Panel and Multilevel Data

Philip Leifeld

GV903: Advanced Research Methods, Week 9

University of Essex

1. Concepts and Data Structures

Multilevel Data

Consider this data structure:

Obs.	Group	DV	IV_1	IV_2
1	A	12.24	0.23	-33.23
2	A	11.09	0.19	-43.01
3	A	13.94	0.56	-28.75
4	B	16.41	0.68	-9.33
5	B	17.64	0.55	-12.46
6	C	15.00	0.72	-19.89

Multilevel Data

Consider this data structure:

Obs.	Group	DV	IV_1	IV_2
1	A	12.24	0.23	-33.23
2	A	11.09	0.19	-43.01
3	A	13.94	0.56	-28.75
4	B	16.41	0.68	-9.33
5	B	17.64	0.55	-12.46
6	C	15.00	0.72	-19.89

Observations are nested in groups.

Multilevel Data

Consider this data structure:

Obs.	Group	DV	IV_1	IV_2
1	A	12.24	0.23	-33.23
2	A	11.09	0.19	-43.01
3	A	13.94	0.56	-28.75
4	B	16.41	0.68	-9.33
5	B	17.64	0.55	-12.46
6	C	15.00	0.72	-19.89

Observations are nested in groups.

Examples: students in classes; voters in districts; candidates in parties; municipalities in states; respondents in scenarios.

Panel Data

Now consider this data structure:

Obs.	Respondent	Time	DV	IV_1	IV_2
1	1	1	12.24	0.23	-33.23
2	1	2	11.09	0.19	-43.01
3	1	3	13.94	0.56	-28.75
4	2	1	16.41	0.68	-9.33
5	2	2	17.64	0.55	-12.46
6	3	1	15.00	0.72	-19.89

Panel Data

Now consider this data structure:

Obs.	Respondent	Time	DV	IV_1	IV_2
1	1	1	12.24	0.23	-33.23
2	1	2	11.09	0.19	-43.01
3	1	3	13.94	0.56	-28.75
4	2	1	16.41	0.68	-9.33
5	2	2	17.64	0.55	-12.46
6	3	1	15.00	0.72	-19.89

Observations are nested in time series per respondent.

Panel Data

Now consider this data structure:

Obs.	Respondent	Time	DV	IV_1	IV_2
1	1	1	12.24	0.23	-33.23
2	1	2	11.09	0.19	-43.01
3	1	3	13.94	0.56	-28.75
4	2	1	16.41	0.68	-9.33
5	2	2	17.64	0.55	-12.46
6	3	1	15.00	0.72	-19.89

Observations are nested in time series per respondent.

Or: observations are nested in time points (if we don't care about the "respondent" variable.)

Panel Data

Now consider this data structure:

Obs.	Respondent	Time	DV	IV_1	IV_2
1	1	1	12.24	0.23	-33.23
2	1	2	11.09	0.19	-43.01
3	1	3	13.94	0.56	-28.75
4	2	1	16.41	0.68	-9.33
5	2	2	17.64	0.55	-12.46
6	3	1	15.00	0.72	-19.89

Observations are nested in time series per respondent.

Or: observations are nested in time points (if we don't care about the "respondent" variable.)

Examples: voters' party preferences; countries' democracy scores.

What are TSCS Data?

What are TSCS Data?

- Panel data: more respondents than time points.

What are TSCS Data?

- Panel data: more respondents than time points.
- Example: repeated multi-country survey.

What are TSCS Data?

- Panel data: more respondents than time points.
- Example: repeated multi-country survey.
- TSCS (= time series cross-section): more time points than units.

What are TSCS Data?

- Panel data: more respondents than time points.
- Example: repeated multi-country survey.
- TSCS (= time series cross-section): more time points than units.
- Example: annual measurements of OECD countries since 1960.

What are TSCS Data?

- Panel data: more respondents than time points.
- Example: repeated multi-country survey.
- TSCS (= time series cross-section): more time points than units.
- Example: annual measurements of OECD countries since 1960.
- In principle, the same methods for both kinds of data.

What are TSCS Data?

- Panel data: more respondents than time points.
- Example: repeated multi-country survey.
- TSCS (= time series cross-section): more time points than units.
- Example: annual measurements of OECD countries since 1960.
- In principle, the same methods for both kinds of data.
- But: TSCS makes more use of time series methods and spatial regression.

Implications of Multilevel or Panel Data

What are the implications for a regression model (pooled OLS)?

Implications of Multilevel or Panel Data

What are the implications for a regression model (pooled OLS)?

- The errors (= residuals) are not independent of each other.

Implications of Multilevel or Panel Data

What are the implications for a regression model (pooled OLS)?

- The errors (= residuals) are not independent of each other.
- They are clustered in the groups (respondents, time points).

Implications of Multilevel or Panel Data

What are the implications for a regression model (pooled OLS)?

- The errors (= residuals) are not independent of each other.
- They are clustered in the groups (respondents, time points).
- This pattern may explain parts of the variance in the DV.

Implications of Multilevel or Panel Data

What are the implications for a regression model (pooled OLS)?

- The errors (= residuals) are not independent of each other.
- They are clustered in the groups (respondents, time points).
- This pattern may explain parts of the variance in the DV.
- Example: There are good and bad school classes due to factors outside our control, such as neighbourhood deprivation or teacher quality.

Implications of Multilevel or Panel Data

What are the implications for a regression model (pooled OLS)?

- The errors (= residuals) are not independent of each other.
- They are clustered in the groups (respondents, time points).
- This pattern may explain parts of the variance in the DV.
- Example: There are good and bad school classes due to factors outside our control, such as neighbourhood deprivation or teacher quality.
- Example: When measuring democracy in countries, there may be more variation across countries than within each time series (due to serial autocorrelation).

Implications of Multilevel or Panel Data

What are the implications for a regression model (pooled OLS)?

