

GSERM 2021

Analyzing Panel Data

June 15, 2021

One- and Two-Way “Unit Effects”

Two-way variation:

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

→ “two-way” effects:

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

Also: “One-way” effects:

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

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

“Brute force” model:

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

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:

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

Means that:

$$Y_{it}^* = Y_{it} - \bar{Y}_i$$

$$\mathbf{X}_{it}^* = \mathbf{X}_{it} - \bar{\mathbf{X}}_i$$

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

→ 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 $\sim F_{N-1, NT-(N-1)}$.

Example Data: Demonstrations, 1945-2014

Data:

- Data are (a subset of) Banks (2019)
- $N = 180$ countries, $T = 70$ years [1945-2014]
- Variables:
 - Demonstrations: Number of social/political demonstrations in that country in that year
 - POLITY: The country's POLITY IV score that year (-10 = fully autocratic; 10 = fully democratic)
 - POLITY²: POLITY IV squared (expected curvilinear relationship)
 - GDP: The per capita GDP (PPP, in constant \$US) for that country / year
 - Monarch: Whether (=1) or not (=0) that country was a monarchy in that year
 - ColdWar: Indicator variable, coded 1 for the period 1945-1989, 0 otherwise

Regression model:

$$\ln(\text{Demonstrations} + 1)_{it} = \beta_0 + \beta_1 \text{POLITY}_{it} + \beta_2 \text{POLITY}_{it}^2 + \beta_3 \ln(\text{GDP})_{it} + \beta_4 \text{Monarch}_{it} + \beta_5 \text{Cold War}_{it} + u_{it}$$

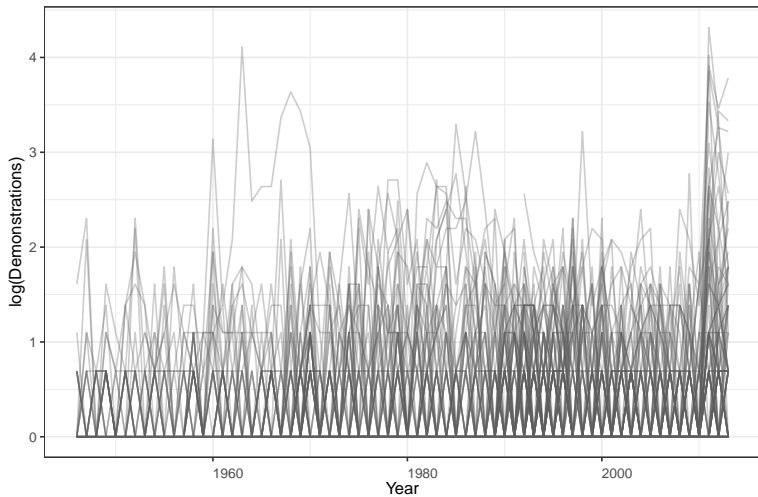
Data Summary

```
> summary(Demos)
```

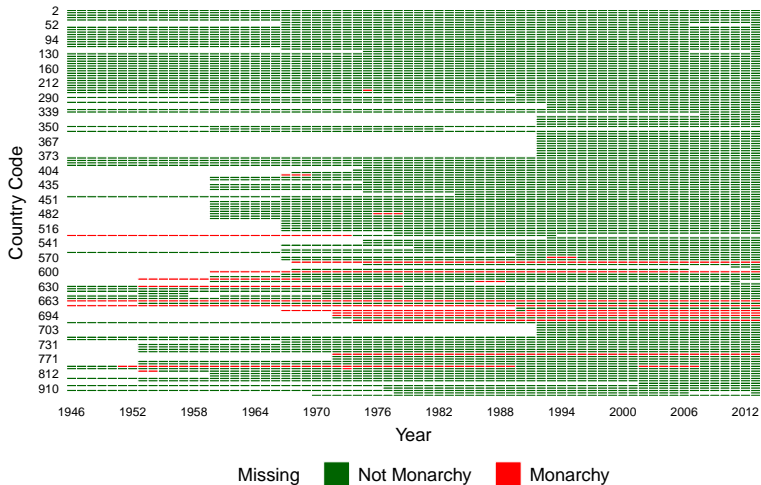
ccode	Year	POLITY	GDP
Min. : 2	Min. :1945	Min. : -10.00	Min. : 185
1st Qu.:235	1st Qu.:1969	1st Qu.: -7.00	1st Qu.: 1580
Median :451	Median :1985	Median : 0.00	Median : 4002
Mean :456	Mean :1984	Mean : 0.63	Mean : 8120
3rd Qu.:663	3rd Qu.:2000	3rd Qu.: 8.00	3rd Qu.: 10365
Max. :950	Max. :2014	Max. : 10.00	Max. :134040
		NA's :111	NA's :2348

