Multilevel Workshop

Andrew Grogan-Kaylor

2026-01-14

Table of contents

List of Figures

List of Tables

1 Introduction

> "Listening to the world. Well, I did that, and I still do it. I still do it." (Mary Oliver in Oliver and Tippett 2015)

This site contains materials for a workshop on multilevel modeling.

1.1 Background

Multilevel models are useful when you have data that are nested or clustered inside social units such as schools, neighborhoods, states, or countries.

Multilevel models are also useful when you have longitudinal data where repeated measures are collected for study participants.

The Importance of Multilevel Models

Multilevel models may improve one's statistical inferences in two important substantive ways.

- Multilevel models adjust standard errors for clustering, and thus calculate appropriate p values. Failure to use a model that accounts for the clustering in the data may lead to improperly calculated standard errors and p values, possibly leading to false attributions of statistical significance (false positives) (see Chapter 2).
- Multilevel models adjust regression coefficients (β's) for the presence of clustering. Failure to use a model that accounts for the clustering in the data may lead to improperly calculated regression coefficients (β's) which may have the wrong magnitude, the wrong statistical significance, and even the wrong sign (see Chapter 3).

1.2 Simulated Multilevel Data

The data used in these workshop materials are *simulated* data on parents, children and families. The data are simulated to come from 30 hypothetical countries around the world. These are the same data used and discussed in my book *Multilevel Thinking: Discovering Variation, Universals, and Particulars in Cross-Cultural Research*.

There are two versions of the data: a *cross-sectional* data set from a single point in time; a *longitudinal* version of the data spanning several time points.

> **The Data Can Be Downloaded Here:**
>
> - Cross Sectional Data
> - Longitudinal Data

Table 1.1: Variables in Simulated Multilevel Data

pos	variable	label
1	country	country id
2	HDI	Human Development Index
3	family	family id
4	id	unique country family id
5	identity	hypothetical identity group variable
6	intervention	recieved intervention
7	physical_punishment	physical punishment in past week
8	warmth	parental warmth in past week
9	outcome	beneficial outcome

Table 1.2: Sample Data From Simulated Multilevel Data

country	HDI	family	id	identity	intervention	physical_punishment	warmth	outcome
15	77	59	15.59	0	0	3	6	57.13
2	83	49	2.49	0	1	3	4	57.51
16	57	50	16.50	1	0	1	5	51
24	82	38	24.38	0	0	1	6	59.41
28	53	32	28.32	0	0	5	2	45.1

2 Estimation of p Values

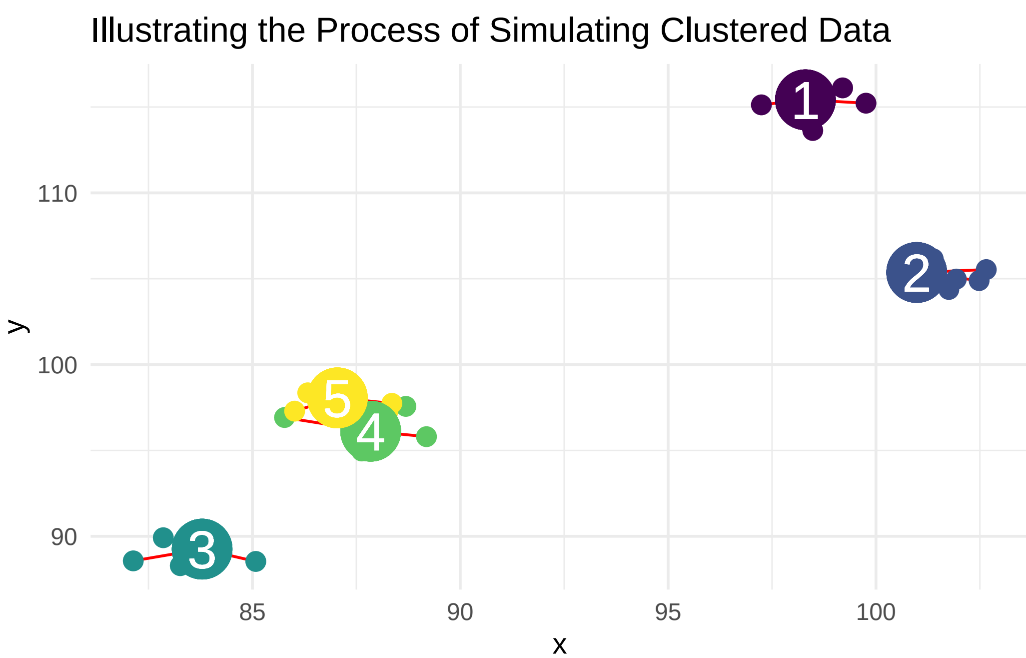
2.1 Grouped and Individual Data

Bland and Altman (1994) suggested the following procedure for simulating some data:

> "The data were generated from random numbers, and there is no relation between X and Y at all. Firstly, values of X and Y were generated for each 'subject,' then a further random number was added to make the individual observation." (Bland and Altman 1994)

So... we follow their procedure.

