
id    0.443   0.897  0.907 0.900   0.907 0.907
time  0.300   0.355  0.398 0.379   0.405 0.408
total 0.293   0.354  0.397 0.378   0.404 0.407

Residuals:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  -63.5    -6.8     2.1     0.2    10.9    32.6

Coefficients:
                      Estimate Std. Error z-value Pr(>|z|)
(Intercept)           19.33798    0.33990    56.9   <2e-16 ***
PopGrowth             -0.22014    0.01210   -18.2   <2e-16 ***
UrbanPopulation        0.13185    0.00239    55.1   <2e-16 ***
FertilityRate         -0.80004    0.02101   -38.1   <2e-16 ***
log(GDPPerCapita)      4.94340    0.03709   133.3   <2e-16 ***
NaturalResourceRents   0.02749    0.00220    12.5   <2e-16 ***
ColdWar              -11.94828    0.04882  -244.8   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    2910000
Residual Sum of Squares: 2020000
R-Squared:      0.322
Adj. R-Squared: 0.322
Chisq: 222859 on 6 DF, p-value: <2e-16
```

A Prettier Table

Table: Models of WBLI

	OLS	FE	BE	RE	TwoWayFE	TwoWayRE
Population Growth	-2.360^{***}	-0.089	-5.810^{***}	-0.184^{*}	-0.250^{***}	-0.220^{***}
	(0.131)	(0.095)	(1.040)	(0.097)	(0.080)	(0.012)
Urban Population	-0.059^{***}	0.306^{***}	-0.042	0.188^{***}	0.036^{**}	0.132^{***}
	(0.010)	(0.020)	(0.055)	(0.019)	(0.017)	(0.002)
Fertility Rate	-2.520^{***}	-2.030^{***}	-0.386	-2.350^{***}	1.230^{***}	-0.800^{***}
	(0.159)	(0.162)	(1.140)	(0.162)	(0.148)	(0.021)
ln(GDP Per Capita)	2.650^{***}	8.720^{***}	2.690^{**}	7.110^{***}	1.750^{***}	4.940^{***}
	(0.194)	(0.300)	(1.150)	(0.284)	(0.276)	(0.037)
Natural Resource Rents	-0.340^{***}	0.065^{***}	-0.332^{***}	0.034^{**}	0.026^{*}	0.027^{***}
	(0.015)	(0.017)	(0.091)	(0.017)	(0.015)	(0.002)
Cold War	-11.000^{***}	-6.870^{***}	-12.000^{**}	-8.170^{***}		-11.900^{***}
	(0.372)	(0.296)	(5.120)	(0.291)		(0.049)
Constant	60.400^{***}		57.500^{***}	3.110		19.300^{***}
	(1.690)		(10.500)	(2.560)		(0.340)
Observations	8,100	8,100	187	8,100	8,100	8,100
R^2	0.501	0.532	0.584	0.516	0.016	0.322
Adjusted R^2	0.501	0.521	0.570	0.516	-0.015	0.322

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

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 panelr package (various commands)
- 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