# Multilevel Multilingual

Multilevel Models in Stata, R and Julia

Andrew Grogan-Kaylor

2024 - 09 - 24

# Table of contents

1	Mul	tilevel Multilingual
	1.1	Introduction
	1.2	The Data
	1.3	An Introduction To Equations and Syntax
		1.3.1 Stata
		1.3.2 R
		1.3.3 Julia
2	Stat	istical Workflows 8
	2.1	Statistical Software Is Best Run Using a Script
	2.2	Scripts
	2.3	Script Flow
	2.4	Good Statistical Workflows Allow Multiple Statistical Packages
	2.5	Good Statistical Workflows Require Safe Workspaces
	2.6	Good Statistical Workflows Require Patience And Can Be Psychologically De-
		manding
	2.7	Good Statistical Workflows Often Allow Multiple Principled Ways Forward 10
3	Stor	ring Statistical Data
	3.1	Spreadsheets
	3.2	Data in Statistical Format
		3.2.1 Describe The Data
		3.2.2 Descriptive Statistics
		3.2.3 Graph
		3.2.4 List Out A Sample Of The Data
	3.3	Data In Spreadsheet Format
		3.3.1 Describe The Data
		3.3.2 Descriptive Statistics
		3.3.3 Graph
	3.4	A Few Final Issues
	3.5	File Organization
4	Des	criptive Statistics 21
	4.1	Descriptive Statistics
		4.1.1. Ctoto

		4.1.2 R
		4.1.3 Julia
	4.2	Interpretation
5	Unc	onditional Models 25
	5.1	Two Level Model
		5.1.1 The Equation
		5.1.2 Run Models
		5.1.3 Interpretation
	5.2	Three Level Model
		5.2.1 The Equation
		5.2.2 Run Models
		5.2.3 Interpretation
6	Cros	ss Sectional Multilevel Models 34
	6.1	The Equation
	6.2	Correlated and Uncorrelated Random Effects
	6.3	Run The Models
		6.3.1 Stata
		6.3.2 R
		6.3.3 Julia
	6.4	Interpretation
7	Lon	gitudinal Multilevel Models 41
•	7.1	The Data
	7.2	The Equation
	7.2	Growth Trajectories
	7.4	Run The Models
	1.4	7.4.1 Stata
		7.4.2 R
		7.4.3 Julia
	7.5	Interpretation
_		•
8		tilevel Logistic Regression 54
	8.1	The Data
	8.2	The Equation
		8.2.1 Stata
	0.0	8.2.2 R
	8.3	Run Models
		8.3.1 Stata
		832 B 57

9	# M	odels With Three or More Levels and Cross-Classified Models 6	60
	9.1	Introduction	30
	9.2	Three Or More Levels	30
		9.2.1 The Data	30
		9.2.2 The Equation	30
		9.2.3 Run The Models	31
		9.2.4 Interpretation	39
	9.3	Cross-Classified Models	39
		9.3.1 The Data	39
		9.3.2 The Equation	39
		9.3.3 Run The Models	70
		9.3.4 Interpretation	76
10	Resh	aping Data 7	77
	10.1	Introduction	77
	10.2	Data in Wide Format	77
	10.3	Data Management	78
	10.4	Reshaping Data From Wide To Long	78
		10.4.1 Stata	79
		10.4.2 R	30
	10.5	Data in Long Format	31
11	Δσσι	egating Data	33
	, ,00.	-0	33
			33
Re	feren	ces 8	35

# **List of Tables**

1.1	Software for Multilevel Modeling	4
1.2	Sample of Simulated Multilevel Data	6
1.2	Table continues below	6
1.3	Sample of Simulated Multilevel Data	6
3.1	Example Data As Stored in A Spreadsheet	11
3.1	Table continues below	11
3.2	Example Data As Stored in A Spreadsheet	12
3.3	Variable Labels	12
3.4	Descriptive Statistics	13
3.4	Table continues below	13
3.5	Descriptive Statistics	13
3.5	Table continues below	13
3.6	Descriptive Statistics	13
3.6	Descriptive Statistics	14
3.7	Sample of Data	15
3.7	Table continues below	15
3.8	Sample of Data	15
3.9	Example Data As Stored in A Spreadsheet	15
3.9	Example Data As Stored in A Spreadsheet	16
3.10	Example Data As Stored in A Spreadsheet	16
3.10	Table continues below	16
3.11	Example Data As Stored in A Spreadsheet	16
3.12	Descriptive Statistics	17
3.12	Table continues below	17
	Descriptive Statistics	17
	Table continues below	17
	Descriptive Statistics	17
3.15	A Spreadsheet Table With Problematic Organization	19
6.1	Correlated and Uncorrelated Random Effects	35
7.1	Slope and Intercept for Each Group	42
10.1	Data in Wide Format	77
10.1	Table continues below	77

10.2	Data in Wide Format		•						•									78
10.2	Table continues below																	78
10.3	Data in Wide Format																	78
10.4	Data in Long Format .																	81
10.4	Table continues below																	81
10.5	Data in Long Format.																	82

# 1 Multilevel Multilingual

"This curious world which we inhabit is more wonderful than it is convenient..." (Thoreau, 1975)

"Mathematics is my secret. My secret weakness. I feel like a stubborn, helpless fool in the middle of a problem. Trapped and crazed. Also, thrilled." (Schanen, 2021)

#### 1.1 Introduction

Below, I describe the use of Stata (StataCorp, 2023), R (Bates et al., 2015; R Core Team, 2023), and Julia (Bates, 2024; Bezanson et al., 2017) to estimate multilevel models.

All of these software packages can estimate multilevel models and can visualize relationships in the data. However, there are substantial differences between the different packages: Stata is proprietary for cost software, which is very well documented and very intuitive. While it costs money to purchase Stata, the price is often very reasonable for academic and educational use. R is free open source software which is less intuitive, but there are many excellent resources for learning R. There is often a cost associated with purchasing books and other materials for learning R, which sometimes feels like it offsets the fact that R is free. Julia is newer open source software, also free, and ostensibly much faster than either Stata or R, which may be an important advantage when running multilevel models with very large data sets. At this point in time, both Stata and R feel much more stable than Julia which is still evolving software.

While any of these software packages can be used for learning and estimating multilevel models, I will offer my own opinion—based upon 15 years of teaching multilevel models at the doctoral level—that Stata offers the quickest pathway for learning the basic and advanced uses of multilevel models. I also believe the intuitive nature of Stata syntax contributes to accurate and replicable work in this area.

Table 1.1: Software for Multilevel Modeling

Software	Cost	Ease of Use
Stata	some cost	learning curve, but very intuitive for both multilevel modeling and graphing.

Software	Cost	Ease of Use
R	free	learning curve: intuitive for multilevel modeling; but steeper learning curve for graphing (ggplot).
Julia	free	steep learning curve in general: steep learning curve for multilevel modeling; and very steep learning curve for graphing. Graphics libraries are very much under development and in flux.

#### Results Will Vary Somewhat

Estimating multilevel models is a complex endeavor. The software details of how this is accomplished are beyond the purview of this book. Suffice it to say that across different software packages there will be differences in estimation routines, resulting in some numerical differences in the results provided by different software packages. Substantively speaking, however, results should agree across software.

#### Multi-Line Commands

Sometimes I have written commands out over multiple lines. I have done this for especially long commands, but have also sometimes done this simply for the sake of clarity. The different software packages have different approaches to multi-line commands.

- 1. By default, Stata ends a command at the end of a line. If you are going to write a multi-line command you should use the /// line continuation characters.
- 2. R is the software that most naturally can be written using multiple lines, as R commands are usually clearly encased in parentheses (()) or continued with + signs.
- 3. Like Stata, Julia expects commands to end at the end of a line. If you are going to write a mult-line command, all commands except for the last line should end in a character that clearly indicates continuation, like a + sign. An alternative is to encase the entire Julia command in an outer set of parentheses (()).

#### Running Statistical Packages in Quarto

I used Quarto (Allaire et al., 2024) (https://quarto.org/) to create this Appendix. Quarto is a programming and publishing environment that can run multiple programming languages, including Stata, R and Julia, and that can write to multiple output formats including HTML, PDF, and MS Word. To run Stata, I used the Statamarkdown library (Hemken, 2023) in R to connect Stata to Quarto. Quarto has a built in connection to R, and runs R without issue. To run Julia, I used the  $\tt JuliaCall$  library (Li, 2019) in R to connect Quarto to Julia.

Of course, each of these programs can be run by itself, if you have them installed on your computer.

#### 1.2 The Data

#### i Datasets

The examples use the simulated\_multilevel\_data.dta and simulated\_multilevel\_longitudinal\_data.dta files.

Here is a direct link to download the cross-sectional data.

Here is a direct link to download the longitudinal data.

Table 1.2: Sample of Simulated Multilevel Data

Table 1.2: Table continues below

country	HDI	family	id	identity	intervention	physical_punishment
1	69	1	1.1	1	0	3
1	69	2	1.2	1	1	2
1	69	3	1.3	0	1	3
1	69	4	1.4	1	0	0
1	69	5	1.5	1	0	4
1	69	6	1.6	0	1	5

Table 1.3: Sample of Simulated Multilevel Data

warmth	outcome
3	57.47
1	50.1
2	52.92
5	60.17
4	55.05
3	49.81

## 1.3 An Introduction To Equations and Syntax

To explain statistical syntax for each software, I consider the general case of a multilevel model with dependent variable y, independent variables x and z, clustering variable group, and a random slope for x. i is the index for the person, while j is the index for the group.

$$y = \beta_0 + \beta_1 x_{ij} + \beta_2 z_{ij} + u_{0j} + u_{1j} \times x_{ij} + e_{ij}$$
(1.1)

#### 1.3.1 Stata

In Stata mixed, the syntax for a multilevel model of the form described in Equation 1.1 is:

```
mixed y x z || group: x
```

#### 1.3.2 R

In R lme4, the syntax for a multilevel model of the form described in Equation 1.1 is:

```
library(lme4)

lmer(y ~ x + z + (1 + x || group), data = ...)
```

#### 1.3.3 Julia

In Julia MixedModels, the syntax for a multilevel model of the form described in Equation 1.1 is:

```
using MixedModels
fit(MixedModel, @formula(y ~ x + z + (1 + x | group)), data)
```

# 2 Statistical Workflows

#### 2.1 Statistical Software Is Best Run Using a Script

Many statistical workflows—whatever the statistical package being used—follow the same conceptual pattern.



