

# PLSC 504 – Fall 2017

## One-Way Unit Effects

October 19, 2017

# 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:

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

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:

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

and

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

Yields:

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

Means that:

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

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

≡ “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  $\sim F_{N-1, NT-(N-1)}$ .

# An Example: Refugee Flows in Africa, 1992-2001

Data:

- 50 African countries  $\rightarrow (50 \times 49 = )$  2450 directed dyads
- Ten years
- $i$  indexes directed dyads,  $t$  indexes years

Model:

$$\ln(\text{Refugees})_{A \rightarrow Bt} = \beta_0 + \beta_1 \text{Population Difference}_{ABt} + \beta_2 \text{Distance}_{AB} + \beta_3 \text{POLITY Difference}_{ABt} + \beta_4 \text{War Difference}_{ABt} + u_{ABt}$$

# Data: Refugee Flows in Africa, 1992-2001

```
> summary(Refugees)
```

dirdyadID	year	ln_ref_flow	pop_diff
Min. :404411	Min. :1992	Min. : -0.6931	Min. : -0.117949
1st Qu.:451461	1st Qu.:1994	1st Qu.: -0.6931	1st Qu.: -0.008848
Median :510520	Median :1996	Median : -0.6931	Median : 0.000000
Mean :512160	Mean :1996	Mean : -0.6011	Mean : 0.000000
3rd Qu.:565553	3rd Qu.:1999	3rd Qu.: -0.6931	3rd Qu.: 0.008848
Max. :651625	Max. :2001	Max. :14.1343	Max. : 0.117949

distance	regimedif	wardiff	pop_between
Min. :0.000	Min. : -1.00	Min. : -4	Min. : -0.109517
1st Qu.:1.299	1st Qu.: -0.25	1st Qu.: 0	1st Qu.: -0.008833
Median :2.169	Median : 0.00	Median : 0	Median : 0.000000
Mean :2.200	Mean : 0.00	Mean : 0	Mean : 0.000000
3rd Qu.:3.066	3rd Qu.: 0.25	3rd Qu.: 0	3rd Qu.: 0.008833
Max. :5.652	Max. : 1.00	Max. : 4	Max. : 0.109517

pop_within	regime_between	regime_within	war_between
Min. : -0.0088492	Min. : -0.955	Min. : -1.180	Min. : -2.3
1st Qu.: -0.0004707	1st Qu.: -0.225	1st Qu.: -0.085	1st Qu.: -0.4
Median : 0.0000000	Median : 0.000	Median : 0.000	Median : 0.0
Mean : 0.0000000	Mean : 0.000	Mean : 0.000	Mean : 0.0
3rd Qu.: 0.0004707	3rd Qu.: 0.225	3rd Qu.: 0.085	3rd Qu.: 0.4
Max. : 0.0088492	Max. : 0.955	Max. : 1.180	Max. : 2.3

war_within
Min. : -2.5
1st Qu.: -0.3
Median : 0.0
Mean : 0.0
3rd Qu.: 0.3
Max. : 2.5

# An Example: Refugee Flows in Africa, 1992-2001

## Pooled OLS:

```
> RefOLS<-lm(ln_ref_flow~pop_diff+distance+regimedif+wardiff, data=Refugees)
> summary(RefOLS)
```

### Residuals:

Min	1Q	Median	3Q	Max
-0.6114	-0.2109	-0.0857	0.0335	14.3756

### Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.3224073	0.0119195	-27.049	<2e-16 ***
pop_diff	-0.1732934	0.2166658	-0.800	0.424
distance	-0.1266528	0.0047016	-26.938	<2e-16 ***
regimedif	-0.0002476	0.0157962	-0.016	0.987
wardiff	0.0743220	0.0068169	10.903	<2e-16 ***

---

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

Residual standard error: 0.9097 on 23613 degrees of freedom

Multiple R-squared: 0.03467, Adjusted R-squared: 0.03451

F-statistic: 212 on 4 and 23613 DF, p-value: < 2.2e-16



# An Example: Refugee Flows in Africa, 1992-2001

“Fixed” effects:

```
> library(plm)
> RefFE<-plm(ln_ref_flow~pop_diff+distance+regimedif+wardiff,
  data=Refugees, effect="individual", model="within")
> summary(RefFE)
Oneway (individual) effect Within Model
```

Unbalanced Panel: n=2450, T=1-10, N=23618

Residuals :

Min.	1st Qu.	Median	3rd Qu.	Max.
-9.03e+00	-5.74e-03	-9.18e-06	5.72e-03	1.14e+01

Coefficients :

	Estimate	Std. Error	t-value	Pr(> t )
pop_diff	6.8642028	2.5516636	2.6901	0.007149 **
regimedif	0.0050497	0.0223160	0.2263	0.820984
wardiff	0.0104144	0.0073673	1.4136	0.157493