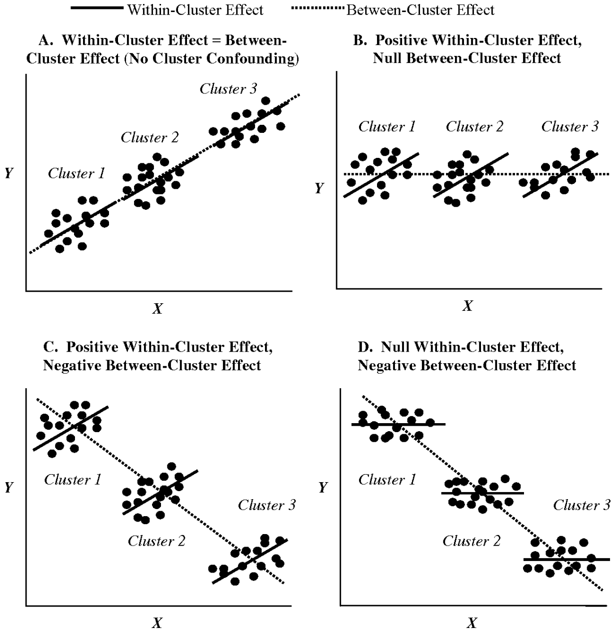
- The errors (= residuals) are not independent of each other.
- They are clustered in the groups (respondents, time points).
- This pattern may explain parts of the variance in the DV.
- Example: There are good and bad school classes due to factors outside our control, such as neighbourhood deprivation or teacher quality.
- Example: When measuring democracy in countries, there may be more variation across countries than within each time series (due to serial autocorrelation).
- The error clustering may also be correlated with the IVs.

Implications of Multilevel or Panel Data

What are the implications for a regression model (pooled OLS)?

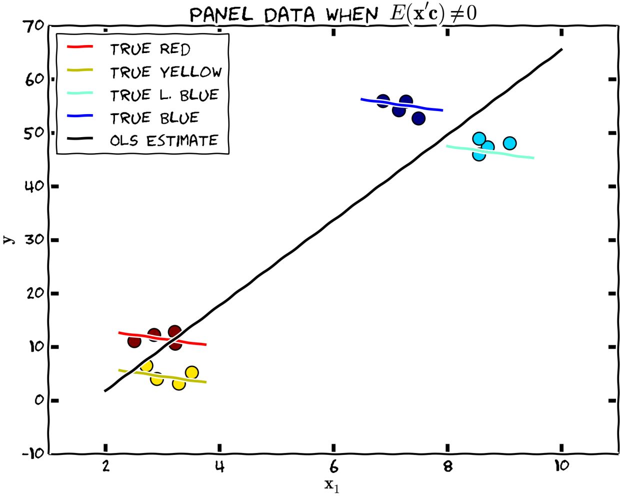
- The errors (= residuals) are not independent of each other.
- They are clustered in the groups (respondents, time points).
- This pattern may explain parts of the variance in the DV.
- Example: There are good and bad school classes due to factors outside our control, such as neighbourhood deprivation or teacher quality.
- Example: When measuring democracy in countries, there may be more variation across countries than within each time series (due to serial autocorrelation).
- The error clustering may also be correlated with the IVs.
- This violates the OLS assumption of unsystematic errors.

The Consequence: Cluster Confounding (Bartels 2011)



Case 1: Slopes (= IVs) are Affected

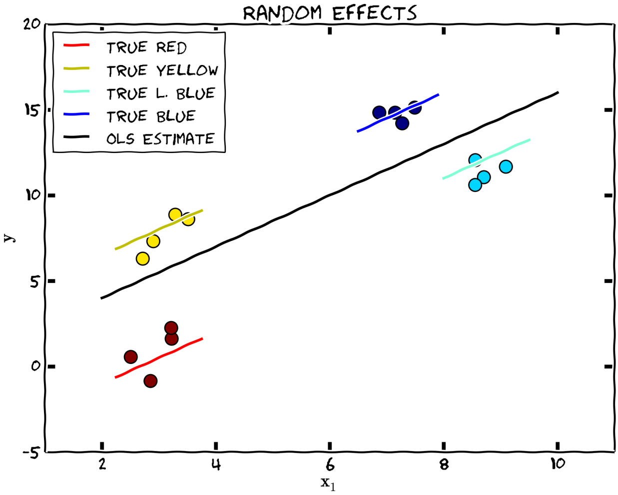
Source: `https://rlhick.people.wm.edu/econ407/notes/intro_panel_data.html`



In this case, we can use a Fixed Effects Model.

Case 2: Only Intercepts are Affected

Source: `https://rlhick.people.wm.edu/econ407/notes/intro_panel_data.html`



In this case, we can use a Random Effects Model.

Exercise

1. Can you come up with a political science example of a panel dataset, TSCS dataset, and multilevel dataset each, along with a research question (DV and IV)?
2. In each of these examples, what are the individual-level units and what are the higher-level units/groups?
3. Do you expect any cluster confounding in these cases?

What Do Fixed and Random Effects Do?

- Fixed effects control explicitly for known groups in the data.

What Do Fixed and Random Effects Do?

- Fixed effects control explicitly for known groups in the data.
- They "eliminate" any between-group variation (hence "fixed").

What Do Fixed and Random Effects Do?

- Fixed effects control explicitly for known groups in the data.
- They "eliminate" any between-group variation (hence "fixed").
- Can be modeled as group dummies, with group demeaning, or via first differences.

What Do Fixed and Random Effects Do?

- Fixed effects control explicitly for known groups in the data.
- They "eliminate" any between-group variation (hence "fixed").
- Can be modeled as group dummies, with group demeaning, or via first differences.
- Random effects rather permit different intercepts by allowing the different groups to have a variance (to be measured from the data).

What Do Fixed and Random Effects Do?

- Fixed effects control explicitly for known groups in the data.
- They "eliminate" any between-group variation (hence "fixed").
- Can be modeled as group dummies, with group demeaning, or via first differences.
- Random effects rather permit different intercepts by allowing the different groups to have a variance (to be measured from the data).
- There are also random-slope models.

What Do Fixed and Random Effects Do?

- Fixed effects control explicitly for known groups in the data.
- They "eliminate" any between-group variation (hence "fixed").
- Can be modeled as group dummies, with group demeaning, or via first differences.
- Random effects rather permit different intercepts by allowing the different groups to have a variance (to be measured from the data).
- There are also random-slope models.
- These and random effects models are both sub-types of mixed-effects models.