Monarch	lnDemos	ColdWar	lnGDP
Min. :0.0	Min. :0.0	Min. :0.000	Min. : 5.2
1st Qu.:0.0	1st Qu.:0.0	1st Qu.:0.000	1st Qu.: 7.4
Median :0.0	Median :0.0	Median :1.000	Median : 8.3
Mean :0.1	Mean :0.3	Mean :0.563	Mean : 8.3
3rd Qu.:0.0	3rd Qu.:0.0	3rd Qu.:1.000	3rd Qu.: 9.2
Max. :1.0	Max. :4.3	Max. :1.000	Max. :11.8
NA's :1198	NA's :1149		NA's :2348

Visualization (using panelView)



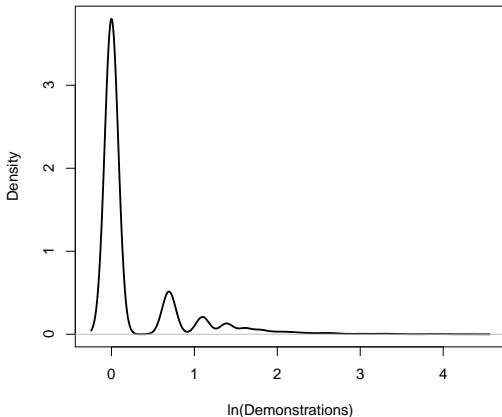
Categorical Variable Visualization



Demonstrations: Total Variation

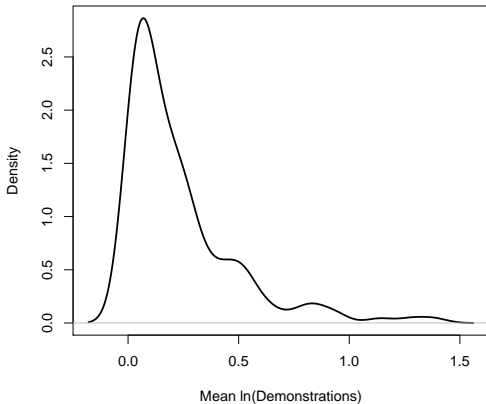
```
> with(Demos, describe(lnDemos)) # all variation
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
X1	1	8219	0.25	0.55	0	0.11	0	0	4.32	4.32	2.54	7.3	0.01



Demonstrations: “Between” Variation

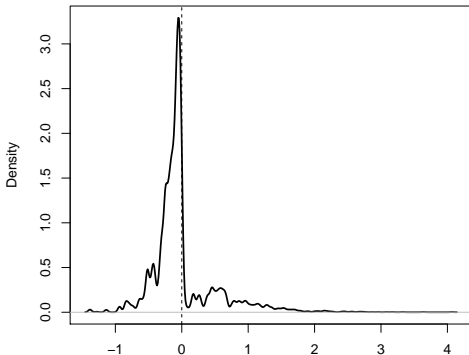
```
> DemonsMeans <- ddply(Demos,.(ccode),summarise,  
+                       DemonsMean = mean(lnDemos,na.rm=TRUE))  
>  
> with(DemonsMeans, describe(DemonsMean)) # "between" variation  
vars  n mean  sd median trimmed mad min  max range skew kurtosis  se  
X1    1 160 0.23 0.25  0.15   0.18 0.15  0 1.38  1.38 1.99    4.55 0.02
```



Demonstrations: “Within” Variation

```
> Demos <- ddpoly(Demos, .(ccode), mutate,  
+               DemonsMean = mean(lnDemos, na.rm=TRUE))  
> Demos$DemosWithin <- with(Demos, lnDemos-DemonsMean)  
>  
> with(Demos, describe(DemonsWithin))
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
X1	1	8219	0	0.48	-0.09	-0.07	0.16	-1.38	4.07	5.45	1.96	6.41	0.01



Demonstrations: Within-Country Variation

Regression: Pooled OLS

```
> OLS<-plm(lnDemons~POLITY+I(POLITY^2)+lnGDP+Monarch+ColdWar,
+          data=PDF,model="pooling")

> summary(OLS)
Pooling Model

Call:
plm(formula = lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch +
     ColdWar, data = PDF, model = "pooling")

Unbalanced Panel: n = 145, T = 1-62, N = 6505

Residuals:
    Min. 1st Qu.  Median 3rd Qu.    Max.
-0.4501 -0.2930 -0.2176 -0.0754  4.1073

Coefficients:
              Estimate Std. Error t-value Pr(>|t|)
(Intercept) -0.124639   0.058208  -2.14    0.032 *
POLITY       0.006296   0.001179   5.34 9.5e-08 ***
I(POLITY^2) -0.002267   0.000255  -8.90 < 2e-16 ***
lnGDP        0.057679   0.007513   7.68 1.9e-14 ***
Monarch      -0.046393   0.028572  -1.62  0.104
ColdWar      0.027883   0.013961   2.00  0.046 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Total Sum of Squares:    1850
Residual Sum of Squares: 1800
R-Squared:               0.0261
Adj. R-Squared: 0.0253
F-statistic: 34.8228 on 5 and 6499 DF, p-value: <2e-16
```