> **ℹ Simulating The Data**
>
> The graph below illustrates the process of simulating the data.



2.2 Analyses

2.2.1 OLS

An OLS analysis indicates that there is a statistically significant association of x and y.

	OLS1
x_individual	1.046 **
Intercept	4.488
Number of observations	25

** p<.01, * p<.05

2.2.2 MLM

In contrast, an MLM analysis (correctly) finds that there is no statistically significant association of x and y.

	MLM1
x_individual	0.039
Intercept	97.005 **
var(_cons)	74.523
var(e)	0.594
Number of observations	25

** p<.01, * p<.05

2.2.3 Compare OLS and MLM

	OLS1	MLM1
x_individual	1.046 **	0.039
Intercept	4.488	97.005 **
var(_cons)		74.523
var(e)		0.594
Number of observations	25	25

** p<.01, * p<.05

3 Estimation of β Coefficients

3.1 Introduction

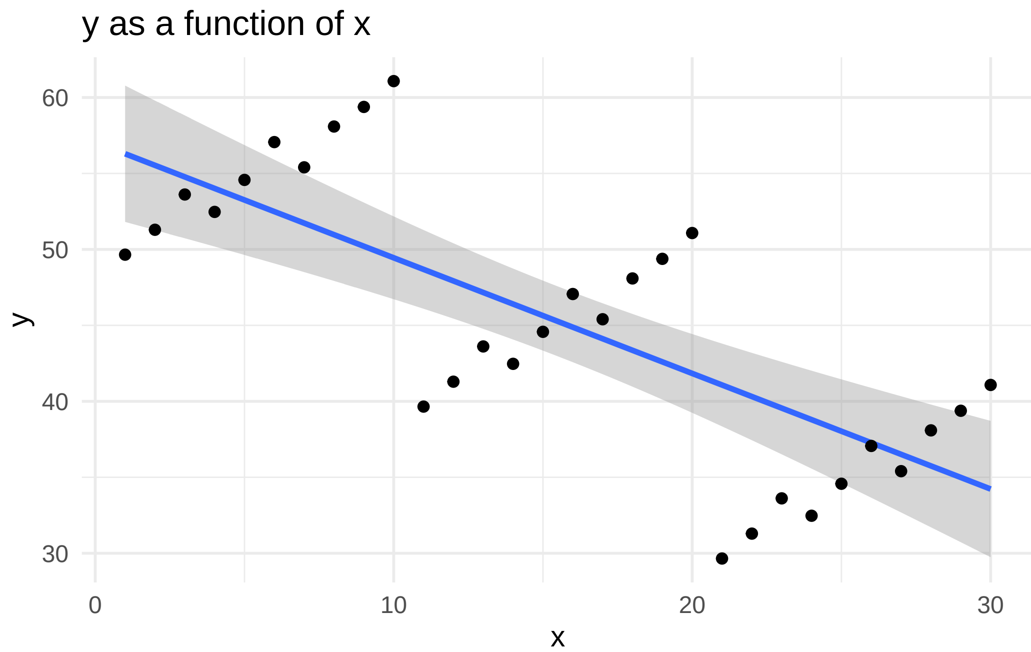
Associations between two variables can be *very different* (or even *reversed*) depending upon whether or not the analysis is "aware" of the grouped, nested, or clustered nature of the data (Nieuwenhuis 2015; Diez Roux 2003; Gelman et al. 2007). In multilevel analysis, the groups are often schools, neighborhoods, communities, or countries.

A model that is "aware" of the clustered nature of the data may provide very different–likely better–substantive conclusions than a model that is not aware of the clustered nature of the data. This phenomena is closely related to the "ecological fallacy": the idea that group level and individual level relationships are not necessarily the same (Firebaugh 2001).