Figure 2.1: A Common Statistical Workflow

Increasingly, we want to think about workflows that are

- documentable, transparent, and auditable: We have a record of what we did if we want to double check our work, clarify a result, or develop a new project with a similar process. We, or others, can find the inevitable errors in our work, and correct them.
- replicable: Others can replicate our findings with the same or new data.
- scalable: We are developing a process that can be as easily used with thousands or millions of rows of data as it can with ten rows of data. We are developing a process that can be easily repeated if we are constantly getting new or updated data, e.g. getting new data every week, or every month.

#### 2.2 Scripts

For most statistical workflows, we will often want to write a script or code. Data analysis scripts can be stored in a Quarto document (Allaire et al., 2024) as they are in this Appendix, or every statistical package has its own unique format for storing scripts as a text file: in Stata, scripts are stored in .do files; in R, scripts are stored in .R files, and in Julia, scripts are stored in .jl files.

#### 2.3 Script Flow

A good practice when writing a script, is to have a script that begins with the raw data, moves through any necessary re-coding or cleaning of the data, generates descriptive statistics, generates the appropriate multivariate results, and then generates any necessary visualizations.

#### 2.4 Good Statistical Workflows Allow Multiple Statistical Packages

While this Appendix focuses on the use of each individual statistical package on its own, it is certainly possible to use multiple statistical packages as part of the same workflow. For example, one might employ Stata to carry out data management tasks, and then possibly use R to run a multilevel model with a more complicated multilevel structure, such as a cross-classified model, or Julia to more quickly run a model with a large data.

### 2.5 Good Statistical Workflows Require Safe Workspaces

It is also very important to be aware that good complex workflows are highly iterative and highly collaborative. Good complex workflows require a safe workspace in which team members feel free to admit their own errors, and help with others' mistakes in a non-judgmental fashion. Such a safe environment is necessary to build an environment where the overall error rate is low.

# 2.6 Good Statistical Workflows Require Patience And Can Be Psychologically Demanding

Developing a good documented and auditable workflow that is implemented in code requires a lot of patience, and often, **many iterations**. Working through these many iterations can be psychologically demanding. It is important to remember that careful attention to getting

the details right early in the research process, while sometimes tiring and frustrating, will pay large dividends later on when the research is reviewed, presented, published and read.

# 2.7 Good Statistical Workflows Often Allow Multiple Principled Ways Forward

One of my most recent ideas about statistical workflows is that there are certainly *wrong* decisions that one can make with data.

For example, I would not want to write the paper that says that smoking prevents lung cancer, nor would I want to write a paper saying physical punishment is good for children.

That being said, I think there are often multiple principled ways forward.

Often the key is not so much to make the 100% correct decision, but to make one of *several possible principled decisions*.

Then after making a *principled decision*, one is *transparent* and *thorough* about describing the decision that one made.

For example, in implementing a multilevel analysis, I would have many choices: I could estimate only a random intercept; estimate one or more random slopes; or estimate all possible random slopes. The random effects could be correlated or uncorrelated. I could estimate only main effects, or could estimate interactions of several variables. Each of these would be a different, yet principled, approach to analyzing the data.

In science and statistics, we often want an answer that provides one clear direction. Instead, I'm increasingly convinced that the best science (and teaching!) often involves engaging in open discussion about the multiple possible alternatives, and then choosing one principled solution, and being transparent about its implementation.

# 3 Storing Statistical Data

#### 3.1 Spreadsheets

Spreadsheets are sometimes used to collect and store data. Spreadsheets may sometimes be used because they are the only program that some individuals or agencies have for storing data. Spreadsheet programs may also be used because spreadsheets can be very intuitive and easy ways of managing small amounts of data.

However, spreadsheets may be problematic as an ultimate data storage solution for a number of reasons detailed below, especially as data sets grow in size. Notably, statistical programs like Stata, R, or Julia can all store additional information with each variable such as: a variable label, describing the contents of the variable, or the survey question that resulted in the variable; and a value label, which attaches qualitative information to each possible value of the response.

Spreadsheets do not generally contain this extra information about each variable, or column of data, which may lead to errors in working with quantitative information.

⚠ If Data Are Stored In A Spreadsheet, Variable Names Should Be Limited To a Single Row of the Spreadsheet

If spreadsheets are used to store data, the first row of the data should be used to list the variable names, as is seen in the example below. Rows other than the first row should not contain additional information about the variables, but should only contain data.

Table 3.1: Example Data As Stored in A Spreadsheet

country	HDI	family	id	identity	intervention	physical_punishment
1	69	1	1.1	1	0	3
1	69	2	1.2	1	1	2
1	69	3	1.3	0	1	3
1	69	4	1.4	1	0	0
1	69	5	1.5	1	0	4
1	69	6	1.6	0	1	5

Table 3.2: Example Data As Stored in A Spreadsheet

warmth	outcome
3	57.47
1	50.1
2	52.92
5	60.17
4	55.05
3	49.81

### 3.2 Data in Statistical Format

I load the data from a statistical program.

#### 3.2.1 Describe The Data

Notice how a description of the data contains information that helps us to understand the variables.

Table 3.3: Variable Labels

pos	variable	label
1	country	country id
2	HDI	Human Development Index
3	family	family id
4	$\operatorname{id}$	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

#### 3.2.2 Descriptive Statistics

#### 💡 Variable Labels and Value Labels Help Us Understand Our Data

Notice how the descriptive statistics and graph are informative in that they contain information on the variable label and value label. These help us to get an intuitive sense of the information in the data. We see this information when we list out the data as well.

Table 3.4: Descriptive Statistics

Table 3.4: Table continues below

country	HDI	family	id
1:100	Min. :33.00	Min.: 1.00	Length:3000
2:100	1st Qu.:53.00	1st Qu.: 25.75	Class:character
3:100	Median $:70.00$	Median: 50.50	Mode :character
4:100	Mean $:64.77$	Mean: 50.50	NA
5:100	3rd Qu.:81.00	3rd Qu.: 75.25	NA
6:100	Max. :87.00	Max. $:100.00$	NA
(Other):2400	NA	NA	NA

Table 3.5: Descriptive Statistics

Table 3.5: Table continues below

identity	intervention	physical_punishment	warmth
Identity B:1507	no intervention:1547	Min. :0.000	Min. :0.000
Identity A:1493	intervention:1453	1st Qu.:2.000	1st Qu.:2.000
NA	NA	Median $:2.000$	Median : 4.000
NA	NA	Mean $:2.479$	Mean $:3.522$
NA	NA	3rd Qu.:3.000	3rd Qu.:5.000
NA	NA	Max. $:5.000$	Max. $:7.000$
NA	NA	NA	NA

Table 3.6: Descriptive Statistics

outcome				
Min. :29.61				
1st Qu.:48.02				

Table 3.6: Descriptive Statistics

### 3.2.3 Graph

# Beneficial Outcome by Intervention Participation

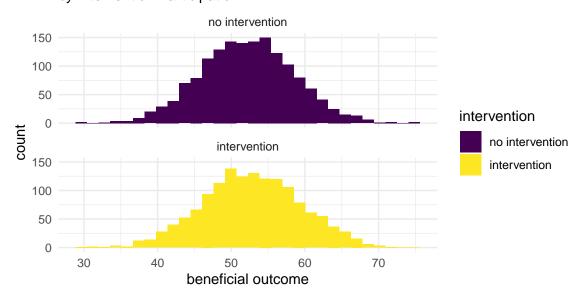


Figure 3.1: Graph from Data Stored in Statistical Software

#### 3.2.4 List Out A Sample Of The Data

Table 3.7: Sample of Data

Table 3.7: Table continues below

country	HDI	family	id	identity	intervention
1	69	1	1.1	Identity A	no intervention
1	69	2	1.2	Identity A	intervention
1	69	3	1.3	Identity B	intervention
1	69	4	1.4	Identity A	no intervention
1	69	5	1.5	Identity A	no intervention
1	69	6	1.6	Identity B	intervention

Table 3.8: Sample of Data

physical_punishment	warmth	outcome
3	3	57.47
2	1	50.1
3	2	52.92
0	5	60.17
4	4	55.05
5	3	49.81

## 3.3 Data In Spreadsheet Format

I now import the spreadsheet data file. I use the first row of data as variable names.

We see right away that the data are less informative.

#### 3.3.1 Describe The Data

Notice how a description of the data no longer contains much of the information that helped us to understand the variables.

Table 3.9: Example Data As Stored in A Spreadsheet

pos	variable	label
1	country	NA
2	HDI	NA
3	family	NA

Table 3.9: Example Data As Stored in A Spreadsheet

pos	variable	label
4	id	NA
5	identity	NA
6	intervention	NA
7	physical_punishment	NA
8	warmth	NA
9	outcome	NA

Table 3.10: Example Data As Stored in A Spreadsheet

Table 3.10: Table continues below

country	HDI	family	id	identity	intervention	physical_punishment
1	69	1	1.1	1	0	3
1	69	2	1.2	1	1	2
1	69	3	1.3	0	1	3
1	69	4	1.4	1	0	0
1	69	5	1.5	1	0	4
1	69	6	1.6	0	1	5

Table 3.11: Example Data As Stored in A Spreadsheet

warmth	outcome
3	57.47
1	50.1
2	52.92
5	60.17
4	55.05
3	49.81

#### ⚠ Warning

Adding this valuable information back into the data set may take a great deal of extra effort.