---

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

Total Sum of Squares: 8149.6

Residual Sum of Squares: 8146

R-Squared : 0.00043949

Adj. R-Squared : 0.00039385

F-statistic: 3.102 on 3 and 21165 DF, p-value: 0.025509

# An Example: Refugee Flows in Africa, 1992-2001

Models of Refugees in Africa		
Variable	OLS	Fixed Effects
Constant	-0.32 (0.01)	-
Population Difference	-0.17 (0.22)	6.86 (2.55)
Distance	-0.13 (0.005)	(dropped)
POLITY Difference	-0.0002 (0.016)	0.005 (0.022)
War Difference	0.074 (0.007)	0.010 (0.007)
$\hat{\rho}$	-	0.61

Note:  $NT = 23618$  ( $N = 2450$ ,  $\bar{T} = 9.6$ )

# Issues (?) with “Fixed” Effects

## Pros:

- Specification Bias
- Intuitive
- Widely Used/Understood

## Cons:

- Can't Estimate  $\beta_B$
- Slowly-Changing  $\mathbf{X}$ s
- (In)Efficiency / Inconsistency (Incidental Parameters)

From:

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

“Between” effects:

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

- Essentially cross-sectional
- Based on  $N$  observations

# Refugee Flows in Africa, 1992-2001

“Between” effects:

```
> RefBE<-plm(ln_ref_flow~pop_diff+distance+regimedif+wardiff, data=Refugees,
  effect="individual", model="between")
> summary(RefBE)
Oneway (individual) effect Between Model
```

Unbalanced Panel: n=2450, T=1-10, N=23618

Residuals :

Min.	1st Qu.	Median	3rd Qu.	Max.
-0.5850	-0.2200	-0.0840	0.0534	9.6500

Coefficients :

	Estimate	Std. Error	t-value	Pr(> t )
(Intercept)	-0.299703	0.029741	-10.0771	< 2.2e-16 ***
pop_diff	-0.246861	0.525232	-0.4700	0.6384
distance	-0.134874	0.011755	-11.4742	< 2.2e-16 ***
regimedif	0.010709	0.045117	0.2374	0.8124
wardiff	0.124185	0.022004	5.6439	1.855e-08 ***

---

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

Total Sum of Squares: 1383.9

Residual Sum of Squares: 1296.7

R-Squared : 0.063042

Adj. R-Squared : 0.062913

F-statistic: 41.1269 on 4 and 2445 DF, p-value: < 2.22e-16

# Refugee Example Redux

Variable	OLS	Fixed ("Within") Effects	Between Effects
Constant	-0.32 (0.01)	-	-0.30 (0.03)
Population Difference	-0.17 (0.22)	6.86 (2.55)	-0.25 (0.53)
Distance	-0.13 (0.005)	(dropped)	-0.13 (0.01)
POLITY Difference	-0.0002 (0.016)	0.005 (0.022)	0.01 (0.05)
War Difference	0.074 (0.007)	0.010 (0.007)	0.12 (0.02)
$\hat{\rho}$	-	0.61	-

Note:  $NT = 23618$  ( $N = 2450$ ,  $\bar{T} = 9.6$ ).

Model:

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

with:

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

and

$$E(\alpha_i) = E(\lambda_t) = E(\eta_{it}) = 0,$$

$$E(\alpha_i\lambda_t) = E(\alpha_i\eta_{it}) = E(\lambda_t\eta_{it}) = 0,$$

$$E(\alpha_i\alpha_j) = \sigma_\alpha^2 \text{ if } i = j, \text{ 0 otherwise,}$$

$$E(\lambda_t\lambda_s) = \sigma_\lambda^2 \text{ if } t = s, \text{ 0 otherwise,}$$

$$E(\eta_{it}\eta_{js}) = \sigma_\eta^2 \text{ if } i = j, t = s, \text{ 0 otherwise,}$$

$$E(\alpha_i\mathbf{X}_{it}) = E(\lambda_t\mathbf{X}_{it}) = E(\eta_{it}\mathbf{X}_{it}) = 0.$$

“Variance Components”:

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

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

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

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

# “Random” Effects: Estimation

Can estimate:

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

where

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

With  $\hat{\theta}$ , calculate:

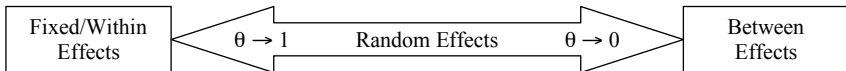
$$\begin{aligned} Y_{it}^* &= Y_{it} - \hat{\theta} \bar{Y}_i \\ X_{it}^* &= X_{it} - \hat{\theta} \bar{X}_i, \end{aligned}$$

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...