3.2 Graphs

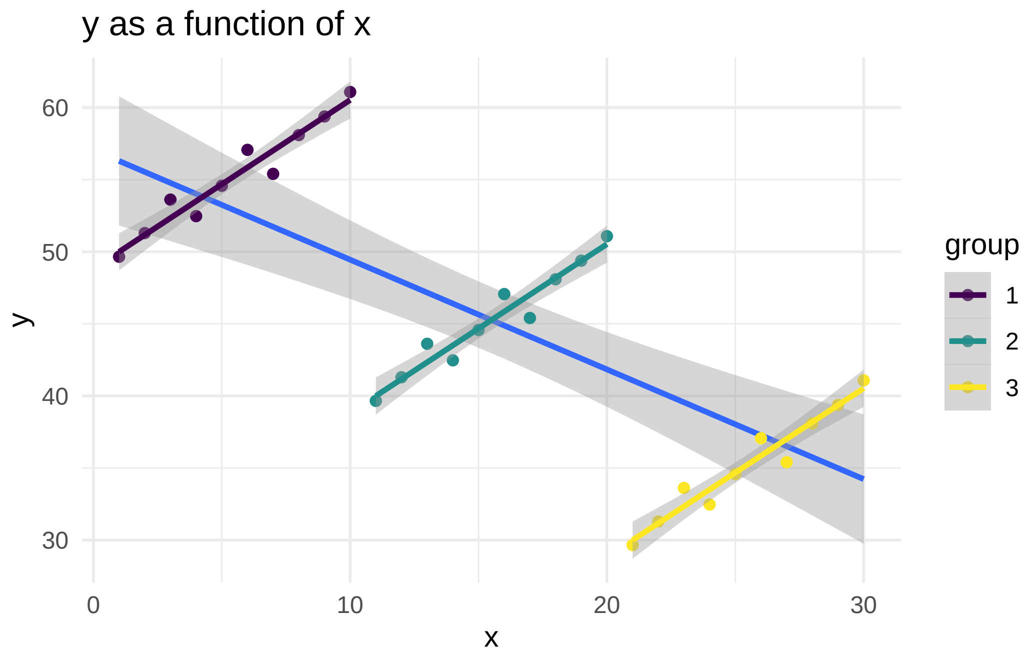
3.2.1 A "Naive" Graph

This "naive" graph is unaware of the grouped nature of the data.



3.2.2 An "Aware" Graph

This "aware" graph is aware of the grouped nature of the data.



3.3 Regressions

3.3.1 A "Naive" OLS Analysis

The OLS model with only x as a covariate is not aware of the grouped structure of the data, and the coefficient for x reflects this.

	OLS2
x	-0.761 **
Intercept	57.057 **
Number of observations	30

** p<.01, * p<.05

3.3.2 An "Aware" MLM Analysis

The multilevel model is aware of the grouped structure of the data, and the coefficient for x reflects this.

	MLM2
x	1.166 **
Intercept	27.192 **
var(_cons)	312.623
var(e)	0.806
Number of observations	30

** p<.01, * p<.05

3.3.3 Compare The Models

	OLS2	MLM2
x	-0.761 **	1.166 **
Intercept	57.057 **	27.192 **
var(_cons)		312.623
var(e)		0.806
Number of observations	30	30

** p<.01, * p<.05

3.4 A Thought Experiment

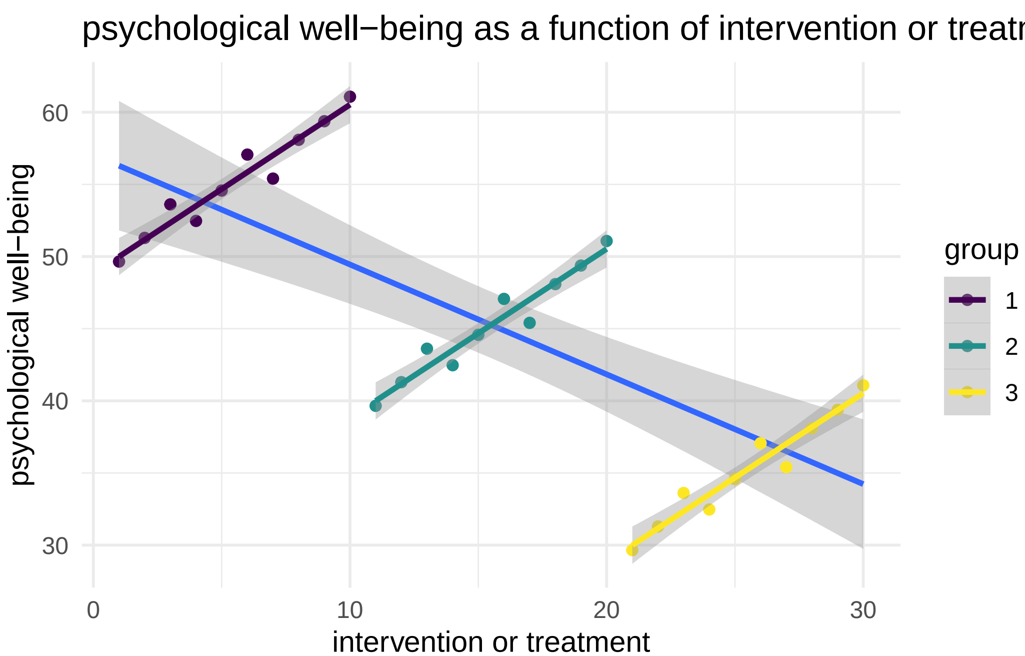
When might a situation like this arise in practice? This is surprisingly difficult to think through.