#### 3.3.2 Descriptive Statistics

Notice here how the descriptive statistics and graph are much less informative. For example, it is now not immediately clear what the values of identity or intervention represent. The information on variable labels and value labels will have to be added back into the data when preparing a final product for dissemination.

Table 3.12: Descriptive Statistics

Table 3.12: Table continues below

country	HDI	family	id
Min. : 1.0	Min. :33.00	Min.: 1.00	Length:3000
1st Qu.: 8.0	1st Qu.:53.00	1st Qu.: 25.75	Class:character
Median:15.5	Median $:70.00$	Median: 50.50	Mode :character
Mean: 15.5	Mean $:64.77$	Mean: 50.50	NA
3rd Qu.:23.0	3rd Qu.:81.00	3rd Qu.: 75.25	NA
Max. $:30.0$	Max. :87.00	Max. :100.00	NA

Table 3.13: Descriptive Statistics

Table 3.13: Table continues below

identity	intervention	physical_punishment	warmth
Min. :0.0000	Min. :0.0000	Min. :0.000	Min. :0.000
1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:2.000	1st Qu.:2.000
Median: 0.0000	Median : 0.0000	Median $:2.000$	Median $:4.000$
Mean $:0.4977$	Mean $:0.4843$	Mean $:2.479$	Mean $:3.522$
3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:3.000	3rd Qu.:5.000
Max. $:1.0000$	Max. $:1.0000$	Max. :5.000	Max. :7.000

Table 3.14: Descriptive Statistics

outcome
Min. :29.61
1st Qu.:48.02
Median: $52.45$
Mean $:52.43$
3rd Qu.:56.86
Max. :74.84

#### 3.3.3 Graph

While the graph has an informative title, as well as informative axis labels, a crucial piece of information is missing: what each status of the intervention represents.

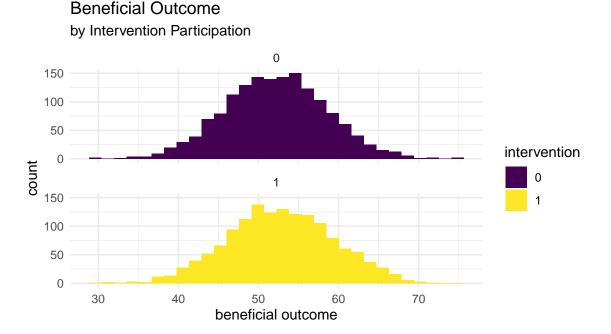


Figure 3.2: Graph from Data Stored in Spreadsheet

#### 3.4 A Few Final Issues

Notice, finally, how spreadsheets don't enforce the idea of whether variables are *numeric*, or *text*, and so would allow storage of different types of information in the same column. Relatedly, *numeric* variables may be improperly stored as *text*, often necessitating recoding before graphical or statistical procedures can be employed.

Second, a spreadsheet would allow some of your columns to have the same name, which might make data difficult to work with in other software.

Lastly, spreadsheets do not enforce the idea that the data have a *structure* wherein the *column* header is a variable name, while the other rows are data.

Table 3.15: A Spreadsheet Table With Problematic Organization

X	У	verylongvariablename	verylongvariablename
100	1	Smith	20
200	2	30	NA
not applicable	X	yes some other random information	60

# 3.5 File Organization

Files for all of your work should not be stored all together in downloads. Ideally, you should have a specific set of folders for your work. Each project, should be stored in its own individual folder. Ideally, each project folder would have a separate sub-folder for separate aspects of the project such as data, code or syntax, and various outputs.

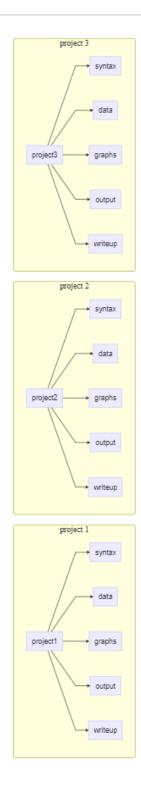


Figure 3.3: A Hypothetical Set of Folders and Subfolders

# **4 Descriptive Statistics**

# 4.1 Descriptive Statistics

#### 4.1.1 Stata

```
use simulated_multilevel_data.dta // use data
```

We use summarize for *continuous* variables, and tabulate for *categorical* variables.

```
summarize outcome warmth physical_punishment HDI
tabulate identity
tabulate intervention
```

Variable		0bs	Mean	Std. dev.	Min	Max
outcome	1	3,000	52.43327	6.530996	29.60798	74.83553
warmth		3,000	3.521667	1.888399	0	7
physical_p~t		3,000	2.478667	1.360942	0	5
HDI		3,000	64.76667	17.24562	33	87

hypothetica	1			
1 identity				
group				
variable		Freq.	Percent	Cum.
	-+			
0	1	1,507	50.23	50.23
1	1	1,493	49.77	100.00
	-+			
Total	1	3,000	100.00	

recieved				
interventio				
n		Freq.	Percent	Cum.
0		1,547	51.57	51.57
1		1,453	48.43	100.00
Total	 	3,000	100.00	

#### 4.1.2 R

```
library(haven) # read data in Stata format

df <- read_dta("simulated_multilevel_data.dta")</pre>
```

R's descriptive statistics functions rely heavily on whether a variable is a *numeric* variable, or a *factor* variable. Below, I convert two variables to factors (factor) before using summary<sup>1</sup> to generate descriptive statistics.

```
df$country <- factor(df$country)

df$identity <- factor(df$identity)

df$intervention <- factor(df$intervention)

summary(df)</pre>
```

country	HDI	family	id	identity
1 : 100	Min. :33.00	Min. : 1.00	Length:3000	0:1507
2 : 100	1st Qu.:53.00	1st Qu.: 25.75	Class :character	1:1493
3 : 100	Median :70.00	Median : 50.50	Mode :character	
4 : 100	Mean :64.77	Mean : 50.50		
5 : 100	3rd Qu.:81.00	3rd Qu.: 75.25		
6 : 100	Max. :87.00	Max. :100.00		
(Other):2400				
intervention	physical_punishme	ent warmth	outcome	
0:1547	Min. :0.000	Min. :0.000	Min. :29.61	
1:1453	1st Qu.:2.000	1st Qu.:2.000	1st Qu.:48.02	

<sup>&</sup>lt;sup>1</sup>skimr is an excellent new alternative library for generating descriptive statistics in R.

```
Median :2.000Median :4.000Median :52.45Mean :2.479Mean :3.522Mean :52.433rd Qu.:3.0003rd Qu.:5.0003rd Qu.:56.86Max. :5.000Max. :7.000Max. :74.84
```

#### 4.1.3 Julia

```
using Tables, MixedModels, MixedModelsExtras, StatFiles, DataFrames, CategoricalArrays, Data
df = DataFrame(load("simulated_multilevel_data.dta"))
```

Similarly to R, Julia relies on the idea of variable type. I use transform to convert the appropriate variables to categorical variables.

```
@transform!(df, :country = categorical(:country))
@transform!(df, :identity = categorical(:identity))
@transform!(df, :intervention = categorical(:intervention))
```

```
describe(df) # descriptive statistics
```

9×7 Da	ataFrame						
Row	variable	mean	min	median	max	nmissing	eltype
	Symbol	Union	Any	Union	Any	Int64	Union
1	country		1.0		30.0	0	Union{
2	HDI	64.7667	33.0	70.0	87.0	0	Union{
3	family	50.5	1.0	50.5	100.0	0	Union{
4	id		1.1		9.99	0	Union{
5	identity		0.0		1.0	0	Union{
6	intervention		0.0		1.0	0	Union{
7	physical_punishment	2.47867	0.0	2.0	5.0	0	Union{
8	warmth	3.52167	0.0	4.0	7.0	0	Union{
9	outcome	52.4333	29.608	52.449	74.8355	0	Union{
						1 colum	nn omitted

## 4.2 Interpretation

Examining descriptive statistics is an important first step in any analysis. It is important to examine your descriptive statistics before skipping ahead to more sophisticated analyses, such as multilevel models.

In examining the descriptive statistics for this data, we get a sense of the data.

- outcome has a mean of approximately 52 and ranges from approximately 30 to 75.
- warmth and physical punishment are both variables that represent the number of times that parents use each of these forms of discipline in a week. The average of the former is about 3.5, while the average of the latter is about 2.5.
- HDI, the Human Development Index has an average of about 65, and a wide range.
- identity is a categorical variable for a hypothetical identity group, and has values of 0 and 1.
- intervention is also a categorical variable, and has values of 0 and 1.

# **5** Unconditional Models

#### 5.1 Two Level Model

An *unconditional* multilevel model is a model with no independent variables. One should always run an unconditional model as the first step of a multilevel model in order to get a sense of the way that variation is apportioned in the model across the different levels.

#### 5.1.1 The Equation

$$outcome_{ij} = \beta_0 + u_{0j} + e_{ij} \tag{5.1}$$

The Intraclass Correlation Coefficient (ICC) is given by:

$$ICC = \frac{var(u_{0j})}{var(u_{0j}) + var(e_{ij})}$$

$$(5.2)$$

In a two level multilevel model, the ICC provides a measure of the proportion of variation attributable to Level 2.