"Fixed" (Within) Effects

```
> FE<-plm(lnDemons~POLITY+I(POLITY^2)+lnGDP+Monarch+ColdWar,
+         data=PDF, effect="individual",model="within")

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

Call:
plm(formula = lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch +
     ColdWar, data = PDF, effect = "individual", model = "within")

Unbalanced Panel: n = 145, T = 1-62, N = 6505

Residuals:
    Min. 1st Qu.  Median 3rd Qu.    Max.
-1.3556 -0.2120 -0.0768  0.0193  4.0496

Coefficients:
              Estimate Std. Error t-value Pr(>|t|)
POLITY         0.001526   0.001553    0.98  0.32604
I(POLITY^2)  -0.001942   0.000296   -6.55  6.1e-11 ***
lnGDP          0.054586   0.015200    3.59  0.00033 ***
Monarch        0.047976   0.068071    0.70  0.48097
ColdWar       -0.035487   0.016235   -2.19  0.02887 *
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Total Sum of Squares:    1410
Residual Sum of Squares: 1400
R-Squared:              0.013
Adj. R-Squared:        -0.0102
F-statistic: 16.7177 on 5 and 6355 DF, p-value: <2e-16
```

A Nicer Table

Table: Models of Demonstrations

	OLS	FE
POLITY	0.006*** (0.001)	0.002 (0.002)
POLITY Squared	-0.002*** (0.0003)	-0.002*** (0.0003)
ln(GDP)	0.058*** (0.008)	0.055*** (0.015)
Monarch	-0.046 (0.029)	0.048 (0.068)
Cold War	0.028** (0.014)	-0.035** (0.016)
Constant	-0.125** (0.058)	
Observations	6,505	6,505
R ²	0.026	0.013
Adjusted R ²	0.025	-0.010
F Statistic	34.820*** (df = 5; 6499)	16.720*** (df = 5; 6355)

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

Time-Period Fixed Effects

The model is:

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

which is estimated via:

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

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

Comparison: Unit vs. Time Fixed Effects

Table: FE Models of Demonstrations

	FE.Units	FE.Time
POLITY	0.002 (0.002)	0.007*** (0.001)
POLITY Squared	-0.002*** (0.0003)	-0.002*** (0.0003)
ln(GDP)	0.055*** (0.015)	0.058*** (0.008)
Monarch	0.048 (0.068)	-0.038 (0.028)
Cold War	-0.035** (0.016)	
Observations	6,505	6,505
R ²	0.013	0.028
Adjusted R ²	-0.010	0.018
F Statistic	16.720*** (df = 5; 6355)	46.270*** (df = 4; 6439)

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

The specification:

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

...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

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

F test for individual effects

```
data: lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch + ColdWar  
F = 13, df1 = 144, df2 = 6355, p-value <2e-16  
alternative hypothesis: significant effects
```

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

Lagrange Multiplier Test - (Breusch-Pagan) for unbalanced panels

```
data: lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch + ColdWar  
chisq = 8016, df = 1, p-value <2e-16  
alternative hypothesis: significant effects
```

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

Lagrange Multiplier Test - (King and Wu) for unbalanced panels

```
data: lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch + ColdWar  
normal = 90, p-value <2e-16  
alternative hypothesis: significant effects
```

Same For Time Effects

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

F test for time effects

```
data:  lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch + ColdWar
F = 3, df1 = 60, df2 = 6439, p-value = 1e-13
alternative hypothesis: significant effects
```

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

Lagrange Multiplier Test - time effects (Breusch-Pagan) for unbalanced panels

```
data:  lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch + ColdWar
chisq = 144, 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) for unbalanced panels

```
data:  lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch + ColdWar
normal = 12, 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 counterfactuals when discussing the substantive impact of a treatment” (2018, 829).

Significance:

- Predictors **X** in FE models typically have both cross-sectional and temporal variation
- FE models only consider *within-unit* variation in **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: Monarchy

Monarchy – All Variation:

```
> with(Demos, sd(Monarch, na.rm=TRUE))  
[1] 0.2601
```

Monarchy – “Within” Variation:

```
> Demos <- ddply(Demos, .(ccode), mutate,  
+               MonMean = mean(Monarch, na.rm=TRUE))  
> Demos$MonarchWithin <- with(Demos, Monarch-MonMean)  
  
> with(Demos, sd(MonarchWithin, na.rm=TRUE)) # "within" variation  
[1] 0.09418
```

“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 β_B
- Slowly-Changing \mathbf{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:

$$Y_{it} = \bar{\mathbf{X}}_i \beta_B + \tilde{\mathbf{X}}_{it} \beta_W + \alpha_i + u_{it}$$

...we can derive a “Between Effects” model:

$$\bar{Y}_i = \bar{\mathbf{X}}_i \beta_B + u_{it}$$

This model:

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

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

“Between” Effects

```
> BE<-plm(lnDemons~POLITY+I(POLITY^2)+lnGDP+Monarch+ColdWar,
+         data=PDF, effect="individual",model="between")
>
> summary(BE)
Oneway (individual) effect Between Model

Call:
plm(formula = lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch +
     ColdWar, data = PDF, effect = "individual", model = "between")

Unbalanced Panel: n = 145, T = 1-62, N = 6505
Observations used in estimation: 145

Coefficients:
              Estimate Std. Error t-value Pr(>|t|)
(Intercept) -0.30601    0.20837   -1.47  0.1442
POLITY       0.00597    0.00489    1.22  0.2244
I(POLITY^2) -0.00302    0.00112   -2.69  0.0079 **
lnGDP       0.06883    0.02734    2.52  0.0130 *
Monarch     -0.04966    0.10320   -0.48  0.6312
ColdWar     0.25872    0.08482    3.05  0.0027 **
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Total Sum of Squares:    8.75
Residual Sum of Squares: 7.64
R-Squared:              0.127
Adj. R-Squared: 0.0961
F-statistic: 4.06164 on 5 and 139 DF, p-value: 0.0018
```

A Nicer Table (Again)

Table: Models of Demonstrations

	OLS	FE	BE
POLITY	0.006*** (0.001)	0.002 (0.002)	0.006 (0.005)
POLITY Squared	-0.002*** (0.0003)	-0.002*** (0.0003)	-0.003*** (0.001)
ln(GDP)	0.058*** (0.008)	0.055*** (0.015)	0.069** (0.027)
Monarch	-0.046 (0.029)	0.048 (0.068)	-0.050 (0.103)
Cold War	0.028** (0.014)	-0.035** (0.016)	0.259*** (0.085)
Constant	-0.125** (0.058)		-0.306 (0.208)
Observations	6,505	6,505	145
R ²	0.026	0.013	0.127
Adjusted R ²	0.025	-0.010	0.096
F Statistic	34.820*** (df = 5; 6499)	16.720*** (df = 5; 6355)	4.062*** (df = 5; 139)

*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

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

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

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

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:

$$\sigma_u^2 = \sigma_{\alpha}^2 + \sigma_{\eta}^2.$$

“Random” Effects: Estimation

The model in 1 will violate the standard OLS assumptions of uncorrelated errors, because the (compound) “errors” within each unit share a common component α_j .

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

$$\begin{aligned} E(\mathbf{u}_i \mathbf{u}_i') \equiv \mathbf{\Sigma}_i &= \sigma_\eta^2 \mathbf{I}_T + \sigma_\alpha^2 \mathbf{1}\mathbf{1}' \\ &= \begin{pmatrix} \sigma_\eta^2 + \sigma_\alpha^2 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\eta^2 + \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \cdots & \sigma_\eta^2 + \sigma_\alpha^2 \end{pmatrix} \end{aligned}$$

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:

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

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:

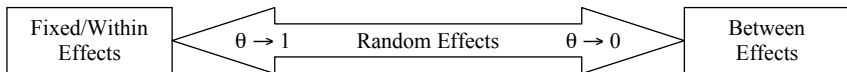
$$\begin{aligned} Y_{it}^* &= Y_{it} - \hat{\theta} \bar{Y}_i \\ X_{it}^* &= X_{it} - \hat{\theta} \bar{X}_i, \end{aligned}$$

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(lnDemons~POLITY+I(POLITY^2)+lnGDP+Monarch+ColdWar,
+         data=PDF, effect="individual", model="random")

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

Call:
plm(formula = lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch +
     ColdWar, data = PDF, effect = "individual", model = "random")

Unbalanced Panel: n = 145, T = 1-62, N = 6505

Effects:
               var std.dev share
idiosyncratic 0.2197 0.4687 0.8
individual    0.0563 0.2373 0.2
theta:
      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.108  0.708   0.736   0.724   0.757   0.757