What Do Fixed and Random Effects Do?

- Fixed effects control explicitly for known groups in the data.
- They "eliminate" any between-group variation (hence "fixed").
- Can be modeled as group dummies, with group demeaning, or via first differences.
- Random effects rather permit different intercepts by allowing the different groups to have a variance (to be measured from the data).
- There are also random-slope models.
- These and random effects models are both sub-types of mixed-effects models.
- Let's start with fixed effects models!

2. Fixed Effects

The Fixed Effects Model

You can control for any between-group variance directly by using dummy variables. I. e., one binary indicator D_j per group j:

$$y_{ij} = \sum_j \delta_j D_j + \beta_1 x_{ij} + \ldots + u_{ij}$$

The Fixed Effects Model

You can control for any between-group variance directly by using dummy variables. I. e., one binary indicator D_j per group j:

$$y_{ij} = \sum_j \delta_j D_j + \beta_1 x_{ij} + \ldots + u_{ij}$$

This explains all variance due to different group memberships effectively. Whatever the correlates of those group memberships may be.

The Fixed Effects Model

You can control for any between-group variance directly by using dummy variables. I. e., one binary indicator D_j per group j:

$$y_{ij} = \sum_j \delta_j D_j + \beta_1 x_{ij} + \ldots + u_{ij}$$

This explains all variance due to different group memberships effectively. Whatever the correlates of those group memberships may be.

BUT: we have to add as many model terms as we have groups to the model. Not very parsimonious...

The Fixed Effects Model

You can control for any between-group variance directly by using dummy variables. I. e., one binary indicator D_j per group j:

$$y_{ij} = \sum_j \delta_j D_j + \beta_1 x_{ij} + \ldots + u_{ij}$$

This explains all variance due to different group memberships effectively. Whatever the correlates of those group memberships may be.

BUT: we have to add as many model terms as we have groups to the model. Not very parsimonious...

There is a fix for this: group demeaning!

The Fixed Effects Model: Group Demeaning

We can decompose the error term into a group-level component (a_j) and the remaining, "idiosyncratic" error (u_{ij}):

$$y_{ij} = \beta_1 x_{ij} + \ldots + a_j + u_{ij}$$

The Fixed Effects Model: Group Demeaning

We can decompose the error term into a group-level component (a_j) and the remaining, "idiosyncratic" error (u_{ij}):

$$y_{ij} = \beta_1 x_{ij} + \ldots + a_j + u_{ij}$$

For each group j, we can average this over the individuals included in the group:

$$\bar{y}_j = \beta_1 \bar{x}_j + \ldots + a_j + \bar{u}_j$$

The Fixed Effects Model: Group Demeaning

We can decompose the error term into a group-level component (a_j) and the remaining, "idiosyncratic" error (u_{ij}):

$$y_{ij} = \beta_1 x_{ij} + \ldots + a_j + u_{ij}$$

For each group j, we can average this over the individuals included in the group:

$$\bar{y}_j = \beta_1 \bar{x}_j + \ldots + a_j + \bar{u}_j$$

We can now do the group demeaning:

$$(y_{ij} - \bar{y}_j) = \beta_1 (x_{ij} - \bar{x}_j) + \ldots + (a_j - a_j) + (u_{ij} - \bar{u}_j)$$

The Fixed Effects Model: Group Demeaning

We can decompose the error term into a group-level component (a_j) and the remaining, "idiosyncratic" error (u_{ij}):

$$y_{ij} = \beta_1 x_{ij} + \ldots + a_j + u_{ij}$$

For each group j, we can average this over the individuals included in the group:

$$\bar{y}_j = \beta_1 \bar{x}_j + \ldots + a_j + \bar{u}_j$$

We can now do the group demeaning:

$$(y_{ij} - \bar{y}_j) = \beta_1 (x_{ij} - \bar{x}_j) + \ldots + (a_j - a_j) + (u_{ij} - \bar{u}_j)$$

This is the most common fixed-effects estimator.

The Fixed Effects Model: Group Demeaning

We can decompose the error term into a group-level component (a_j) and the remaining, "idiosyncratic" error (u_{ij}):

$$y_{ij} = \beta_1 x_{ij} + \ldots + a_j + u_{ij}$$

For each group j, we can average this over the individuals included in the group:

$$\bar{y}_j = \beta_1 \bar{x}_j + \ldots + a_j + \bar{u}_j$$

We can now do the group demeaning:

$$(y_{ij} - \bar{y}_j) = \beta_1 (x_{ij} - \bar{x}_j) + \ldots + (a_j - a_j) + (u_{ij} - \bar{u}_j)$$

This is the most common fixed-effects estimator.

Note how a_j drops out of the equation. We got rid of any group-level variation and do not have to include dummies.

Practical Meaning of Fixed Effects

- In a multilevel context, this means we express each observation as a difference from its group mean.

Practical Meaning of Fixed Effects

- In a multilevel context, this means we express each observation as a difference from its group mean.
- This makes all observations directly comparable across groups. It kills all between-group effects.

Practical Meaning of Fixed Effects

- In a multilevel context, this means we express each observation as a difference from its group mean.
- This makes all observations directly comparable across groups. It kills all between-group effects.
- It yields the same result as including dummies, except the dummies did not have to be included explicitly as separate terms.

Practical Meaning of Fixed Effects

- In a multilevel context, this means we express each observation as a difference from its group mean.
- This makes all observations directly comparable across groups. It kills all between-group effects.
- It yields the same result as including dummies, except the dummies did not have to be included explicitly as separate terms.
- In a panel context, we do time demeaning instead: for each observation, we subtract the individual's mean over time.