Imagine that x is a protective factor, or an intervention or treatment. Imagine that y is a desirable outcome, like improved mental health or psychological well being.

Now imagine that people provide more of the protective factor or more of the intervention in communities where there are lower levels of the desirable outcome. If we think about it, this is a very plausible situation.

A Naive Analysis Would Misconstrue The Results

A naive analysis that was unaware of the grouped nature of the data would therefore misconstrue the results, suggesting that the intervention was harmful, when it was in fact helpful.



These data are constructed to provide this kind of extreme example, but it easy to see how multilevel analysis may provide better answers than we would get if we ignored the grouped nature of the data.

4 Two Level Cross Sectional; And Three Level Longitudinal Models

4.1 Cross Sectional Model

4.1.1 Get Data

```
use "simulated_multilevel_data.dta", clear
```

4.1.2 The Equation

$$\begin{aligned} \text{outcome}_{ij} = \beta_0 + \beta_1 \text{parental warmth} + \beta_2 \text{physical punishment} + \\ \beta_3 \text{identity} + \beta_4 \text{intervention} + \beta_5 HDI + \\ u_{0j} + u_{1j} \times \text{parental warmth} + e_{ij} \end{aligned} \tag{4.1}$$

4.1.3 Descriptive Statistics

```
summarize // descriptive statistics
```

```
    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
     country |      3,000        15.5    8.656884          1         30
         HDI |      3,000    64.76667    17.24562         33         87
      family |      3,000        50.5    28.87088          1        100
          id |          0
    identity |      3,000    .4976667    .5000779          0          1
-------------+---------------------------------------------------------
intervention |      3,000    .4843333    .4998378          0          1
physical_p~t |      3,000    2.478667    1.360942          0          5
      warmth |      3,000    3.521667    1.888399          0          7
     outcome |      3,000    52.43327    6.530996   29.60798   74.83553
```

4.1.4 Spaghetti Plot

```
spagplot outcome warmth, id(country) scheme(stcolor)

graph export spagplot1.png, width(1000) replace
```

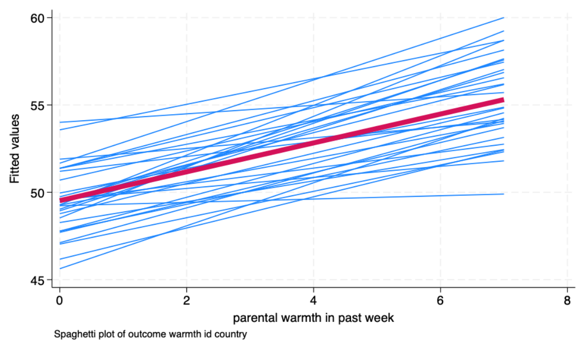


Figure 4.1: Spaghetti Plot of Outcome by Warmth by Country

> Random Slopes and Random Intercepts
>
> This spaghetti plot is illustrating the idea of *random intercepts* and *random slopes*. The multilevel model is estimating *group specific regression lines*, in this case *country specific regression lines*. Each *country specific regression line* has its own intercept and its own slope. The results presented in the multilevel model are the result of *averaging* these estimates into one *overall estimate*.