Coefficients:
              Estimate Std. Error z-value Pr(>|z|)
(Intercept) -0.131708   0.104987  -1.25   0.210
POLITY       0.002574   0.001450   1.78   0.076 .
I(POLITY^2)  -0.001953   0.000287  -6.81 9.6e-12 ***
lnGDP        0.057117   0.012443   4.59 4.4e-06 ***
Monarch      -0.006937   0.053291  -0.13 0.896
ColdWar      -0.023580   0.015008  -1.57 0.116
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Total Sum of Squares: 1440
Residual Sum of Squares: 1430
R-Squared: 0.0124
Adj. R-Squared: 0.0117
Chisq: 81.0788 on 5 DF, p-value: 4.99e-16
```

A Nicer Table (Yet Again)

Table: Models of Demonstrations

	OLS	FE	BE	RE
POLITY	0.006*** (0.001)	0.002 (0.002)	0.006 (0.005)	0.003* (0.001)
POLITY Squared	-0.002*** (0.0003)	-0.002*** (0.0003)	-0.003*** (0.001)	-0.002*** (0.0003)
ln(GDP)	0.058*** (0.008)	0.055*** (0.015)	0.069** (0.027)	0.057*** (0.012)
Monarch	-0.046 (0.029)	0.048 (0.068)	-0.050 (0.103)	-0.007 (0.053)
Cold War	0.028** (0.014)	-0.035** (0.016)	0.259*** (0.085)	-0.024 (0.015)
Constant	-0.125** (0.058)		-0.306 (0.208)	-0.132 (0.105)
Observations	6,505	6,505	145	6,505
R ²	0.026	0.013	0.127	0.012
Adjusted R ²	0.025	-0.010	0.096	0.012
F Statistic	34.820*** (df = 5; 6499)	16.720*** (df = 5; 6355)	4.062*** (df = 5; 139)	81.080***

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

“Random” Effects: Testing

Hausman test (FE vs. RE):

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

$$W \sim \chi_k^2$$

Issues:

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

Hausman Test: Intuition

If the data-generating process is the result of “random” effects [that is, if $\text{Cov}(\mathbf{X}_{it}, \alpha_i = 0)$]:

- the random-effects estimate $\hat{\beta}_{RE}$ will be consistent *and* (more) efficient
- the fixed-effects / within-unit estimator $\hat{\beta}_{FE}$ will be consistent but *inefficient*

BUT...

If the data-generating process is in fact the result of “fixed” effects [that is, if $\text{Cov}(\mathbf{X}_{it}, \alpha_i \neq 0)$]:

- the fixed-effects / within-unit estimator $\hat{\beta}_{FE}$ will be consistent and efficient
- the random-effects estimate $\hat{\beta}_{RE}$ will be *inconsistent*

Hausman Test Results

Hausman test (FE vs. RE):

```
> phtest(FE, RE)
```

Hausman Test

```
data: lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch + ColdWar  
chisq = 11, df = 5, p-value = 0.05  
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

HLM Starting Points

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

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

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

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

where β_{0j} is a “constant” term, \mathbf{X}_{ij} is a $NJ \times K$ matrix of K covariates, β_j is a $K \times 1$ vector of parameters, and $u_{ij} \sim \text{i.i.d. } N(0, \sigma_u^2)$ 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:

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

for the “intercept” and

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

for the “slopes” of \mathbf{X} . The ε s 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:

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

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

Model Assumptions

- Linearity / Additivity
- Normality of u_s
- 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

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(lnDemos~POLITY+I(POLITY^2)+lnGDP+Monarch+ColdWar+
+             (1|ccode), data=Demos)
>
> summary(AltRE)
Linear mixed model fit by REML ['lmerMod']
Formula: lnDemos ~ POLITY + I(POLITY^2) + lnGDP + Monarch + ColdWar +
          (1 | ccode)
Data: Demos
```

REML criterion at convergence: 9005

Random effects:

Groups	Name	Variance	Std.Dev.
ccode	(Intercept)	0.0536	0.232
Residual		0.2200	0.469

Number of obs: 6507, groups: ccode, 145

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-0.133634	0.104246	-1.28
POLITY	0.002623	0.001447	1.81
I(POLITY^2)	-0.001972	0.000287	-6.88
lnGDP	0.057513	0.012371	4.65
Monarch	-0.015175	0.052863	-0.29
ColdWar	-0.022225	0.014986	-1.48

Correlation of Fixed Effects:

	(Intr)	POLITY	I(POLI	lnGDP	Monrch
POLITY	0.109				
I(POLITY^2)	0.134	-0.135			
lnGDP	-0.968	-0.140	-0.270		
Monarch	0.004	0.172	-0.163	-0.022	
ColdWar	-0.391	0.387	-0.210	0.351	0.014

Q: Are They The Same? [A: Yes]

Table: RE Models of Demonstrations

	RE	AltRE
POLITY	0.003* (0.001)	0.003* (0.001)
POLITY Squared	-0.002*** (0.0003)	-0.002*** (0.0003)
ln(GDP)	0.057*** (0.012)	0.057*** (0.012)
Monarch	-0.007 (0.053)	-0.008 (0.053)
Cold War	-0.024 (0.015)	-0.023 (0.015)
Constant	-0.132 (0.105)	-0.132 (0.104)
Observations	6,505	6,505
R ²	0.012	
Adjusted R ²	0.012	
Log Likelihood		-4,496.000
Akaike Inf. Crit.		9,009.000
Bayesian Inf. Crit.		9,063.000
F Statistic	81.080***	

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

HLM with Country-Level Random β s for lnGDP

```
> HLM1<-lmer(lnDemos~POLITY+I(POLITY^2)+lnGDP+(lnGDP|ccode)+
+           Monarch+ColdWar, data=Demos,
+           control=lmerControl(optimizer="bobyqa"))