# “Random” Effects: An Alternative View



# Refugees Redux

```
> RefRE<-plm(ln_ref_flow~pop_diff+distance+regimedif+wardiff, data=Refugees,
  effect="individual", model="random")
> summary(RefRE)
Oneway (individual) effect Random Effect Model
(Swamy-Arora's transformation)
```

Unbalanced Panel: n=2450, T=1-10, N=23618

Effects:

	var	std.dev	share
idiosyncratic	0.3849	0.6204	0.466
individual	0.4416	0.6645	0.534

theta :

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.3176	0.7168	0.7168	0.7141	0.7168	0.7168

Coefficients :

	Estimate	Std. Error	t-value	Pr(> t )
(Intercept)	-0.3063941	0.0285299	-10.7394	< 2.2e-16 ***
pop_diff	0.0638665	0.4974613	0.1284	0.897845
distance	-0.1324536	0.0112685	-11.7544	< 2.2e-16 ***
regimedif	0.0005633	0.0198580	0.0284	0.977370
wardiff	0.0228523	0.0069775	3.2751	0.001058 **

---

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

Total Sum of Squares: 9216.6

Residual Sum of Squares: 9158.9

R-Squared : 0.0062699

Adj. R-Squared : 0.0062686

F-statistic: 37.177 on 4 and 23613 DF, p-value: < 2.22e-16

# Refugees Redux, Remix

```
> library(lme4)

> AltRefRE<-lmer(ln_ref_flow~pop_diff+distance+regimedif+wardiff+(1|dirtyadID), data=Refugees)
> summary(AltRefRE)
Linear mixed model fit by REML
Formula: ln_ref_flow ~ pop_diff + distance + regimedif + wardiff + (1 |      dirtyadID)
Data: Refugees
      AIC      BIC logLik deviance REMLdev
50733 50790 -25360    50692    50719
Random effects:
Groups      Name      Variance Std.Dev.
dirtyadID (Intercept) 0.46653  0.68303
Residual              0.38592  0.62123
Number of obs: 23618, groups: dirtyadID, 2450

Fixed effects:
              Estimate Std. Error t value
(Intercept) -0.3061471  0.0291477 -10.503
pop_diff     0.0758989  0.5075942   0.150
distance     -0.1325429  0.0115127 -11.513
regimedif     0.0007138  0.0199078   0.036
wardiff       0.0223476  0.0069779   3.203

Correlation of Fixed Effects:
      (Intr) pp_dff distnc regmdf
pop_diff    0.000
distance   -0.869  0.000
regimedif    0.000  0.036  0.000
wardiff     0.000 -0.004  0.000  0.109
```

Variable	OLS	Fixed Effects	Between Effects	Random Effects
Constant	-0.32 (0.01)	-	-0.30 (0.03)	-0.31 (0.03)
Population Difference	-0.17 (0.22)	6.86 (2.55)	-0.25 (0.53)	0.09 (0.52)
Distance	-0.13 (0.005)	(dropped)	-0.13 (0.01)	-0.13 (0.01)
POLITY Difference	-0.0002 (0.016)	0.005 (0.022)	0.01 (0.05)	0.0005 (0.0199)
War Difference	0.074 (0.007)	0.010 (0.007)	0.12 (0.02)	0.023 (0.007)
$\hat{\rho}$	-	0.61	-	0.56

Note:  $NT = 23618$  ( $N = 2450$ ,  $\bar{T} = 9.6$ ).

# “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 (FE vs. RE):

```
> phtest(RefFE, AltRefRE)
```

Hausman Test

```
data: ln_ref_flow ~ pop_diff + distance + regimedif + wardiff  
chisq = 34.712, df = 3, p-value = 0.0000001401  
alternative hypothesis: one model is inconsistent
```



# Practical “Fixed” vs. “Random” Effects

- “Panel” vs. “TSCS” Data
- Data-Generating Process
- Covariate Effects

# Separating Within and Between Effects

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

- Simple...
- Easy interpretation
- Easy to test  $\hat{\boldsymbol{\beta}}_B = \hat{\boldsymbol{\beta}}_W$

## Again With The Refugees...

Variable	Estimate
Constant	-0.32 (0.01)
Distance	-0.13 (0.004)
Between (Mean) Population Difference	-0.22 (0.22)
Within Population Difference	6.86 (3.74)
Between (Mean) POLITY Difference	0.01 (0.02)
Within POLITY Difference	0.005 (0.032)
Between (Mean) War Difference	0.12 (0.01)
Within War Difference	0.01 (0.01)

Note:  $NT = 23618$  ( $N = 2450$ ,  $\bar{T} = 9.6$ ).

R :

- the `lme4` package; command is `lmer`
- the `plm` package; `plm` command
- the `nlme` package; command `lme`

Stata : `xtreg`

- the `re` (the default) = random effects
- the `fe` = fixed (within) effects
- the `be` = between-effects