4.1.5 Unconditional Model

4.1.5.1 Model

```
mixed outcome || country: // unconditional model
```

```
Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:   Log likelihood = -9802.8371
Iteration 1:   Log likelihood = -9802.8371
```

```
Computing standard errors ...

Mixed-effects ML regression                         Number of obs    = 3,000
Group variable: country                             Number of groups =    30
                                                    Obs per group:
                                                                 min =   100
                                                                 avg = 100.0
                                                                 max =   100
                                                    Wald chi2(0)     =     .
Log likelihood = -9802.8371                         Prob > chi2      =     .

------------------------------------------------------------------------------
     outcome | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       _cons |   52.43327   .3451217   151.93   0.000     51.75685     53.1097
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
country: Identity            |
                  var(_cons) |   3.178658   .9226736      1.799552    5.614658
-----------------------------+------------------------------------------------
               var(Residual) |   39.46106   1.024013      37.50421       41.52
------------------------------------------------------------------------------
LR test vs. linear model: chibar2(01) = 166.31        Prob >= chibar2 = 0.0000
```

4.1.5.2 ICC

```
estat icc
```

```
Intraclass correlation

------------------------------------------------------------------------------
                       Level |        ICC   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
                     country |   .0745469   .0201254      .0434963    .1248696
------------------------------------------------------------------------------
```

4.1.6 Conditional Model

```
mixed outcome warmth physical_punishment identity i.intervention HDI || country: warmth // mu

est store crosssectional // store estimates
```

```
Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:  Log likelihood = -9626.6279
Iteration 1:  Log likelihood =  -9626.607
Iteration 2:  Log likelihood =  -9626.607

Computing standard errors ...

Mixed-effects ML regression                     Number of obs     =  3,000
Group variable: country                         Number of groups  =     30
                                                Obs per group:
                                                              min =    100
                                                              avg =  100.0
                                                              max =    100
                                                Wald chi2(5)      = 334.14
Log likelihood =  -9626.607                     Prob > chi2       = 0.0000

------------------------------------------------------------------------------------
           outcome | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
            warmth |   .8345368   .0637213    13.10   0.000     .7096453    .9594282
physical_punishm~t |  -.9916657   .0797906   -12.43   0.000    -1.148052   -.8352791
          identity |  -.3004767   .2170295    -1.38   0.166    -.7258466    .1248933
    1.intervention |   .6396427   .2174519     2.94   0.003     .2134448    1.065841
               HDI |   -.003228   .0199257    -0.16   0.871    -.0422817    .0358256
             _cons |   51.99991   1.371257    37.92   0.000      49.3123    54.68753
------------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
country: Independent         |
                 var(warmth) |   .0227504   .0257784      .0024689    .2096436
                  var(_cons) |   2.963975   .9737647      1.556777    5.643163
```

```
-----------------------------+------------------------------------------------
               var(Residual) |   34.97499   .9097109      33.23668    36.80422
------------------------------------------------------------------------------
LR test vs. linear model: chi2(2) = 205.74                Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

4.2 Longitudinal Model

4.2.1 Get Data

```
use "simulated_multilevel_longitudinal_data.dta", clear
```

4.2.2 The Equation

$$\begin{aligned} \text{outcome}_{ij} = \beta_0 + \beta_1 \text{parental warmth} + \beta_2 \text{physical punishment} + \beta_3 \text{time} + \\ \beta_4 \text{identity}_2 + \beta_5 \text{intervention} + \beta_5 HDI + \\ u_{0j} + u_{1j} \times \text{parental warmth} + \\ v_{0i} + v_{1i} \times t + e_{ij} \end{aligned} \tag{4.2}$$

4.2.3 Descriptive Statistics

```
summarize // descriptive statistics
```

```
    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
     country |      9,000        15.5    8.655922          1         30
         HDI |      9,000    64.76667     17.2437         33         87
      family |      9,000        50.5    28.86767          1        100
          id |          0
    identity |      9,000    .4976667    .5000223          0          1
-------------+---------------------------------------------------------
intervention |      9,000    .4843333    .4997823          0          1
           t |      9,000           2    .8165419          1          3
physical_p~t |      9,000    2.485333    1.373639          0          5
```

```
     warmth |      9,000    3.514222      1.8839          0          7
    outcome |      9,000    53.37768    6.572285   29.60798   79.02199
```

4.2.4 Alternate Plot