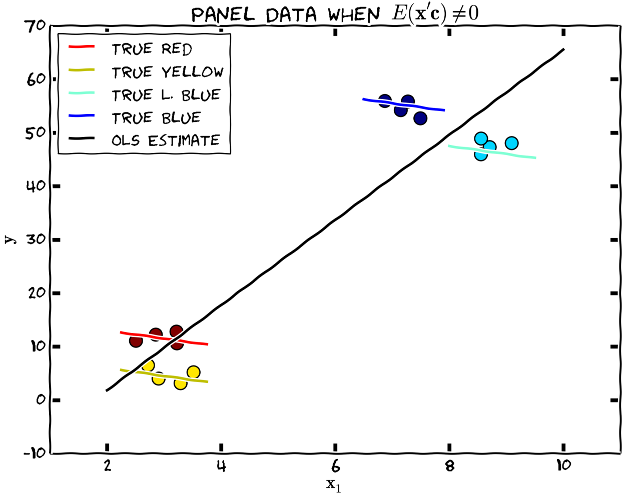
Practical Meaning of Fixed Effects

- In a multilevel context, this means we express each observation as a difference from its group mean.
- This makes all observations directly comparable across groups. It kills all between-group effects.
- It yields the same result as including dummies, except the dummies did not have to be included explicitly as separate terms.
- In a panel context, we do time demeaning instead: for each observation, we subtract the individual's mean over time.
- This puts the different individuals' time series on a comparable scale.

Practical Meaning of Fixed Effects

- In a multilevel context, this means we express each observation as a difference from its group mean.
- This makes all observations directly comparable across groups. It kills all between-group effects.
- It yields the same result as including dummies, except the dummies did not have to be included explicitly as separate terms.
- In a panel context, we do time demeaning instead: for each observation, we subtract the individual's mean over time.
- This puts the different individuals' time series on a comparable scale.
- Then we can compare the remaining variation over different time points because we have fixed all individual variation.

Fixed Effects



Think of the demeaning as moving all these groups or time series onto a common scale (on both axes)! Then estimate pooled OLS.

Let's Do Fixed Effects in R!

Source: https://www.princeton.edu/~otorres/Panel101R.pdf

First, load some panel data.

```
library("foreign")
Panel <- read.dta(
  "http://dss.princeton.edu/training/Panel101.dta")
```

Let's Do Fixed Effects in R!

Source: https://www.princeton.edu/~otorres/Panel101R.pdf

First, load some panel data.

```
library("foreign")
Panel <- read.dta(
  "http://dss.princeton.edu/training/Panel101.dta")
```

Let's fit a fixed-effects model with dummies:

```
fixed.dum <- lm(y ~ x1 + factor(country) - 1,
                data = Panel)
```

Let's Do Fixed Effects in R!

Source: `https://www.princeton.edu/~otorres/Panel101R.pdf`

First, load some panel data.

```r
library("foreign")
Panel <- read.dta(
  "http://dss.princeton.edu/training/Panel101.dta")
```

Let's fit a fixed-effects model with dummies:

```r
fixed.dum <- lm(y ~ x1 + factor(country) - 1,
                data = Panel)
```

Now with country demeaning. `plm` is a package for panel data.

```r
library("plm")
fixed <- plm(y ~ x1, data = Panel,
             index = c("country", "year"),
             model = "within")
```

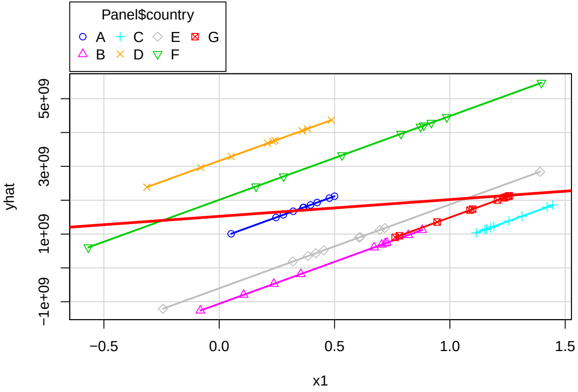
Results of the Two Models

	Model 1	Model 2
x1	2475617827.10*	2475617827.10*
	(1106675593.60)	(1106675593.60)
factor(country)A	880542403.99	
	(961807052.24)	
factor(country)B	−1057858363.16	
	(1051067684.19)	
factor(country)C	−1722810754.55	
	(1631513751.40)	
factor(country)D	3162826897.32***	
	(909459149.66)	
factor(country)E	−602622000.33	
	(1064291684.41)	
factor(country)F	2010731793.24	
	(1122809097.35)	
factor(country)G	−984717493.45	
	(1492723118.24)	
R^2	0.44	0.07
Adj. R^2	0.37	−0.03
Num. obs.	70	70

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

We can Inspect the Dummy Variable Groups

```
yhat <- predict(fixed.dum)
library("car")
scatterplot(yhat ~ Panel$x1|Panel$country, boxplots = FALSE,
            xlab = "x1", ylab = "yhat", smooth = FALSE)
abline(lm(Panel$y ~ Panel$x1), lwd = 3, col = "red")
```



Within- and Between-Estimation

On the previous slides, we did within-estimation. That means we want to look at the variation within each time series and fix the cross-series variance (or generally, the cross-group variance).

We could also do between-estimation, thereby fixing the group-internal or time-series-internal variance (e. g., time demeaning) and focusing on the variance between groups or time series.

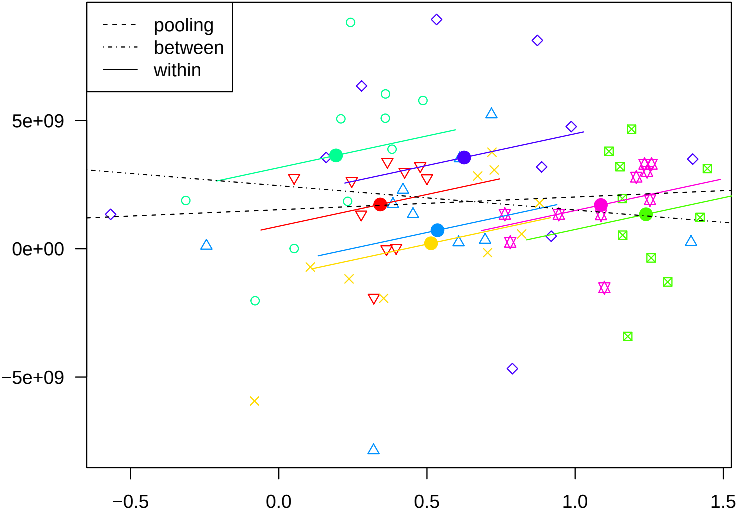
This would be equivalent to including time-point dummies in the panel case.

```
fixed2 <- plm(y ~ x1, data = Panel,
              index = c("country", "year"),
              model = "between")
```

Plotting Within- and Between-Estimation

This time the observed values, not fitted values.

```
plot(fixed, within = TRUE, between = TRUE)
```



The Hausman Test

The Hausman test can assess whether fixed or random effects should be used.