#### 5.1.2 Run Models

#### 5.1.2.1 Stata

```
use simulated_multilevel_data.dta // use data
mixed outcome || country: // unconditional model
estat icc // ICC
```

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 = 3,000Number of groups = 30 Group variable: country Obs per group: min = 100avg = 100.0max = 100Wald chi2(0) Log likelihood = -9802.8371Prob > chi2 outcome | Coefficient Std. err. z P>|z| [95% conf. interval] \_\_\_\_\_\_ \_cons | 52.43327 .3451217 151.93 0.000 51.75685 Random-effects parameters | Estimate Std. err. [95% conf. interval] \_\_\_\_\_\_ country: Identity var(\_cons) | 3.178658 .9226737 1.799552 5.614658 \_\_\_\_\_\_ var(Residual) | 39.46106 1.024013 37.50421 \_\_\_\_\_\_ LR test vs. linear model: chibar2(01) = 166.31 Prob >= chibar2 = 0.0000 Intraclass correlation -----Level | ICC Std. err. [95% conf. interval] ----country | .0745469 .0201254 .0434963 .1248696

Performing EM optimization ...

#### 5.1.2.2 R

```
library(haven)
df <- read_dta("simulated_multilevel_data.dta")</pre>
library(lme4) # estimate multilevel models
fit0 <- lmer(outcome ~ (1 | country),</pre>
            data = df) # unconditional model
summary(fit0)
Linear mixed model fit by REML ['lmerMod']
Formula: outcome ~ (1 | country)
  Data: df
REML criterion at convergence: 19605.9
Scaled residuals:
   Min 1Q Median
                          3Q
                                   Max
-3.3844 -0.6655 -0.0086 0.6725 3.6626
Random effects:
 Groups Name
                     Variance Std.Dev.
 country (Intercept) 3.302 1.817
 Residual
                     39.461
                              6.282
Number of obs: 3000, groups: country, 30
Fixed effects:
           Estimate Std. Error t value
(Intercept) 52.433 0.351 149.4
library(performance)
performance::icc(fit0) # ICC
# Intraclass Correlation Coefficient
    Adjusted ICC: 0.077
  Unadjusted ICC: 0.077
```

#### 5.1.2.3 Julia

```
using Tables, MixedModels, MixedModelsExtras,
StatFiles, DataFrames, CategoricalArrays, DataFramesMeta
df = DataFrame(load("simulated_multilevel_data.dta"))
@transform!(df, :country = categorical(:country))
m0 = fit(MixedModel,
         @formula(outcome ~ (1 | country)), df) # unconditional model
Linear mixed model fit by maximum likelihood
 outcome ~ 1 + (1 | country)
          -2 logLik
                          AIC
                                    AICc
                                                BIC
   logLik
 -9802.8371 19605.6742 19611.6742 19611.6822 19629.6933
Variance components:
            Column
                     Variance Std.Dev.
country
         (Intercept)
                       3.17863 1.78287
Residual
                      39.46106 6.28180
 Number of obs: 3000; levels of grouping factors: 30
 Fixed-effects parameters:
               Coef. Std. Error
                                       z Pr(>|z|)
(Intercept) 52.4333
                        0.345121 151.93
                                            <1e-99
icc(m0) # ICC
```

```
0.07454637475695493
```

#### 5.1.3 Interpretation

In each case, the software finds that nearly 8% of the variation in the outcome is explainable by the clustering of the observations in each country.

#### 5.2 Three Level Model

#### 5.2.1 The Equation

$$outcome_{ij} = \beta_0 + u_{0j} + v_{0i} + e_{ij}$$
 (5.3)

As discussed in the main text, in a three level model, there are two intraclass correlation coefficients (StataCorp, 2023). The formulas for the Intraclass Correlation Coefficient (ICC) are given by (StataCorp, 2023):

$$ICC = \frac{var(u_{0j})}{var(u_{0j}) + var(v_{0i}) + var(e_{ij})}$$

$$(5.4)$$

Following StataCorp (2023), Equation 5.4 is the correlation of responses for person-timepoints from the same country but different persons.

$$ICC = \frac{var(u_{0j}) + var(v_{0i})}{var(u_{0j}) + var(v_{0i}) + var(e_{ij})}$$
(5.5)

Again, closely following StataCorp (2023), Equation 5.5 is the correlation of responses for person-timepoints from the same country and same person.

#### 5.2.2 Run Models

#### 5.2.2.1 Stata

```
use simulated_multilevel_longitudinal_data.dta // use data
mixed outcome || country: || id: // unconditional model
estat icc // ICC
```

Performing EM optimization ...

Performing gradient-based optimization: Iteration 0: Log likelihood = -29058.266 Iteration 1: Log likelihood = -29058.259 Iteration 2: Log likelihood = -29058.259

#### Mixed-effects ML regression

Number of obs = 9,000

Prob > chi2 = 0.0000

#### Grouping information

	No. of	Obser	vations per	group
Group variable	groups	Minimum	Average	Maximum
country id	30	300 3	300.0	300
	3,000	3		

outcome | Coefficient Std. err. z P>|z| [95% conf. interval]

\_cons | 53.37768 .3387943 157.55 0.000 52.71366 54.04171

Random-effects parameters		Std. err.		interval]
country: Identity	   3.232092	.8891367	1.885043	5.54174
id: Identity	   11.72403	.5747501	10.64996	12.90641
var(Residual)		.5154843	27.24178	29.26287
	======================================			<b></b>

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

LR test vs. linear model: chi2(2) = 1314.88

#### Intraclass correlation

Level | ICC Std. err. [95% conf. interval]

country | .0748336 .0190847 .0450028 .1219141 id|country | .3462837 .0171461 .3134867 .3806097

-----

#### 5.2.2.2 R

In R, the ICC for a three level model is easiest to estimate "by hand".

```
library(haven)
dfL <- read_dta("simulated_multilevel_longitudinal_data.dta")</pre>
library(lme4) # estimate multilevel models
fit0L <- lmer(outcome ~ (1 | country/id),</pre>
            data = dfL) # unconditional model
summary(fit0L)
Linear mixed model fit by REML ['lmerMod']
Formula: outcome ~ (1 | country/id)
  Data: dfL
REML criterion at convergence: 58116.8
Scaled residuals:
            1Q Median 3Q
                                   Max
-3.7858 -0.6059 -0.0062 0.6017 3.4348
Random effects:
 Groups
        Name Variance Std.Dev.
 id:country (Intercept) 11.724 3.424
 country (Intercept) 3.351 1.830
 Residual
                       28.234 5.314
Number of obs: 9000, groups: id:country, 3000; country, 30
Fixed effects:
           Estimate Std. Error t value
(Intercept) 53.3777 0.3446 154.9
```

```
3.351 / (11.724 + 3.351 + 28.234)
[1] 0.07737422
(3.351 + 11.724) / (11.724 + 3.351 + 28.234)
[1] 0.3480801
5.2.2.3 Julia
In Julia, the ICC for a three level model is also easiest to estimate "by hand".
using Tables, MixedModels, StatFiles, DataFrames, CategoricalArrays, DataFramesMeta
dfL = DataFrame(load("simulated_multilevel_longitudinal_data.dta"))
@transform!(dfL, :country = categorical(:country))
mOL = fit(MixedModel, @formula(outcome ~
                                  (1 | country) +
                                  (1 | id)), dfL)
Linear mixed model fit by maximum likelihood
 outcome ~ 1 + (1 | country) + (1 | id)
    logLik
             -2 logLik
                            AIC
                                         AICc
                                                     BIC
 -29058.2592 58116.5184 58124.5184 58124.5229 58152.9384
Variance components:
            Column
                     Variance Std.Dev.
id
         (Intercept) 11.72401 3.42403
country (Intercept)
                       3.23190 1.79775
Residual
                      28.23426 5.31359
 Number of obs: 9000; levels of grouping factors: 3000, 30
  Fixed-effects parameters:
               Coef. Std. Error
                                        z Pr(>|z|)
```

<1e-99

0.338785 157.56

(Intercept) 53.3777

```
3.23190 / (11.72401 + 3.23190 + 28.23426)
```

0.07482952718176382

```
(3.23190 + 11.72401) / (11.72401 + 3.23190 + 28.23426)
```

0.34628041519632824

#### 5.2.3 Interpretation

Each software suggests that almost 8% of the variation in the outcome is within time points for different individuals within the same country, while almost 35% of the variation in the outcome is within time points for the same individual within the same country.

# 6 Cross Sectional Multilevel Models

# 6.1 The Equation

Recall the general model of Equation 1.1, and the syntax outlined in Section 1.3. Below in Equation 6.1, we consider a more substantive example.

$$outcome_{ij} = \beta_0 + \beta_1 warmth_{ij} +$$
(6.1)

 $\beta_2$ physical punishment<sub>ij</sub>+

$$\beta_3 \mathrm{identity}_{ij} + \beta_4 \mathrm{intervention}_{ij} + \beta_5 \mathrm{HDI}_j +$$

$$u_{0j} + u_{1j} \times \text{warmth}_{ij} + e_{ij}$$

# 6.2 Correlated and Uncorrelated Random Effects

Consider the covariance matrix of random effects (e.g.  $u_{0j}$  and  $u_{1j}$ ). In Equation 6.2 the covariances of the random effects are constrained to be zero.

$$\begin{bmatrix} var(u_{0j}) & 0\\ 0 & var(u_{1j}) \end{bmatrix} \tag{6.2}$$

As discussed in the Chapter on multilevel models with cross-sectional data, however, one can consider a multilevel model in which the random effects are correlated, as is the case in Equation 6.3.

$$\begin{bmatrix} var(u_{0j}) & cov(u_{0j}, u_{1j}) \\ cov(u_{0j}, u_{1j}) & var(u_{1j}) \end{bmatrix}$$
 (6.3)

Procedures for estimating models with uncorrelated and correlated random effects are detailed below (Bates et al., 2015; Bates, 2024; StataCorp, 2023).

Table 6.1: Correlated and Uncorrelated Random Effects

Software	Uncorrelated Random Effects	Correlated Random Effects
Stata R	default separate random effects from grouping variable with	add option: , cov(uns) separate random effects from grouping variable with
Julia	separate terms for each random effect e.g. (1   group) + (0 + x   group)	separate random effects from grouping variable with  .