```
encode id, generate(idNUMERIC) // numeric version of id

* spagplot outcome t if idNUMERIC <= 10, id(idNUMERIC) scheme(stcolor)

twoway (lfit outcome t) (scatter outcome t) if idNUMERIC <= 10, by(idNUMERIC) scheme(stcolor)

graph export spagplot2.png, width(1000) replace
```

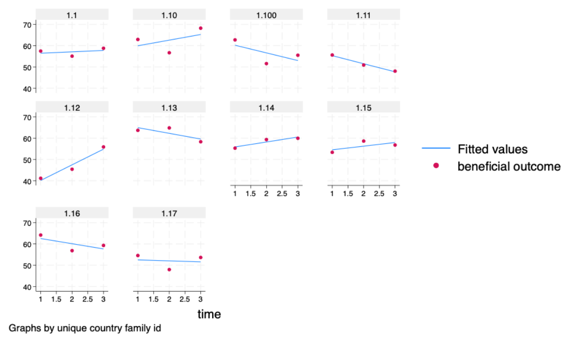


Figure 4.2: Alternate Plot of Outcome by Time by Individual; First 10 Observations

4.2.5 Unconditional Model

4.2.5.1 Model

```
mixed outcome || country: || id: // unconditional model
```

4.2.5.2 ICC

```
estat icc
```

```
Intraclass correlation

------------------------------------------------------------------------------
                       Level |        ICC   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
                     country |   .0748336   .0190847      .0450028    .1219141
                  id|country |   .3462837   .0171461      .3134867    .3806097
------------------------------------------------------------------------------
```

4.2.6 Conditional Model

```
mixed outcome t warmth physical_punishment i.identity i.intervention HDI || country: warmth

est store longitudinal // store estimates
```

```
Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:  Log likelihood =  -28523.49
Iteration 1:  Log likelihood = -28499.953
Iteration 2:  Log likelihood = -28499.735
Iteration 3:  Log likelihood = -28499.604
Iteration 4:  Log likelihood = -28499.603

Computing standard errors ...

Mixed-effects ML regression                         Number of obs =   9,000

        Grouping information
        -------------------------------------------------------------
                        |     No. of       Observations per group
         Group variable |     groups    Minimum    Average    Maximum
        ----------------+--------------------------------------------
                country |         30        300      300.0        300
                     id |      3,000          3        3.0          3
        -------------------------------------------------------------

                                                    Wald chi2(6)  = 1096.15
Log likelihood = -28499.603                         Prob > chi2   =  0.0000
```

```
------------------------------------------------------------------------------
             outcome | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
--------------------+---------------------------------------------------------------
                   t |    .943864   .0658716    14.33   0.000      .814758    1.07297
              warmth |   .9134959   .0423732    21.56   0.000     .8304461   .9965458
physical_punishm~t |  -1.007897   .0497622   -20.25   0.000    -1.105429  -.9103647
          1.identity |  -.1276926   .1515835    -0.84   0.400    -.4247908   .1694056
      1.intervention |   .8589966   .1519094     5.65   0.000     .5612596   1.156734
                 HDI |  -.0005657   .0196437    -0.03   0.977    -.0390666   .0379352
               _cons |   50.46724   1.338318    37.71   0.000     47.84418   53.09029
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
country: Independent         |
                var(warmth) |   .0107585   .0127845      .0010477    .1104712
                 var(_cons) |   3.167087   .9146767      1.798155    5.578185
-----------------------------+------------------------------------------------
id: Independent              |
                     var(t) |   5.69e-10   1.29e-07      1.1e-202    3.0e+183
                 var(_cons) |   8.387268   .4724189      7.510624    9.366236
-----------------------------+------------------------------------------------
              var(Residual) |   26.02734   .4753703      25.11211    26.97592
------------------------------------------------------------------------------
LR test vs. linear model: chi2(4) = 1247.03              Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

4.2.7 Nice Table of Results

```
etable, estimates(crosssectional longitudinal) ///
showstars showstarsnote /// show stars and note
column(estimate) // column is modelname
```

```
                                     crosssectional longitudinal
----------------------------------------------------------------
parental warmth in past week                 0.835 **      0.913 **
                                           (0.064)       (0.042)
physical punishment in past week            -0.992 **     -1.008 **
```

	(0.080)	(0.050)
hypothetical identity group variable	-0.300	
	(0.217)	
recieved intervention		
1	0.640 **	0.859 **
	(0.217)	(0.152)
Human Development Index	-0.003	-0.001
	(0.020)	(0.020)
time		0.944 **
		(0.066)
hypothetical identity group variable		
1		-0.128
		(0.152)
Intercept	52.000 **	50.467 **
	(1.371)	(1.338)
var(warmth)	0.023	0.011
	(0.026)	(0.013)
var(_cons)	2.964	3.167
	(0.974)	(0.915)
var(e)	34.975	26.027
	(0.910)	(0.475)
var(_cons)		8.387
		(0.472)
var(t)		0.000
		(0.000)
Number of observations	3000	9000

** p<.01, * p<.05

4.3 QUESTIONS???

5 Cross-Classified Models

5.1 Introduction

A two level multilevel model imagines that *Level 1* units are nested in *Level 2* units. A three level multilevel model imagines that *Level 1* units are nested in *Level 2* units, which are in turn nested in *Level 3* units.