The Hausman Test

The Hausman test can assess whether fixed or random effects should be used.

If significant, there is a correlation between X and a_j, which is incompatible with the random-effects model.

The Hausman Test

The Hausman test can assess whether fixed or random effects should be used.

If significant, there is a correlation between X and a_j, which is incompatible with the random-effects model.

```
ranef <- plm(y ~ x1, data = Panel,
             index = c("country", "year"),
             model = "random")
phtest(fixed, ranef)
##
##  Hausman Test
##
## data:  y ~ x1
## chisq = 3.674, df = 1, p-value = 0.05527
## alternative hypothesis: one model is inconsistent
```

The Hausman Test

The Hausman test can assess whether fixed or random effects should be used.

If significant, there is a correlation between X and a_j, which is incompatible with the random-effects model.

```
ranef <- plm(y ~ x1, data = Panel,
             index = c("country", "year"),
             model = "random")
phtest(fixed, ranef)
##
##  Hausman Test
##
## data:  y ~ x1
## chisq = 3.674, df = 1, p-value = 0.05527
## alternative hypothesis: one model is inconsistent
```

Close call. . . probably random effects are sufficient.

Difference-in-Difference Analysis

Difference-in-difference analysis is a causal inference technique to assess the effect of treatment conditions (e. g., natural experiments, policy changes. . .) in panel data.

Difference-in-Difference Analysis

Difference-in-difference analysis is a causal inference technique to assess the effect of treatment conditions (e. g., natural experiments, policy changes. . .) in panel data.

We do not have to assume repeated observations. There can be different individuals at different time points.

Difference-in-Difference Analysis

Difference-in-difference analysis is a causal inference technique to assess the effect of treatment conditions (e. g., natural experiments, policy changes. . .) in panel data.

We do not have to assume repeated observations. There can be different individuals at different time points.

Example: Do disgusting pictures on cigarette packs lead to lower cigarette consumption?

Difference-in-Difference Analysis

Difference-in-difference analysis is a causal inference technique to assess the effect of treatment conditions (e. g., natural experiments, policy changes. . .) in panel data.

We do not have to assume repeated observations. There can be different individuals at different time points.

Example: Do disgusting pictures on cigarette packs lead to lower cigarette consumption?

We need respondents from states where this policy was adopted and where it was not adopted, both before and after the (potential) adoption.

Difference-in-Difference Analysis

Difference-in-difference analysis is a causal inference technique to assess the effect of treatment conditions (e. g., natural experiments, policy changes. . .) in panel data.

We do not have to assume repeated observations. There can be different individuals at different time points.

Example: Do disgusting pictures on cigarette packs lead to lower cigarette consumption?

We need respondents from states where this policy was adopted and where it was not adopted, both before and after the (potential) adoption. Then the DiD estimator is:

$$cons = \beta_0 + \beta_1 picState + \beta_2 time2 + \beta_3 picState \cdot time2 + u$$

Difference-in-Difference Analysis

Difference-in-difference analysis is a causal inference technique to assess the effect of treatment conditions (e. g., natural experiments, policy changes. . .) in panel data.

We do not have to assume repeated observations. There can be different individuals at different time points.

Example: Do disgusting pictures on cigarette packs lead to lower cigarette consumption?

We need respondents from states where this policy was adopted and where it was not adopted, both before and after the (potential) adoption. Then the DiD estimator is:

$$cons = \beta_0 + \beta_1 picState + \beta_2 time2 + \beta_3 picState \cdot time2 + u$$

We can include further controls. The interaction tells us if the intervention works.

Exercise

1. Can you come up with an example where fixed-effects estimation makes sense?
2. Describe the DV and IV, the nesting structure of the data, and the research question.
3. Would you need a within- or between-estimator to answer this question?
4. Describe a natural or field experiment in which you could use a difference-in-difference estimator. How would you do this from an analysis perspective? That is, try to write down the regression equation to identify the causal effect of the treatment.

Fixed Effects – Concluding Thoughts

- R^2 not directly comparable – refers only to demeaned data (i. e., part of the variance)!

Fixed Effects – Concluding Thoughts

- R^2 not directly comparable – refers only to demeaned data (i. e., part of the variance)!
- Can be difficult to wrap your head around, but can be effective in some situations.

Fixed Effects – Concluding Thoughts

- R^2 not directly comparable – refers only to demeaned data (i. e., part of the variance)!
- Can be difficult to wrap your head around, but can be effective in some situations.
- Need to decide among within- and between-estimation.

Fixed Effects – Concluding Thoughts

- R^2 not directly comparable – refers only to demeaned data (i. e., part of the variance)!
- Can be difficult to wrap your head around, but can be effective in some situations.
- Need to decide among within- and between-estimation.
- Some people (e. g., Bell and Jones 2014, PSRM) argue that fixed effects are never needed because mixed-effects models can do all of this, but better.

Fixed Effects – Concluding Thoughts

- R^2 not directly comparable – refers only to demeaned data (i. e., part of the variance)!
- Can be difficult to wrap your head around, but can be effective in some situations.
- Need to decide among within- and between-estimation.
- Some people (e. g., Bell and Jones 2014, PSRM) argue that fixed effects are never needed because mixed-effects models can do all of this, but better.
- You may want to play around with the `plm` package for panel data.

Fixed Effects – Concluding Thoughts

- R^2 not directly comparable – refers only to demeaned data (i. e., part of the variance)!
- Can be difficult to wrap your head around, but can be effective in some situations.
- Need to decide among within- and between-estimation.
- Some people (e. g., Bell and Jones 2014, PSRM) argue that fixed effects are never needed because mixed-effects models can do all of this, but better.
- You may want to play around with the `plm` package for panel data.
- For mixed/random effects, packages `nlme` and `lme4` are more general and more powerful.