All models in the examples below are run with uncorrelated random effects, but could just as easily be run with *correlated* random effects.

# 6.3 Run The Models



Continuous and Categorical Variables

Statistically-as noted in the main text-it is important to be clear on whether independent variables in one's model are continuous or categorical. Continuous variables can be entered straightforwardly into statistical syntax. Categorical variables, on the other hand usually require specific attention in statistical software. In Stata, categorical variables are indicated in a statistical model by prefixing them with an i.. In R, categorical variables are distinguished by making them into factors e.g. x <- factor(x). In Julia, categorical variables are created by using the Otransform syntax detailed below.

# 6.3.1 Stata

# 6.3.1.1 Get The Data

use simulated\_multilevel\_data.dta

# 6.3.1.2 Run The Model

mixed outcome warmth physical\_punishment i.identity i.intervention HDI || ///
country: warmth

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

Group variable: country

Number of obs = 3,000

Number of groups = 30

Obs per group:

min = 100 avg = 100.0 max = 100

Wald chi2(5) = 334.14Prob > chi2 = 0.0000

Log likelihood = -9626.607

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) var(_cons)		.0227504 2.963975	.0257784 .9737647	.0024689 1.556777	.2096436

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.

# 6.3.2 R

# 6.3.2.1 Get The Data

```
library(haven)
df <- read_dta("simulated_multilevel_data.dta")</pre>
```

# 6.3.2.2 Change Some Variables To Categorical

```
df$identity <- factor(df$identity)</pre>
df$intervention <- factor(df$intervention)</pre>
```

# 6.3.2.3 Run The Model



# Caution

lme4 does not directly provide p values in results, because of some disagreement over exactly how these p values should be calculated. Therefore, in this Appendix, I also call library lmerTest to provide p values for lme4 results.



R prefers to use scientific notation when possible. I find that the use of scientific notation can be confusing in reading results. I turn off scientific notation by setting a penalty for its use: options(scipen = 999).

```
library(lme4)
library(lmerTest)
```

```
options(scipen = 999)
fit1 <- lmer(outcome ~ warmth + physical_punishment +</pre>
              identity + intervention + HDI +
              (1 + warmth || country),
            data = df)
summary(fit1)
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula: outcome ~ warmth + physical punishment + identity + intervention +
   HDI + (1 + warmth || country)
  Data: df
REML criterion at convergence: 19268.8
Scaled residuals:
   Min
            1Q Median
                           3Q
                                  Max
-3.9774 -0.6563 0.0186 0.6645 3.6730
Random effects:
Groups
          Name
                     Variance Std.Dev.
country
          (Intercept) 3.19120 1.786
country.1 warmth
                      0.02464 0.157
                     35.01779 5.918
Residual
Number of obs: 3000, groups: country, 30
Fixed effects:
                     Estimate Std. Error
                                                  df t value
(Intercept)
                    52.011324 1.414976
                                           30.293141 36.758
warmth
                     0.834562 0.064250 41.896457 12.989
physical_punishment -0.991893 0.079845 2968.012381 -12.423
identity1
                    intervention1
                     0.639060 0.217603 2971.186718 2.937
HDT
                    -0.003394 0.020598
                                           27.592814 -0.165
                              Pr(>|t|)
                   < 0.00000000000000000000 ***
(Intercept)
                   0.00000000000000277 ***
warmth
physical_punishment < 0.000000000000000 ***
identity1
                               0.16678
```

# 6.3.3 Julia

# 6.3.3.1 Get The Data

```
using Tables, MixedModels, StatFiles, DataFrames, CategoricalArrays, DataFramesMeta
df = DataFrame(load("simulated_multilevel_data.dta"))
```

# 6.3.3.2 Change Some Variables To Categorical

```
@transform!(df, :country = categorical(:country))
@transform!(df, :identity = categorical(:identity))
@transform!(df, :intervention = categorical(:intervention))
```

# 6.3.3.3 Run The Model

```
Linear mixed model fit by maximum likelihood
outcome ~ 1 + warmth + physical_punishment + identity + intervention + HDI + (1 | country)
logLik -2 logLik AIC AICc BIC
-9626.6070 19253.2140 19271.2140 19271.2742 19325.2713
```

# Variance components:

Column Variance Std.Dev. Corr.

country (Intercept) 2.963849 1.721583

warmth 0.022756 0.150852 Residual 34.974984 5.913965

Number of obs: 3000; levels of grouping factors: 30

# Fixed-effects parameters:

	Coef.	Std. Error	Z	Pr(> z )
(Intercept)	51.9999	1.37124	37.92	<1e-99
warmth	0.834537	0.0637228	13.10	<1e-38
physical_punishment	-0.991665	0.0797906	-12.43	<1e-34
identity: 1.0	-0.300475	0.217029	-1.38	0.1662
intervention: 1.0	0.639641	0.217452	2.94	0.0033
HDI	-0.0032286	0.0199255	-0.16	0.8713

# 6.4 Interpretation

Models suggest that parental warmth is associated with increases in the beneficial outcome, while physical punishment is associated with decreases in the beneficial outcome. The intervention is associated with increases in the outcome. There is insufficient evidence that either identity group or the Human Development Index are associated with the outcome.

# 7 Longitudinal Multilevel Models

# 7.1 The Data

The data employed in these examples are a longitudinal extension of the data described in Section 1.2.

# 7.2 The Equation

outcome
$$_{itj} = \beta_0 + \beta_1$$
parental warmth $_{itj} + \beta_2$ physical punishment $_{itj} + \beta_3$ time $_{itj} +$  (7.1) 
$$\beta_4 \text{identity}_{itj} + \beta_5 \text{intervention}_{itj} + \beta_6 \text{HDI}_j +$$
 
$$u_{0j} + u_{1j} \times \text{parental warmth}_{itj} +$$
 
$$v_{0ij} + v_{1ij} \times \text{time}_{itj} + e_{itj}$$

# 7.3 Growth Trajectories

Remember, following the discussion in the main text, that in longitudinal multilevel models, the variable for *time* assumes an important role as we are often thinking of a *growth trajectory* over time.

As discussed in the main text, think about a model where *identity* is a (1/0) variable for membership in one of two groups:

$$\text{outcome} = \beta_0 + \beta_t \\ \text{time} + \beta_{\text{identity}} \\ \text{identity} + \beta_{\text{interaction}} \\ \text{identity} \\ \times \\ \text{time} + u_{0i} + e_{it}$$

Then, each identity group has its own intercept and time trajectory:

Table 7.1: Slope and Intercept for Each Group

Group	Intercept	Slope (Time Trajectory)
0	$\beta_0$	$eta_t$
1	$\beta_0 + \beta_{\text{identity}}$	$eta_t + eta_{ ext{interaction}}$

# Main Effects and Interactions

Thus, again following the main text, in longitudinal multilevel models, main effects modify the intercept of the time trajectory, while interactions with time, modify the slope of the time trajectory. Below, we run models with main effects only, then models with main effects, and interactions with time.

# 7.4 Run The Models



# Warning

Remember that we are estimating a model in which time points are nested inside families, who are in turn nested inside countries. For each software package, it is accordingly important to specify the way in which different levels of the data are nested. Pay careful attention to the syntax examples below with regard to id and country

# 7.4.1 Stata

# 7.4.1.1 Get The Data

```
use simulated_multilevel_longitudinal_data.dta
```

# 7.4.1.2 Run The Models

# 7.4.1.2.1 Main Effects Only

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

Performing EM optimization  $\dots$ 

Performing gradient-based optimization:

Iteration 0: Log likelihood = -28523.888
Iteration 1: Log likelihood = -28500.378
Iteration 2: Log likelihood = -28500.105
Iteration 3: Log likelihood = -28500.086
Iteration 4: Log likelihood = -28500.086

Computing standard errors ...

 ${\tt Mixed-effects}\ {\tt ML}\ {\tt regression}$ 

Number of obs = 9,000

# Grouping information

	No. of	Obser	vations per	group
Group variable	groups	Minimum	Average	Maximum
	+			
country	30	300	300.0	300
id	3,000	3	3.0	3

Log likelihood = -28500.086

Wald chi2(6) = 1208.49Prob > chi2 = 0.0000

outcome	Coefficient +	Std. err.	z	P> z	[95% conf.	interval]
t	.9433792	.0658667	14.32	0.000	.8142828	1.072476
warmth	.9140251	.0379154	24.11	0.000	.8397122	.988338
physical_punishment	-1.00861	.0497766	-20.26	0.000	-1.106171	9110499
1.identity	1319026	.1516462	-0.87	0.384	4291236	.1653184
1.intervention	.8592402	.1519616	5.65	0.000	.5614009	1.15708
HDI	.0007913	.0200615	0.04	0.969	0385285	.040111
_cons	50.38381	1.367464	36.84	0.000	47.70363	53.06399

Random-effects parameters	Est	imate St	d. err.	[95% conf. :	interval]
country: Identity   var(_cons)	3.4	18565 .9	9268849	2.009349	5.816108

id: Independent | var(t) | 1.27e-08 2.25e-06 7.3e-160 2.2e+143 | var(\_cons) | 8.42116 .4720261 7.545013 9.399046 | var(Residual) | 26.02918 .4753157 25.11405 26.97765 | LR test vs. linear model: chi2(3) = 1246.06 | Prob > chi2 = 0.0000

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

# 7.4.1.2.2 Interactions With Time

mixed outcome c.t##(c.warmth c.physical\_punishment i.identity i.intervention c.HDI) || count

Performing EM optimization ...

Performing gradient-based optimization:

Iteration 0: Log likelihood = -28522.21
Iteration 1: Log likelihood = -28498.677
Iteration 2: Log likelihood = -28498.468
Iteration 3: Log likelihood = -28498.31
Iteration 4: Log likelihood = -28498.309

Computing standard errors ...