> summary(HLM1)
Linear mixed model fit by REML ['lmerMod']
Formula: lnDemos ~ POLITY + I(POLITY^2) + lnGDP + (lnGDP | ccode) + Monarch +
  ColdWar
Data: Demos
Control: lmerControl(optimizer = "bobyqa")
```

Scaled residuals:

	Min	1Q	Median	3Q	Max
	-2.821	-0.442	-0.175	-0.005	8.530

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
ccode	(Intercept)	1.377	1.174	
	lnGDP	0.021	0.145	-0.98
Residual		0.214	0.462	

Number of obs: 6505, groups: ccode, 145

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	-0.21563	0.15416	-1.40
POLITY	0.00198	0.00149	1.33
I(POLITY^2)	-0.00206	0.00030	-6.88
lnGDP	0.06938	0.01894	3.66
Monarch	0.02635	0.05681	0.46
ColdWar	-0.01394	0.01542	-0.90

Correlation of Fixed Effects:

	(Intr)	POLITY	I(POLI	lnGDP	Monrch
POLITY		0.093			
I(POLITY^2)	0.127	-0.096			
lnGDP	-0.984	-0.114	-0.214		
Monarch	-0.015	0.136	-0.167	0.006	
ColdWar	-0.286	0.368	-0.215	0.246	0.019

```
> anova(AltRE,HLM1)

refitting model(s) with ML (instead of REML)
Data: Demos
Models:
AltRE: lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch + ColdWar +
AltRE:      (1 | ccode)
HLM1: lnDemons ~ POLITY + I(POLITY^2) + lnGDP + (lnGDP | ccode) + Monarch +
HLM1:      ColdWar
      npar  AIC  BIC logLik deviance Chisq Df Pr(>Chisq)
AltRE    8 8959 9013  -4471     8943
HLM1    10 8887 8955  -4434     8867  75.6  2    <2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1
```

```
> VarCorr(HLM1)
Groups   Name             Std.Dev. Corr
ccode    (Intercept)  1.174
          lnGDP        0.145   -0.98
Residual                      0.462
```

Random Coefficients

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

```
> head(Bs)
```

	ccode..Intercept.	ccode.POLITY	ccode.I.POLITY.2.	ccode.lnGDP	ccode.Monarch	ccode.ColdWar
2	0.45231	0.001981	-0.002065	0.09778	0.02635	-0.01394
20	0.12457	0.001981	-0.002065	0.02329	0.02635	-0.01394
42	-0.45766	0.001981	-0.002065	0.10707	0.02635	-0.01394
51	-0.04898	0.001981	-0.002065	0.05320	0.02635	-0.01394
52	0.44058	0.001981	-0.002065	-0.02591	0.02635	-0.01394
70	-1.34265	0.001981	-0.002065	0.20813	0.02635	-0.01394

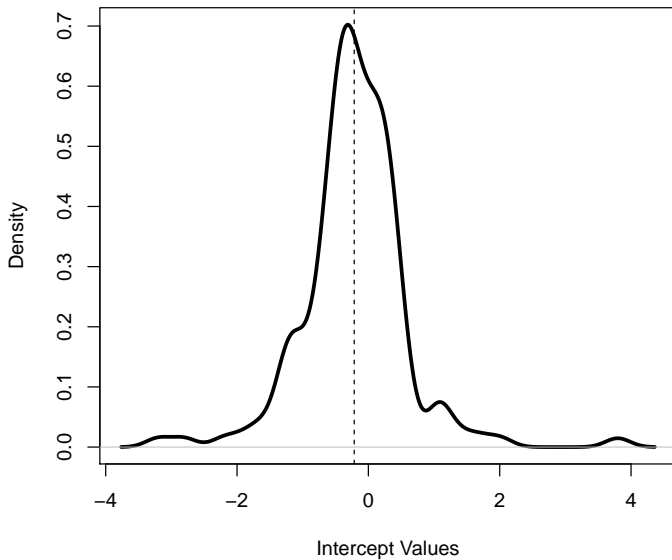
```
> mean(Bs$ccode..Intercept.)
```

```
[1] -0.2156
```

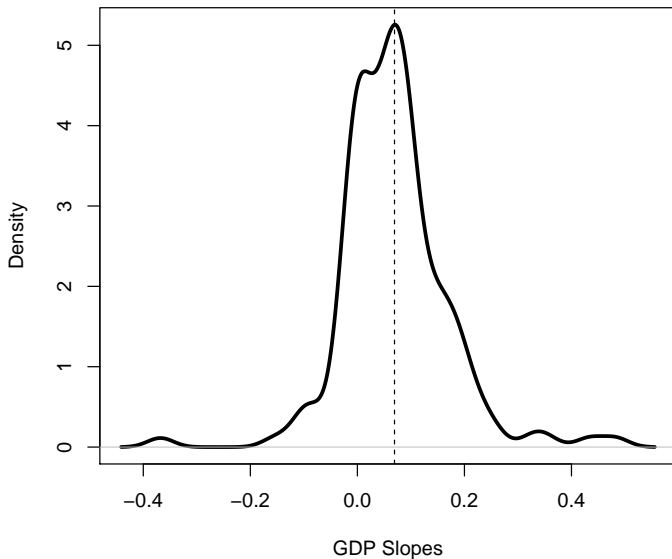
```
> mean(Bs$ccode.lnGDP)
```