Fixed Effects – Concluding Thoughts

- R^2 not directly comparable – refers only to demeaned data (i. e., part of the variance)!
- Can be difficult to wrap your head around, but can be effective in some situations.
- Need to decide among within- and between-estimation.
- Some people (e. g., Bell and Jones 2014, PSRM) argue that fixed effects are never needed because mixed-effects models can do all of this, but better.
- You may want to play around with the `plm` package for panel data.
- For mixed/random effects, packages `nlme` and `lme4` are more general and more powerful.
- Let's look at random effects with `lme4`...

3. Random Effects and Mixed Effects

`lme4` Syntax

The `lme4` package contains the `lmer` function to fit random- and mixed-effects models.

`lme4` Syntax

The `lme4` package contains the `lmer` function to fit random- and mixed-effects models.

```
model <- lmer(dv ~ x1 + (1|x2), data = d)
```

lme4 Syntax

The `lme4` package contains the `lmer` function to fit random- and mixed-effects models.

```
model <- lmer(dv ~ x1 + (1|x2), data = d)
```

`(1|x2)` introduces random intercepts for different groups in `x2`.

`lme4` Syntax

The `lme4` package contains the `lmer` function to fit random- and mixed-effects models.

```
model <- lmer(dv ~ x1 + (1|x2), data = d)
```

`(1|x2)` introduces random intercepts for different groups in `x2`.

```
model <- lmer(dv ~ (1 + x1|x2), data = d)
```

lme4 Syntax

The `lme4` package contains the `lmer` function to fit random- and mixed-effects models.

```
model <- lmer(dv ~ x1 + (1|x2), data = d)
```

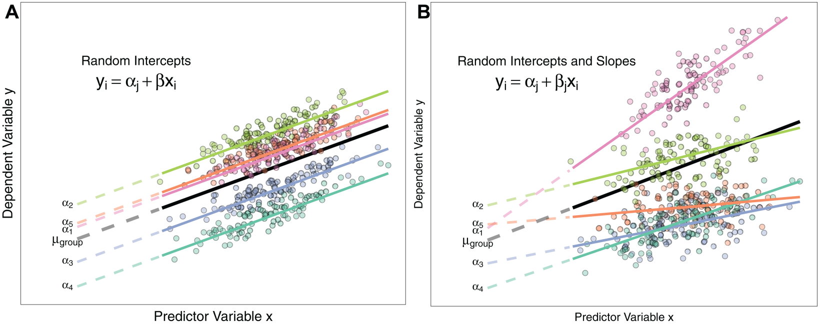
`(1|x2)` introduces random intercepts for different groups in `x2`.

```
model <- lmer(dv ~ (1 + x1|x2), data = d)
```

`(1 + x1|x2)` introduces random intercepts (`1`) *and* `x1` slopes for the `x2` grouping variable.

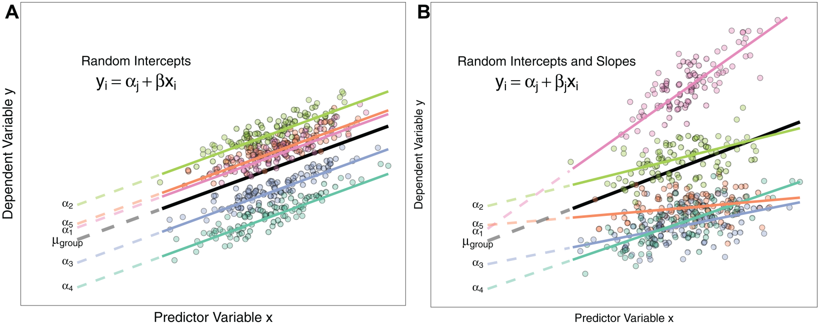
Random Intercepts and Random Slopes

Source: Harrison et al. (2018), *PeerJ*. `https://peerj.com/articles/4794/`



Random Intercepts and Random Slopes

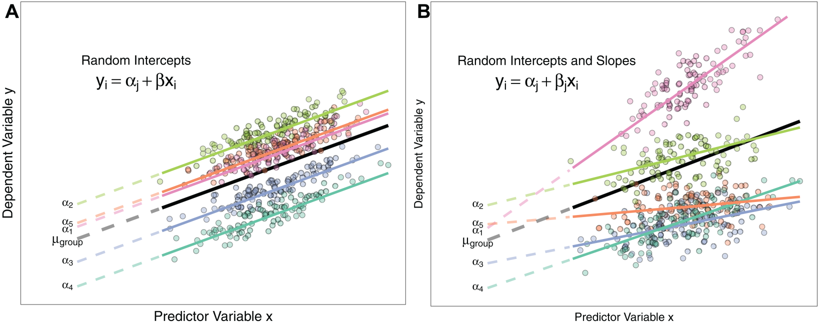
Source: Harrison et al. (2018), *PeerJ*. `https://peerj.com/articles/4794/`



Example: Pupils' grades are a function of class attendance. Pupils are nested in schools with all sorts of characteristics, which we can't control for but which affect grades. Hence random intercepts for the schools.

Random Intercepts and Random Slopes

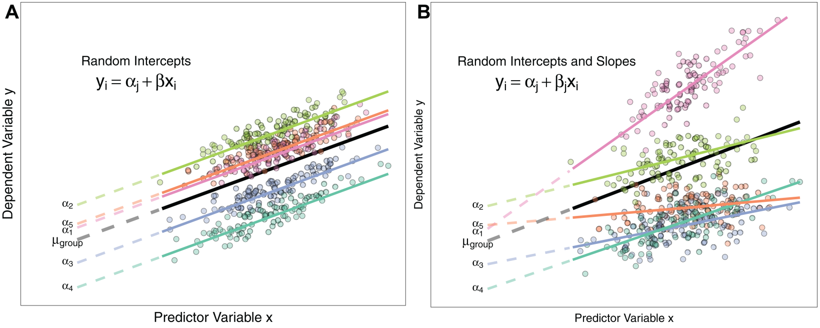
Source: Harrison et al. (2018), *PeerJ*. `https://peerj.com/articles/4794/`



If we think the link between attendance and grades differs by school, perhaps because teacher quality varies, we need random intercepts *and* slopes.