Mixed-effects ML regression

Number of obs = 9,000

Grouping information

Group variable		No. of groups	Obser Minimum	vations per Average	group Maximum
country id		30 3,000	300 3	300.0	300 3

Wald chi2(11) = 1100.25Prob > chi2 = 0.0000

Log likelihood = -28498.309

outcome | Coefficient Std. err. z P>|z| [95% conf. interval]

+						
t	.7582075	.326177	2.32	0.020	.1189123	1.397503
warmth	.8170757	.082662	9.88	0.000	.6550611	.9790903
physical_punishment	-1.009031	.1112932	-9.07	0.000	-1.227162	7909007
1.identity	2387167	.3039964	-0.79	0.432	8345387	.3571053
1.intervention	.6607606	.3044503	2.17	0.030	.064049	1.257472
HDI	.0013614	.0210842	0.06	0.949	0399628	.0426856
I						
c.t#c.warmth	.0483637	.0356074	1.36	0.174	0214255	.1181529
I						
c.t#						
.physical_punishment	.0005421	.0494354	0.01	0.991	0963496	.0974338
1						
identity#c.t						
1	.0554389	.1317444	0.42	0.674	2027754	.3136532
I						
intervention#c.t						
1 l	.0992811	.131925	0.75	0.452	1592872	.3578493
I						
c.t#c.HDI	0009551	.0038216	-0.25	0.803	0084453	.0065352
_cons	50.83632	1.483548	34.27	0.000	47.92862	53.74402
HDI   c.t#c.warmth   c.t#  c.t#  .physical_punishment   identity#c.t   1   intervention#c.t   1   c.t#c.HDI	.0013614 .0483637 .0005421 .0554389 .0992811 0009551	.0210842 .0356074 .0494354 .1317444 .131925 .0038216	0.06 1.36 0.01 0.42 0.75 -0.25	0.949 0.174 0.991 0.674 0.452 0.803	039962802142550963496202775415928720084453	.04268 .11818 .09743 .31368 .35784

Random-effects parameters			[95% conf.	_
country: Independent				
var(warmth)	.0106014	.0127458	.0010046	.1118779
var(_cons)	3.170088	.9153355	1.80009	5.582753
id: Independent				
var(t)	9.47e-10	2.07e-07	1.5e-195	6.0e+176
· <b>-</b>	8.39189		7.515234	9.370809
var(Residual)			25.10101	26.964
LR test vs. linear model: chi2	(4) = 1247.84	4	Prob > chi	2 = 0.0000

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

# 7.4.2 R

# 7.4.2.1 Get The Data

```
library(haven)
dfL <- read_dta("simulated_multilevel_longitudinal_data.dta")</pre>
```

# 7.4.2.2 Change Some Variables To Categorical

```
dfL$identity <- factor(dfL$identity)</pre>
dfL$intervention <- factor(dfL$intervention)</pre>
```

# 7.4.2.3 Run The Models



# Caution

lme4 does not directly provide p values in results, because of some disagreement over exactly how these p values should be calculated. Therefore, in this Appendix, I also call library lmerTest to provide p values for lme4 results.



🕊 Tip

R prefers to use scientific notation when possible. I find that the use of scientific notation can be confusing in reading results. I turn off scientific notation by setting a penalty for its use: options(scipen = 999).

# 7.4.2.3.1 Main Effects Only

```
library(lme4)
library(lmerTest)
options(scipen = 999)
fit2A <- lmer(outcome ~ t + warmth + physical_punishment +</pre>
```

```
identity + intervention + HDI +
              (1 | country/id),
            data = dfL)
summary(fit2A)
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula:
outcome ~ t + warmth + physical punishment + identity + intervention +
   HDI + (1 | country/id)
  Data: dfL
REML criterion at convergence: 57022.7
Scaled residuals:
   Min
            1Q Median
                           3Q
                                  Max
-3.6850 -0.6094 -0.0035 0.6133 3.6792
Random effects:
Groups
           Name
                      Variance Std.Dev.
id:country (Intercept) 8.438
                               2.905
          (Intercept) 3.675
                               1.917
country
                      26.036
Residual
                               5.103
Number of obs: 9000, groups: id:country, 3000; country, 30
Fixed effects:
                                 Std. Error
                      Estimate
                                                     df t value
(Intercept)
                    50.3842343 1.4139114
                                             29.8246912 35.635
                     0.9433806
                                  0.0658755 5998.3764548 14.321
t
                     warmth
                    -1.0087537
physical_punishment
                                 0.0497972 6483.6771808 -20.257
identity1
                    -0.1319548
                                  0.1517350 2968.7828107 -0.870
intervention1
                     0.8591494
                                  0.1520510 2971.8111995 5.650
HDI
                     0.0007909
                                  0.0207656
                                             28.0001855 0.038
                              Pr(>|t|)
(Intercept)
                   < 0.000000000000000 ***
                   < 0.000000000000000 ***
t
                   < 0.000000000000000 ***
physical_punishment < 0.000000000000000 ***</pre>
identity1
                                 0.385
```

0.000000175 \*\*\*

intervention1

```
HDI
                                   0.970
___
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Correlation of Fixed Effects:
           (Intr) t
                          warmth physc_ idntt1 intrv1
           -0.092
           -0.091 -0.002
warmth
physcl_pnsh -0.092 -0.007 -0.012
          -0.051 0.000 -0.013 -0.003
identity1
interventn1 -0.058 0.000 0.039 0.019 -0.018
HDI
           -0.951 0.000 -0.004 0.005 0.000 0.002
7.4.2.3.2 Interactions With Time
fit2B <- lmer(outcome ~ t *(warmth + physical_punishment +</pre>
               identity + intervention + HDI) +
               (1 | country/id),
             data = dfL
summary(fit2B)
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula:
outcome ~ t * (warmth + physical_punishment + identity + intervention +
    HDI) + (1 | country/id)
   Data: dfL
REML criterion at convergence: 57042.8
Scaled residuals:
    Min
             1Q Median
                             3Q
                                    Max
-3.7118 -0.6092 -0.0024 0.6150 3.6779
Random effects:
 Groups
                        Variance Std.Dev.
            Name
 id:country (Intercept) 8.436
                                 2.905
 country
            (Intercept) 3.675
                               1.917
                                 5.104
 Residual
                        26.046
Number of obs: 9000, groups: id:country, 3000; country, 30
```

# Fixed effects:

```
Estimate
                                    Std. Error
                                                         df t value
(Intercept)
                       50.7590272
                                     1.5518360
                                                 43.2608620 32.709
                                                              2.315
t
                        0.7552909
                                     0.3263028 6176.7440549
                                     0.0805355 8274.9995422 10.146
warmth
                        0.8170912
physical punishment
                       -1.0097729
                                     0.1113557 8084.6084915 -9.068
identity1
                       -0.2446453
                                     0.3041604 8695.8966197 -0.804
intervention1
                        0.6604671
                                     0.3046286 8697.0843430
                                                              2.168
HDT
                        0.0026692
                                     0.0221295
                                                 36.1037733
                                                              0.121
                                     0.0356217 6404.8723333
t:warmth
                        0.0486211
                                                              1.365
                                     0.0494590 6753.0158441
                                                              0.010
t:physical_punishment
                        0.0004964
                        0.0563140
                                     0.1318043 5993.4518022
                                                              0.427
t:identity1
                                     0.1319917 5994.1433001
t:intervention1
                        0.0995037
                                                              0.754
                                     0.0038233 5993.9090880 -0.245
t:HDI
                       -0.0009379
                                Pr(>|t|)
(Intercept)
                     0.0207 *
                     <0.000000000000000000002 ***
warmth
                     <0.0000000000000000 ***
physical_punishment
identity1
                                  0.4212
intervention1
                                  0.0302 *
HDT
                                  0.9047
t:warmth
                                  0.1723
t:physical_punishment
                                  0.9920
t:identity1
                                  0.6692
                                  0.4510
t:intervention1
t:HDI
                                  0.8062
___
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Correlation of Fixed Effects:
            (Intr) t
                         warmth physc_ idntt1 intrv1 HDI
                                                           t:wrmt t:phy_
           -0.421
t.
warmth
           -0.178 0.331
physcl pnsh -0.190 0.360 -0.005
identity1
           -0.093 0.166 -0.013 -0.002
interventn1 -0.107 0.192 0.039 0.019 -0.017
HDI
           -0.925 0.264 -0.007 0.012 -0.001 0.003
t:warmth
            0.158 -0.377 -0.882 0.001 0.011 -0.035 0.006
t:physcl_pn 0.170 -0.402 0.004 -0.894 -0.001 -0.017 -0.010 -0.003
t:identity1 0.081 -0.192 0.011 0.000 -0.867 0.014 0.001 -0.013 0.002
t:intrvntn1 0.093 -0.222 -0.035 -0.017 0.014 -0.867 -0.003 0.041 0.019
```

# 7.4.3 Julia

# 7.4.3.1 Get The Data

```
using Tables, MixedModels, StatFiles, DataFrames, CategoricalArrays, DataFramesMeta

dfL = DataFrame(load("simulated_multilevel_longitudinal_data.dta"))
```

# 7.4.3.2 Change Some Variables To Categorical

```
@transform!(dfL, :country = categorical(:country))
@transform!(dfL, :identity = categorical(:identity))
@transform!(dfL, :intervention = categorical(:intervention))
```

# 7.4.3.3 Run The Models

# 7.4.3.3.1 Main Effects Only

```
(1 | country) +
  (0 + warmth | country) +
  (1 | id)), dfL)
```

```
Linear mixed model fit by maximum likelihood
outcome ~ 1 + t + warmth + physical_punishment + identity + intervention + HDI + (1 | counts
logLik -2 logLik AIC AICc BIC
-28499.6031 56999.2063 57021.2063 57021.2356 57099.3610
```

# Variance components:

Column Variance Std.Dev. Corr.

id (Intercept) 8.387214 2.896069

country (Intercept) 3.167143 1.779647

warmth 0.010762 0.103739 .