```
[1] 0.06938
```

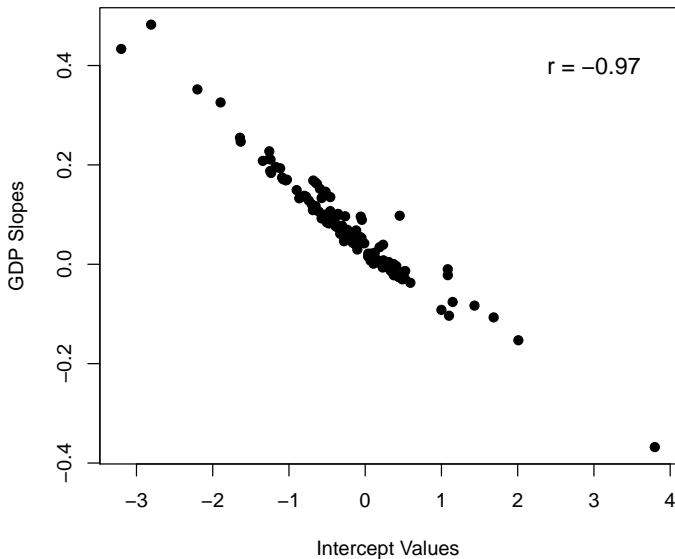

Random Intercepts (Plotted)



Random Slopes for $\ln\text{GDP}$ (Plotted)



Scatterplot: Random Intercepts and Slopes



Separating Within and Between Effects

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

$$Y_{it} = \bar{\mathbf{X}}_i \beta_B + (\mathbf{X}_{it} - \bar{\mathbf{X}}_i) \beta_W + u_{it} \quad (5)$$

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

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

Example data: Separate effects for within- and between-country wealth (GDP)...

Combining Within- and Between-Effects

Table: BE + WE Model of Demonstrations

	WEBE.OLS
POLITY	0.007*** (0.001)
POLITY Squared	-0.002*** (0.0003)
Within-Country ln(GDP)	0.082*** (0.016)
Between-Country ln(GDP)	0.053*** (0.008)
Monarch	-0.042 (0.029)
Cold War	0.040** (0.016)
Constant	-0.091 (0.062)
Observations	6,505
R ²	0.027
Adjusted R ²	0.026
Residual Std. Error	0.526 (df = 6498)
F Statistic	29.510*** (df = 6; 6498)

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

Two-Way Unit Effects

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:

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

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

We can also consider the partial hypotheses:

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

and

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

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 $N = T = 2$ (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(lnDemons~POLITY+I(POLITY^2)+lnGDP+Monarch+ColdWar,  
+ data=PDF,effect="twoway",model="within")  
>  
> summary(TwoWayFE)
```

Twoways effects Within Model

Call:

```
plm(formula = lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch +  
ColdWar, data = PDF, effect = "twoway", model = "within")
```

Unbalanced Panel: n = 145, T = 1-62, N = 6505

Residuals:

Min.	1st Qu.	Median	3rd Qu.	Max.
-1.4299	-0.2154	-0.0790	0.0674	3.7220

Coefficients:

	Estimate	Std. Error	t-value	Pr(> t)
POLITY	0.001864	0.001552	1.20	0.230
I(POLITY^2)	-0.002135	0.000293	-7.29	3.4e-13 ***
lnGDP	-0.013235	0.018399	-0.72	0.472
Monarch	0.112024	0.067260	1.67	0.096 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares: 1350

Residual Sum of Squares: 1330

R-Squared: 0.009

Adj. R-Squared: -0.0239

F-statistic: 14.2917 on 4 and 6295 DF, p-value: 1.29e-11

Two-Way Effects: Testing

```
> # Two-way effects:
>
> pFtest(lnDemons~POLITY+I(POLITY^2)+lnGDP+Monarch+ColdWar,
+       data=PDF,effect="tway",model="within")
```

F test for twoways effects

```
data: lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch + ColdWar
F = 11, df1 = 204, df2 = 6295, p-value <2e-16
alternative hypothesis: significant effects
```

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

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

```
data: lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch + ColdWar
normal = 59, 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) for unbalanced panels

```
data: lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch + ColdWar
normal = 90, p-value <2e-16
alternative hypothesis: significant effects
```

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

Lagrange Multiplier Test - time effects (King and Wu) for unbalanced panels