Random Intercepts and Random Slopes

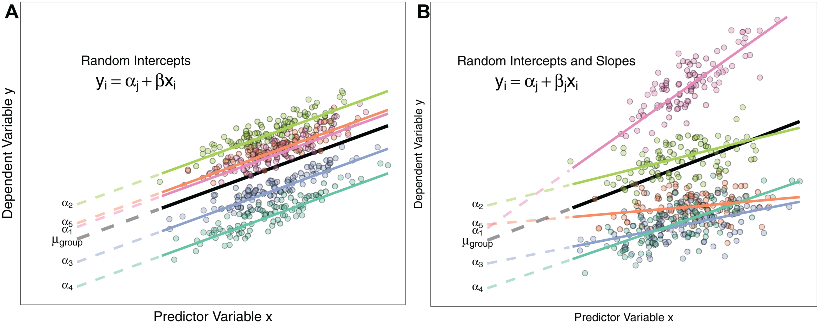
Source: Harrison et al. (2018), *PeerJ*. `https://peerj.com/articles/4794/`



Another example: Countries' foreign direct investment (FDI) over time. If we explain FDI by lagged democracy, we need to account for baseline differences in FDI across countries using random intercepts.

Random Intercepts and Random Slopes

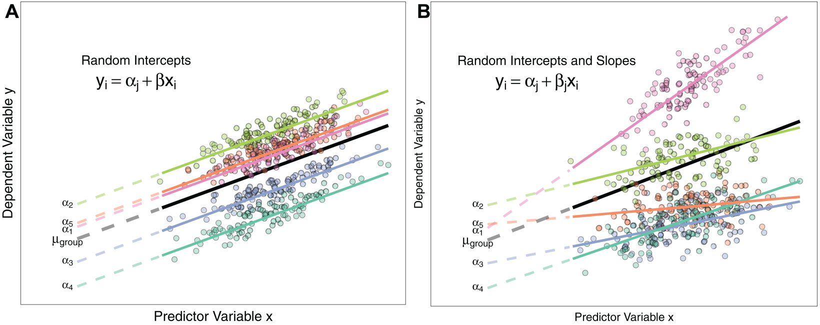
Source: Harrison et al. (2018), *PeerJ*. `https://peerj.com/articles/4794/`



If we think the link between democracy and FDI differs by country, we need random intercepts *and* slopes.

Random Intercepts and Random Slopes

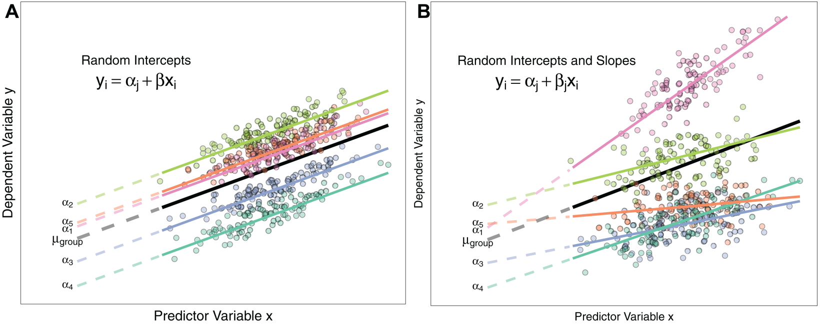
Source: Harrison et al. (2018), *PeerJ*. `https://peerj.com/articles/4794/`



μ_{group} is the mean of the intercept of all groups. The model output reports the variance across groups around this random effect. If we fit random slopes, we also get a variance for those.

Random Intercepts and Random Slopes

Source: Harrison et al. (2018), *PeerJ*. `https://peerj.com/articles/4794/`



Note how β is replaced by β_j in the random slopes equation.
Note also how the random intercepts are denoted by a_j.
Estimation is complicated; `lme4` uses Bayesian estimation.

Interviewer Ratings of Respondent Political Knowledge

Data from the 2000 American National Election Study

```
library("pscl")
data("politicalInformation")
head(politicalInformation)
##              y collegeDegree female age homeOwn govt length id
## 1 Fairly High           Yes     No  49     Yes   No  58.40  1
## 2     Average            No    Yes  35     Yes   No  46.15  2
## 3   Very High            No    Yes  57     Yes   No  89.52  3
## 4     Average            No     No  63     Yes   No  92.63  4
## 5 Fairly High           Yes    Yes  40     Yes   No  58.85  4
## 6     Average            No     No  77     Yes   No  53.82  4

politicalInformation$y2 <- as.numeric(politicalInformation$y)
```

Interviewers may have different baseline rates of judging respondents' knowledge, perhaps conditioned on their own expertise, age, gender etc. Better use a random intercept for interviewers.

Interviewer Ratings of Respondent Political Knowledge

Data from the 2000 American National Election Study

```
library("lme4")
model1 <- lm(y2 ~ collegeDegree + female + age + homeOwn + govt +
  length, data = politicalInformation)
model2 <- lmer(y2 ~ collegeDegree + female + age + homeOwn +
  govt + length + (1|id), data = politicalInformation)
```

	model1	model2
(Intercept)	$2.28\ (0.10)^{***}$	$2.29\ (0.10)^{***}$
collegeDegreeYes	$0.82\ (0.05)^{***}$	$0.83\ (0.05)^{***}$
femaleYes	$-0.36\ (0.05)^{***}$	$-0.39\ (0.05)^{***}$
age	$0.00\ (0.00)^{**}$	$0.00\ (0.00)^{*}$
homeOwnYes	$0.27\ (0.05)^{***}$	$0.27\ (0.05)^{***}$
govtYes	$0.10\ (0.08)$	$0.10\ (0.07)$
length	$0.01\ (0.00)^{***}$	$0.01\ (0.00)^{***}$

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

There is not a big difference here. But there could have been...

Do We Need a Random Effect Here?