Residual 26.027363 5.101702

Number of obs: 9000; levels of grouping factors: 3000, 30

# Fixed-effects parameters:

	Coef.	Std. Error	Z	Pr(> z )
(T	E0 4070	4 00000	07 74	
(Intercept)	50.4673	1.33833	37.71	<1e-99
t	0.943864	0.0658717	14.33	<1e-45
warmth	0.913496	0.0423744	21.56	<1e-99
<pre>physical_punishment</pre>	-1.0079	0.0497622	-20.25	<1e-90
identity: 1.0	-0.127692	0.151583	-0.84	0.3996
intervention: 1.0	0.858997	0.151909	5.65	<1e-07
HDI	-0.000566026	0.0196439	-0.03	0.9770

# 7.4.3.3.2 Interactions With Time

Linear mixed model fit by maximum likelihood

```
outcome ~ 1 + t + warmth + physical_punishment + identity + intervention + HDI + t & warmth logLik -2 logLik AIC AICc BIC -28498.3091 56996.6182 57028.6182 57028.6788 57142.2979
```

# Variance components:

Column Variance Std.Dev. Corr.

id (Intercept) 8.391746 2.896851

country (Intercept) 3.170032 1.780458

warmth 0.010609 0.102999 .

Residual 26.015906 5.100579

Number of obs: 9000; levels of grouping factors: 3000, 30

# Fixed-effects parameters:

	Coef.	Std. Error	z	Pr(> z )
(Intercept)	50.8364	1.48355	34.27	<1e-99
t	0.758209	0.326177	2.32	0.0201
warmth	0.817076	0.0826636	9.88	<1e-22
physical_punishment	-1.00903	0.111293	-9.07	<1e-18
identity: 1.0	-0.238714	0.303996	-0.79	0.4323
intervention: 1.0	0.660761	0.30445	2.17	0.0300
HDI	0.00136065	0.0210842	0.06	0.9485
t & warmth	0.0483635	0.0356074	1.36	0.1744
t & physical_punishment	0.000542203	0.0494355	0.01	0.9912
t & identity: 1.0	0.0554385	0.131745	0.42	0.6739
t & intervention: 1.0	0.0992809	0.131925	0.75	0.4517
t & HDI	-0.000955067	0.00382162	-0.25	0.8027

# 7.5 Interpretation

The main effects only model suggests that time is associated with increases in the outcome. In the main effects model, main effects other than time, indicate whether a particular variable is associated with higher or lower intercepts of the time trajectory, at the beginning of the study time. Warmth is again associated with increases in the outcome, while physical punishment is associated with decreases in the outcome. Identity is again not associated with the outcome, while the intervention is associated with higher levels of the outcome. The Human Development Index is again not associated with the outcome.

The second model adds interactions with time to the first model. Results are largely similar to the prior model. However, here we not only examine whether main effects other than time

are associated with higher or lower time trajectories, but also whether particular variables are associated with differences in the slope of the time trajectory. In this case, we find insufficient evidence that any independent variable is associated with changes in the slope of the time trajectory.

# • Which Interactions To Test?

In this example—for the sake of illustration—I test the interaction of every independent variable with time. In many cases, it may make sense to test only only one or two interactions of time with particular variables of key interest. Also, after finding, as I did in this model, that none of the interactions of other independent variables with time are significant, I might report the model with interactions, or might report only the results of the model with only main effects.

It may be illustrative to imagine how we would interpret the results had a particular interaction term been statistically significant. Let us consider one of the interaction terms with the largest coefficient, intervention#time. The interaction of the intervention with time is positive. Had this coefficient been statistically significant, it would have indicated that the intervention was associated with more rapid increases in the outcome over time in addition to the fact that the intervention is associated with higher initial levels of the outcome.

# 8 Multilevel Logistic Regression

Below, I detail the procedure for multilevel logistic regression models in Stata and R.

# 8.1 The Data

The data employed in these examples are the cross-sectional data described in Section 1.2.

# 8.2 The Equation

To explain statistical syntax for Stata and R, I consider the general case of a multilevel model with categorical dependent variable y, independent variables x and z, clustering variable group, and a random slope for x. i is the index for the person, while j is the index for the group.

$$\ln\left(\frac{p(y)}{1 - p(y)}\right) = \beta_0 + \beta_1 x_{ij} + \beta_2 z_{ij} + u_{0j}$$
(8.1)

۵

Correlated and Uncorrelated Random Effects in Logistic Regression

The reader is referred to the discussion of correlated and uncorrelated random effects in Section 6.2

# 8.2.1 Stata

In Stata mixed, the syntax for a multilevel model of the form described in Equation 8.1 is:

melogit y x z || group:

# 8.2.2 R

In R lme4, the syntax for a multilevel model of the form described in Equation 8.1 is:

```
library(lme4)
glmer(y \sim x + z + (1 \mid group), data = ...)
```

# 8.3 Run Models



Less Variation In Logistic Than Linear Models

Note that in logistic regression models, there is less variation to work with–due to the fact that the outcome is 1/0, than there is in linear models. Therefore, in the models below, I do not attempt to estimate a random slope in addition to a random intercept, as I do in Section 6.

# 8.3.1 Stata

#### 8.3.1.1 Get The Data

```
use simulated_multilevel_data.dta
generate outcome_category = outcome > 52 // dichotomous outcome
```

# 8.3.1.2 Run The Model

As suggested in Equation 8.1, odds ratios are obtained by exponentiating the  $\beta$  coefficients:  $e^{\beta}$ . Stata provides the odds ratios automatically with option, or.

```
melogit outcome_category warmth physical_punishment i.identity i.intervention HDI || ///
country:, or
```

Fitting fixed-effects model:

```
Iteration 0: Log likelihood = -1965.6466
Iteration 1: Log likelihood = -1963.7805
```

Iteration 2: Log likelihood = -1963.7791Iteration 3: Log likelihood = -1963.7791

# Refining starting values:

Grid node 0: Log likelihood = -1908.9697

Iteration 0: Log likelihood = -1908.9697 (not concave)

# Fitting full model:

Iteration 1:	Log likelihood = -1903.703		
Iteration 2:	Log likelihood = -1902.2851		
Iteration 3:	Log likelihood = -1901.3176		
Iteration 4:	Log likelihood = -1901.2662		
Iteration 5:	Log likelihood = -1901.2661		
Mixed-effects	logistic regression	Number of obs	= 3,000
Group variable	e: country	Number of groups =	= 30
		Obs per group:	
		min =	= 100
		avg =	= 100.0
		max =	= 100

Integration method: mvagherm	ite Integration pts. = 7	
------------------------------	--------------------------	--

	Wald chi2(5)	=	219.75
Log likelihood = -1901.2661	Prob > chi2	=	0.0000

outcome_category	Odds ratio	Std. err.	z	P> z	[95% conf.	interval]
warmth	1.292603	.0278565	11.91	0.000	1.239142	1.34837
physical_punishment	.7524276	.0222773	-9.61	0.000	.7100077	.797382
1.identity	.9517262	.0748541	-0.63	0.529	.8157636	1.11035
1.intervention	1.191581	.0940459	2.22	0.026	1.020803	1.390929
HDI	.9990491	.0061371	-0.15	0.877	.9870928	1.01115
_cons	.9115548	.3901774	-0.22	0.829	.3939478	2.109244
country	+ 					
var(_cons)	. 2897697	.0880892			.1596945	.5257944

Note: Estimates are transformed only in the first equation to odds ratios.

Note: \_cons estimates baseline odds (conditional on zero random effects).

LR test vs. logistic model: chibar2(01) = 125.03 Prob >= chibar2 = 0.0000

# 8.3.2 R

# 8.3.2.1 Get The Data

```
library(haven)
df <- read_dta("simulated_multilevel_data.dta")</pre>
df$outcome_category <- 0 # initialize to 0</pre>
df$outcome_category[df$outcome > 52] <- 1 # dichotomous outcome</pre>
```

# 8.3.2.2 Change Some Variables To Categorical

```
df$identity <- factor(df$identity)</pre>
df$intervention <- factor(df$intervention)</pre>
```

# 8.3.2.3 Run The Model



# Caution

lme4 does not directly provide p values in results, because of some disagreement over exactly how these p values should be calculated. Therefore, in this Appendix, I also call library lmerTest to provide p values for lme4 results.



R prefers to use scientific notation when possible. I find that the use of scientific notation can be confusing in reading results. I turn off scientific notation by setting a penalty for its use: options(scipen = 999).

```
library(lme4)
library(lmerTest)
```

```
options(scipen = 999)
fit3 <- glmer(outcome_category ~ warmth + physical_punishment +</pre>
               identity + intervention + HDI +
               (1 | country),
             family = binomial(link = "logit"),
             data = df
summary(fit3)
Generalized linear mixed model fit by maximum likelihood (Laplace
  Approximation) [glmerMod]
Family: binomial (logit)
Formula: outcome_category ~ warmth + physical_punishment + identity +
   intervention + HDI + (1 | country)
  Data: df
    ATC
             BIC logLik deviance df.resid
 3816.6
          3858.7 -1901.3
                           3802.6
                                     2993
Scaled residuals:
            1Q Median
                           3Q
                                  Max
-3.0109 -0.8798 0.4369 0.8428 2.8223
Random effects:
Groups Name
                   Variance Std.Dev.
country (Intercept) 0.2894 0.5379
Number of obs: 3000, groups: country, 30
Fixed effects:
                    Estimate Std. Error z value
                                                         Pr(>|z|)
(Intercept)
                 -0.0926371 0.4277643 -0.217
                                                           0.8286
                   warmth
physical_punishment -0.2844595 0.0295990 -9.610 <0.0000000000000000 ***
identity1
                  -0.0494765 0.0786286 -0.629
                                                           0.5292
intervention1
                  0.1752879 0.0789030 2.222
                                                          0.0263 *
HDI
                  -0.0009513 0.0061388 -0.155
                                                          0.8769
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Correlation of Fixed Effects:
```

```
(Intr) warmth physc_ idntt1 intrv1
warmth -0.158
physcl_pnsh -0.170 -0.082
identity1 -0.086 -0.014 0.002
interventn1 -0.102 0.055 0.006 -0.020
HDI -0.930 -0.007 0.012 -0.001 0.004
```

# 8.3.2.4 Calculate Odds Ratios

R requires one to use a bit of extra syntax to extract the odds ratios. As suggested in Equation 8.1, odds ratios are obtained by exponentiating the  $\beta$  coefficients:  $e^{\beta}$ .