```
data: lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch + ColdWar
normal = 12, p-value <2e-16
alternative hypothesis: significant effects
```

Two-Way Fixed Effects via `lm`

```
> TwoWayFE.BF<-lm(lnDemons~POLITY+I(POLITY^2)+lnGDP+Monarch+
+               factor(ccode)+factor(Year),data=PDF)

> summary(TwoWayFE.BF)

Call:
lm(formula = lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch +
    factor(ccode) + factor(Year), data = PDF)

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.353540   0.195636    6.92 5.0e-12 ***
POLITY          0.001864   0.001552    1.20 0.22984
I(POLITY^2)    -0.002135   0.000293   -7.29 3.4e-13 ***
lnGDP          -0.013235   0.018399   -0.72 0.47198
Monarch         0.112024   0.067260    1.67 0.09586 .
factor(ccode)20 -1.150899   0.082756  -13.91 < 2e-16 ***
factor(ccode)42 -1.111681   0.091558  -12.14 < 2e-16 ***
factor(ccode)51 -1.172966   0.096647  -12.14 < 2e-16 ***
.
.
.
factor(Year)2000 0.361636   0.088496    4.09 4.4e-05 ***
factor(Year)2001 0.249294   0.088601    2.81 0.00491 **
[ reached getOption("max.print") -- omitted 10 rows ]
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.46 on 6295 degrees of freedom
(2863 observations deleted due to missingness)
Multiple R-squared:  0.278, Adjusted R-squared:  0.254
F-statistic: 11.6 on 209 and 6295 DF,  p-value: <2e-16
```

Example: Two-Way Random Effects

```
> TwoWayRE<-plm(lnDemons~POLITY+I(POLITY^2)+lnGDP+Monarch+ColdWar,  
> data=PDF,effect="twoway",model="random")  
>  
> summary(TwoWayRE)
```

```
Twoways effects Random Effect Model  
(Swamy-Arora's transformation)
```

```
Call:  
plm(formula = lnDemons ~ POLITY + I(POLITY^2) + lnGDP + Monarch +  
ColdWar, data = PDF, effect = "twoway", model = "random")
```

```
Unbalanced Panel: n = 145, T = 1-62, N = 6505
```

```
Effects:
```

	var	std.dev	share
idiosyncratic	0.21186	0.46028	0.77
individual	0.05647	0.23763	0.21
time	0.00527	0.07258	0.02

theta:

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
id	0.11143	0.7136	0.7406	0.7289	0.7611	0.7611
time	0.27832	0.4773	0.4942	0.4840	0.5263	0.5289
total	0.09787	0.4485	0.4720	0.4571	0.4943	0.5081

```
Residuals:
```

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
	-0.41	-0.25	-0.18	0.03	-0.07	4.09

```
Coefficients:
```

	Estimate	Std. Error	z-value	Pr(> z)
(Intercept)	-0.003886	0.244584	-0.02	0.98732
POLITY	0.003325	0.003132	1.06	0.28847
I(POLITY^2)	-0.002048	0.000615	-3.33	0.00088 ***
lnGDP	0.042258	0.028693	1.47	0.14082
Monarch	0.021130	0.114872	0.18	0.85405
ColdWar	-0.045119	0.053826	-0.84	0.40190

```
---  
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1
```

```
Total Sum of Squares: 1850  
Residual Sum of Squares: 1820  
R-Squared: 0.0181  
Adj. R-Squared: 0.0174  
Chisq: 92.4819 on 5 DF, p-value: <2e-16
```

Table: Models of Demonstrations

	OLS	FE	BE	RE	TwoWayFE	TwoWayRE
POLITY	0.006*** (0.001)	0.002 (0.002)	0.006 (0.005)	0.003* (0.001)	0.002 (0.002)	0.003 (0.003)
POLITY Squared	-0.002*** (0.0003)	-0.002*** (0.0003)	-0.003*** (0.001)	-0.002*** (0.0003)	-0.002*** (0.0003)	-0.002*** (0.001)
ln(GDP)	0.058*** (0.008)	0.055*** (0.015)	0.069** (0.027)	0.057*** (0.012)	-0.013 (0.018)	0.042 (0.029)
Monarch	-0.046 (0.029)	0.048 (0.068)	-0.050 (0.103)	-0.007 (0.053)	0.112* (0.067)	0.021 (0.115)
Cold War	0.028** (0.014)	-0.035** (0.016)	0.259*** (0.085)	-0.024 (0.015)		-0.045 (0.054)
Constant	-0.125** (0.058)		-0.306 (0.208)	-0.132 (0.105)		-0.004 (0.245)
Observations	6,505	6,505	145	6,505	6,505	6,505
R ²	0.026	0.013	0.127	0.012	0.009	0.018
Adjusted R ²	0.025	-0.010	0.096	0.012	-0.024	0.017
Residual Std. Error	0.526 (df = 6499)					

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

“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_i s
- 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