Some additional output from `summary(model2)`:

```
Random effects:
 Groups   Name        Variance Std.Dev.
 id       (Intercept) 0.1349   0.3673
 Residual             0.8907   0.9438
Number of obs: 1790, groups:  id, 115

Fixed effects:
                 Estimate Std. Error t value
(Intercept)      2.292365   0.103488  22.151
collegeDegreeYes 0.825025   0.048383  17.052
femaleYes       -0.385396   0.046078  -8.364
age              0.003748   0.001473   2.545
homeOwnYes       0.266954   0.051635   5.170
govtYes          0.095319   0.072731   1.311
length           0.008518   0.001114   7.645
```

Do We Need a Random Effect Here?

Some additional output from `summary(model2)`:

```
Random effects:
 Groups   Name        Variance Std.Dev.
 id       (Intercept) 0.1349   0.3673
 Residual             0.8907   0.9438
Number of obs: 1790, groups:  id, 115

Fixed effects:
                  Estimate Std. Error t value
(Intercept)       2.292365   0.103488  22.151
collegeDegreeYes  0.825025   0.048383  17.052
femaleYes        -0.385396   0.046078  -8.364
age               0.003748   0.001473   2.545
homeOwnYes        0.266954   0.051635   5.170
govtYes           0.095319   0.072731   1.311
length            0.008518   0.001114   7.645
```

The variance of 0.13 indicates it is better than not to include a random effect. We also have theoretical reasons to do it.

When Should You Use Random versus Fixed Effects?

Dieleman and Templin (2014), *PLoS ONE*.

Both RE and FE estimation rely on the assumptions of OLS. The estimated models [. . .] must be correctly specified, each variable of x must be strictly exogenous and linearly independent, and the residual must be independently and identically distributed. When these conditions are met, theory states that FE estimation is unbiased and consistent. RE estimation requires an additional assumption—the group-level effect and the included explanatory variables must be independent in order to avoid [omitted variable bias]. When this assumption is met, RE estimation is unbiased, consistent, and, because it utilized both the within- and between-group variation, efficient. Under this assumption, FE estimation is not efficient because it only utilizes the within-group variation [. . .]. Thus, the correlation between the explanatory variable(s) and group-level effects distinguishes which of these two estimators to utilize.

When Should You Use Random versus Fixed Effects?

Dieleman and Templin (2014), *PLoS ONE*.

> *Both RE and FE estimation rely on the assumptions of OLS. The estimated models [. . .] must be correctly specified, each variable of x must be strictly exogenous and linearly independent, and the residual must be independently and identically distributed. When these conditions are met, theory states that FE estimation is unbiased and consistent. RE estimation requires an additional assumption—the group-level effect and the included explanatory variables must be independent in order to avoid [omitted variable bias]. When this assumption is met, RE estimation is unbiased, consistent, and, because it utilized both the within- and between-group variation, efficient. Under this assumption, FE estimation is not efficient because it only utilizes the within-group variation [. . .]. Thus, the correlation between the explanatory variable(s) and group-level effects distinguishes which of these two estimators to utilize.*

However, simulations have shown that RE is robust to even extreme correlations. So people in practice mostly use RE because more flexible. (See Clark/Linzer and Bell/Jones in PSRM.)

Mixed Effects: Concluding Thoughts

Mixed Effects: Concluding Thoughts

- Use fixed or random/mixed effects whenever you have cross-sectional data that are additionally nested in groups (multilevel data) or time series (panel or TSCS data).

Mixed Effects: Concluding Thoughts

- Use fixed or random/mixed effects whenever you have cross-sectional data that are additionally nested in groups (multilevel data) or time series (panel or TSCS data).
- With TSCS data, you should additionally be aware of temporal, spatial, or network dependence that can arise and take care of it if possible.

Mixed Effects: Concluding Thoughts

- Use fixed or random/mixed effects whenever you have cross-sectional data that are additionally nested in groups (multilevel data) or time series (panel or TSCS data).
- With TSCS data, you should additionally be aware of temporal, spatial, or network dependence that can arise and take care of it if possible.
- Temporal or multilevel data offer larger sample sizes and advantages for causal identification (e. g., difference-in-difference approach).

Mixed Effects: Concluding Thoughts

- Use fixed or random/mixed effects whenever you have cross-sectional data that are additionally nested in groups (multilevel data) or time series (panel or TSCS data).
- With TSCS data, you should additionally be aware of temporal, spatial, or network dependence that can arise and take care of it if possible.
- Temporal or multilevel data offer larger sample sizes and advantages for causal identification (e. g., difference-in-difference approach).
- But the nesting needs to be accounted for in terms of the variances of groups/time points/individuals, otherwise there may be bias and inefficiency.

Mixed Effects: Concluding Thoughts

- Use fixed or random/mixed effects whenever you have cross-sectional data that are additionally nested in groups (multilevel data) or time series (panel or TSCS data).
- With TSCS data, you should additionally be aware of temporal, spatial, or network dependence that can arise and take care of it if possible.
- Temporal or multilevel data offer larger sample sizes and advantages for causal identification (e. g., difference-in-difference approach).
- But the nesting needs to be accounted for in terms of the variances of groups/time points/individuals, otherwise there may be bias and inefficiency.
- Mixed effects are more flexible, but fixed effects make more sense when you want to completely eliminate any between- or within-variation.

Mixed Effects: Concluding Thoughts

- Use fixed or random/mixed effects whenever you have cross-sectional data that are additionally nested in groups (multilevel data) or time series (panel or TSCS data).
- With TSCS data, you should additionally be aware of temporal, spatial, or network dependence that can arise and take care of it if possible.
- Temporal or multilevel data offer larger sample sizes and advantages for causal identification (e. g., difference-in-difference approach).
- But the nesting needs to be accounted for in terms of the variances of groups/time points/individuals, otherwise there may be bias and inefficiency.
- Mixed effects are more flexible, but fixed effects make more sense when you want to completely eliminate any between- or within-variation.
- Kreft and de Leeuw (1998): "Once you know that hierarchies exist, you see them everywhere."

Exercise

1. Can you find a political science example where you need a random intercept?
2. Can you find a political science example where you might need random slopes?
3. Write down the `lme4` formula for estimating the (imaginary) model.