A cross-classified model imagines that the nesting is not hierarchical, but rather that there are two sets of clusters or nestings which overlap, but are not hierarchical.

5.2 Get Data

```
use "simulated_multilevel_longitudinal_data.dta", clear
```

5.3 Cross Classified Model

We can treat these random effects as being *cross classified*.

This might be useful if we had data where individuals lived in different countries at different times.

However, because `id` is in fact nested inside `country`, in this case, estimating the random effects as cross classified will be more time consuming, but will give us equivalent results to a three level model.

5.3.1 Standard (Less Computationally Efficient) Syntax

The below syntax will take a very long time to run with the full sample, and thus we have commented it out.

```
* mixed outcome t warmth physical_punishment || _all: R.country || _all: R.id

* est store crossed1
```

The documentation notes that we can use a *much* more computationally efficient version of the above command, which is what we do in these notes. The user can verify that both versions of the command will produce equivalent results.

In fact, at the end of handout we verify the similarity of both sets of syntax using a random sample.

5.3.2 Cross Classified With Computationally Efficient Syntax

```
mixed outcome t warmth physical_punishment || _all: R.country || id:

est store crossed2 // store crossed effects result
```

```
Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:   Log likelihood = -28516.314
Iteration 1:   Log likelihood = -28516.277
Iteration 2:   Log likelihood = -28516.277

Computing standard errors ...

Mixed-effects ML regression                          Number of obs =   9,000

        Grouping information
        -------------------------------------------------------------
                        |     No. of       Observations per group
         Group variable |     groups    Minimum    Average    Maximum
        ----------------+--------------------------------------------
                   _all |          1      9,000    9,000.0      9,000
                     id |      3,000          3        3.0          3
        -------------------------------------------------------------

                                                     Wald chi2(3)  = 1168.69
Log likelihood = -28516.277                          Prob > chi2   =  0.0000
```

```
------------------------------------------------------------------------------------
           outcome | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
                 t |   .9434605    .065866    14.32   0.000     .8143654    1.072556
            warmth |   .9053924   .0380439    23.80   0.000     .8308277    .9799572
physical_punishm~t |  -1.014385   .0499354   -20.31   0.000    -1.112257   -.916514
             _cons |    50.8301   .4123007   123.28   0.000       50.022    51.63819
------------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
_all: Identity               |
              var(R.country) |   3.429974    .930313      2.015668    5.836634
-----------------------------+------------------------------------------------
id: Identity                 |
                  var(_cons) |   8.608872   .4757699      7.725107     9.59374
-----------------------------+------------------------------------------------
               var(Residual) |   26.02862   .4752444      25.11363    26.97695
------------------------------------------------------------------------------
LR test vs. linear model: chi2(2) = 1260.84           Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

5.4 Three Level Model

```
mixed outcome t warmth physical_punishment || country: || id:  // 3 level w/ random intercep

est store threelevel // store random intercept model
```

```
Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:  Log likelihood = -28516.314
Iteration 1:  Log likelihood = -28516.277
Iteration 2:  Log likelihood = -28516.277

Computing standard errors ...
```

```
Mixed-effects ML regression                        Number of obs =   9,000

        Grouping information
        -------------------------------------------------------------
                        |     No. of       Observations per group
         Group variable |     groups    Minimum    Average    Maximum
        ----------------+--------------------------------------------
                country |         30        300      300.0        300
                     id |      3,000          3        3.0          3
        -------------------------------------------------------------

                                                   Wald chi2(3)  = 1168.69
Log likelihood = -28516.277                        Prob > chi2   =  0.0000

------------------------------------------------------------------------------------
           outcome | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------------+----------------------------------------------------------------
                 t |   .9434605    .065866    14.32   0.000     .8143654    1.072556
            warmth |   .9053924   .0380439    23.80   0.000     .8308277    .9799572
physical_punishm~t |  -1.014385   .0499354   -20.31   0.000    -1.112257   -.916514
             _cons |    50.8301   .4123007   123.28   0.000       50.022    51.63819
------------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects parameters  |   Estimate   Std. err.     [95% conf. interval]
-----------------------------+------------------------------------------------
country: Identity            |
                  var(_cons) |   3.429974    .930313      2.015668    5.836634
-----------------------------+------------------------------------------------
id: Identity                 |
                  var(_cons) |   8.608872   .4757699      7.725107     9.59374
-----------------------------+------------------------------------------------
               var(Residual) |   26.02862   .4752444      25.11363    26.97695
------------------------------------------------------------------------------
LR test vs. linear model: chi2(2) = 1260.84           Prob > chi2 = 0.0000

Note: LR test is conservative and provided only for reference.
```

5.5 Nice Table of Results of Three Level and Cross Classified Model

```
etable, estimates(threelevel crossed2), ///
showstars showstarsnote /// show stars and note
column(estimate) // column is modelname
```

```
invalid 'showstars'
r(198);

r(198);
```

5.6 Verification of Syntax Equivalence for Cross Classified Model

```
keep if family <= 5 // random sample of families

quietly mixed outcome t warmth physical_punishment || _all: R.country || _all: R.id

est store crossed1A // less efficient syntax

quietly mixed outcome t warmth physical_punishment || _all: R.country || id:

est store crossed2A // more efficient syntax

etable, estimates(crossed1A crossed2A) ///
showstars showstarsnote /// show stars and note
column(estimate) // column is modelname
```

```
(8,550 observations deleted)




-------------------------------------------------------
                                   crossed1A  crossed2A
```

time	0.745 **	0.745 **
	(0.281)	(0.281)
parental warmth in past week	0.871 **	0.871 **
	(0.160)	(0.160)
physical punishment in past week	-1.262 **	-1.262 **
	(0.206)	(0.206)
Intercept	51.755 **	51.755 **
	(1.009)	(1.009)
var(R_country)	2.245	2.245
	(1.319)	(1.319)
var(R_id)	5.425	
	(1.843)	
var(e)	23.638	23.638
	(1.933)	(1.933)
var(_cons)		5.425
		(1.843)
Number of observations	450	450

** p<.01, * p<.05

5.7 QUESTIONS???

References

Bland, J M, and D G Altman. 1994. "Statistics Notes: Correlation, Regression, and Repeated Data." *BMJ* 308 (April): 896. https://doi.org/10.1136/bmj.308.6933.896.

Diez Roux, Ana. 2003. "Potentialities and Limitations of Multilevel Analysis in Public Health and Epidemiology." In *Methodology and Epistemology of Multilevel Analysis: Approaches from Different Social Sciences*, edited by Daniel Courgeau, 93–119. Kluwer Academic Publishers.

Firebaugh, Glen. 2001. "Ecological Fallacy, Statistics Of." In *International Encyclopedia of the Social & Behavioral Sciences*, edited by Neil J. Smelser and Paul B. Baltes, 4023–26. Oxford: Pergamon. https://doi.org/https://doi.org/10.1016/B0-08-043076-7/00765-8.

Gelman, Andrew, Boris Shor, Joseph Bafumi, and David Park. 2007. "Rich State, Poor State, Red State, Blue State: What's the Matter with Connecticut?" *Quarterly Journal of Political Science* 2 (November): 345–67. https://doi.org/10.2139/ssrn.1010426.

Nieuwenhuis, Rense. 2015. "Association, Aggregation, and Paradoxes: On the Positive Correlation Between Fertility and Women's Employment." *Demographic Research* 32 (March). https://www.demographic-research.org/volumes/vol32/23/.

Oliver, Mary, and Krista Tippett. 2015. "Mary Oliver: 'I Got Saved by the Beauty of the World.'" The On Being Project. https://onbeing.org/programs/mary-oliver-i-got-saved-by-the-beauty-of-the-world/.

License and Citation

License

Citation

For attribution, please cite as:

Grogan-Kaylor (2025). *Multilevel Workshop*. Retrieved from https://agrogan1.github.io/multilevel-workshop

BibTeX Citation

```
@book{,
   author = {Andrew Grogan-Kaylor},
   city = {Ann Arbor, MI},
   title = {Multilevel Workshop},
   url = {https://agrogan1.github.io/multilevel-workshop},
   year = {2025},
}
```