# exp(fixef(fit3))

(Intercept)	warmth	<pre>physical_punishment</pre>	identity1
0.9115242	1.2926176	0.7524208	0.9517275
intervention1	HDI		
1.1915893	0.9990492		

# 9 # Models With Three or More Levels and Cross-Classified Models

# 9.1 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*. As more levels are added to the model (e.g. *Level 4*), we imagine all of these levels to be hierarchically nested.

A cross classified model imagines that the nesting is not hierarchical, but rather that there are two sets of clusters or nestings in which individuals may be nested.

Below, I describe the use of Stata, R, and Julia to estimate these models.

# 9.2 Three Or More Levels

# 9.2.1 The Data

I use the *longitudinal* data from *Multilevel Thinking* (Section 1.2) to which I have added an extra level of *United Nations Region* (Arel-Bundock et al., 2018). This data thus requires a four level model.

# 9.2.2 The Equation

 $outcome_{itjk} = \beta_0 + \beta_1 parental warmth_{itjk} + \beta_2 physical punishment_{itjk} + \beta_3 time_{itjk} + (9.1)$ 

$$\beta_4$$
identity<sub>itik</sub> +  $\beta_5$ intervention<sub>itik</sub> +  $\beta_6$ HDI<sub>itik</sub>+

$$w_{0k} + u_{0j} + v_{0i} + e_{itjk}$$

Here we imagine  $w_{0k}$  (region),  $u_{0j}$  (country) and  $v_{0i}$  (family) are hierarchically nested effects.

# 9.2.3 Run The Models

# 9.2.3.1 Stata

# 9.2.3.1.1 Get The Data

```
use "fourlevel.dta", clear
```

# 9.2.3.1.2 Unconditional Model

```
mixed outcome || UNregion: || country: || family:
```

Performing EM optimization ...

Performing gradient-based optimization:

Iteration 0: Log likelihood = -29061.686
Iteration 1: Log likelihood = -29061.679
Iteration 2: Log likelihood = -29061.679

Computing standard errors ...

Mixed-effects ML regression

Number of obs = 9,000

# Grouping information

Group variable	No. of groups	Obser Minimum	vations per Average	group Maximum
UNregion	5	600	1,800.0	3,600
country	30	300	300.0	300
family	3,000	3	3.0	3

outcome | Coefficient Std. err. z P>|z| [95% conf. interval]

\_cons | 54.05906 .987367 54.75 0.000 52.12385 55.99426

Random-effects parameters			[95% conf.	
UNregion: Identity var(_cons)	   4.172687	3.187885	.9334852	18.65194
<pre>country: Identity      var(_cons)</pre>	     2.849348	.8710225	1.565093	5.187414
family: Identity var(_cons)	   11.72403	.57475	10.64997	12.90641
var(Residual)			27.24177	
LR test vs. linear model: chi	2(3) = 1843.4	4	Prob > chi	2 = 0.0000

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

# 9.2.3.1.3 Conditional Model

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

Performing EM optimization ...

Performing gradient-based optimization: Iteration 0: Log likelihood = -28503.082 Iteration 1: Log likelihood = -28503.039 Iteration 2: Log likelihood = -28503.039

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

	+			
UNregion	J 5	600	1,800.0	3,600
country	J 30	300	300.0	300
id	3,000	3	3.0	3

Log likelihood = -28503.039

Wald chi2(6) = 1209.42Prob > chi2 = 0.0000

outcome		Coefficient	Std. err.	z	P> z	[95% conf.	interval]
t		.9433791	.0658667	14.32	0.000	.8142827	1.072476
warmth		.9140704	.0379156	24.11	0.000	.8397571	.9883837
physical_punishment		-1.008615	.0497772	-20.26	0.000	-1.106176	9110531
1.identity		1332133	.1516437	-0.88	0.380	4304294	.1640028
1.intervention		.8589263	.1519619	5.65	0.000	.5610865	1.156766
HDI		.0148561	.0196605	0.76	0.450	0236777	.0533899
_cons	١	50.16426	1.675219	29.94	0.000	46.88089	53.44763

Random-effects parameters | Estimate Std. err. [95% conf. interval]

UNregion: Identity | var(\_cons) | 4.722007 3.585939 1.065898 20.91884

country: Identity | var(\_cons) | 2.863495 .8656459 1.583342 5.178668

id: Identity | var(\_cons) | 8.421131 .4711947 7.546445 9.397199

var(Residual) | 26.02919 .4752587 25.11417 26.97755

LR test vs. linear model: chi2(3) = 1844.00 Prob > chi2 = 0.0000

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

# 9.2.3.2 R

# 9.2.3.2.1 Get The Data

```
library(haven)
df4 <- read dta("fourlevel.dta")</pre>
```

# 9.2.3.2.2 Change Some Variables To Categorical

```
df4$identity <- factor(df4$identity)</pre>
df4$intervention <- factor(df4$intervention)
```

# 9.2.3.2.3 Unconditional Model



# △ Caution

lme4 does not directly provide p values in results, because of some disagreement over exactly how these p values should be calculated. Therefore, in this Appendix, I also call library lmerTest to provide p values for lme4 results.



R prefers to use scientific notation when possible. I find that the use of scientific notation can be confusing in reading results. I turn off scientific notation by setting a penalty for its use: options(scipen = 999).

```
library(lme4)
library(lmerTest)
options(scipen = 999)
fit4A <- lmer(outcome ~ (1 | UNregion/country/id),</pre>
              data = df4)
summary(fit4A)
```

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula: outcome ~ (1 | UNregion/country/id)
   Data: df4
```

```
REML criterion at convergence: 58121.4
Scaled residuals:
   Min 1Q Median 3Q
                                   Max
-3.7850 -0.6064 -0.0047 0.6020 3.4399
Random effects:
 Groups
                      Name
                                  Variance Std.Dev.
 id:(country:UNregion) (Intercept) 11.724 3.424
 country:UNregion
                      (Intercept) 2.842 1.686
 UNregion
                      (Intercept) 5.478 2.340
 Residual
                                  28.234 5.314
Number of obs: 9000, groups:
id:(country:UNregion), 3000; country:UNregion, 30; UNregion, 5
Fixed effects:
           Estimate Std. Error
                                 df t value Pr(>|t|)
(Intercept) 54.061 1.112 3.777 48.6 0.00000201 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
9.2.3.2.4 Conditional Model
fit4B <- lmer(outcome ~ t + warmth + physical_punishment +</pre>
               identity + intervention + HDI +
               (1 | UNregion/country/id),
             data = df4
summary(fit4B)
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula:
outcome ~ t + warmth + physical_punishment + identity + intervention +
    HDI + (1 | UNregion/country/id)
   Data: df4
REML criterion at convergence: 57026.4
Scaled residuals:
```

```
-3.6846 -0.6096 -0.0038 0.6138 3.6850
Random effects:
                                  Variance Std.Dev.
 Groups
                      Name
 id:(country:UNregion) (Intercept) 8.438
                                           2.905
 country: UNregion
                       (Intercept)
                                   2.979
                                           1.726
 UNregion
                       (Intercept) 6.178
                                           2.486
 Residual
                                  26.036
                                           5.103
Number of obs: 9000, groups:
id:(country:UNregion), 3000; country:UNregion, 30; UNregion, 5
Fixed effects:
                                                 df t value
                      Estimate Std. Error
                                 1.78086
                                           15.79112 28.143
(Intercept)
                      50.11857
                      0.94338
                                 0.06588 5998.37756 14.321
warmth
                      0.91406 0.03793 4745.28492 24.096
physical_punishment
                     -1.00876 0.04980 6483.46337 -20.257
                     -0.13324
                                 0.15173 2969.00938 -0.878
identity1
intervention1
                      0.85872
                                 0.15205 2971.85430
                                                      5.648
                                 0.02006
HDI
                      0.01560
                                           24.39852
                                                      0.778
                               Pr(>|t|)
                     0.0000000000000641 ***
(Intercept)
                    < 0.0000000000000000 ***
                    < 0.000000000000000 ***
warmth
physical_punishment < 0.0000000000000000 ***
identity1
                                   0.380
intervention1
                    0.0000001780521096 ***
HDI
                                   0.444
___
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Correlation of Fixed Effects:
            (Intr) t
                         warmth physc_ idntt1 intrv1
            -0.073
t
warmth
            -0.071 -0.002
physcl pnsh -0.073 -0.007 -0.012
identity1
           -0.040 0.000 -0.013 -0.003
```

interventn1 -0.045 0.000 0.039 0.019 -0.018

HDT

3Q

Max

1Q Median

Min

-0.738 0.000 -0.005 0.005 -0.001 0.001

# 9.2.3.3 Julia

# 9.2.3.3.1 Get The Data

```
using Tables, MixedModels, StatFiles, DataFrames, CategoricalArrays, DataFramesMeta
df4 = DataFrame(load("fourlevel.dta"))
```

# 9.2.3.3.2 Change Some Variables To Categorical

```
@transform!(df4, :country = categorical(:country))
@transform!(df4, :UNregion = categorical(:UNregion))
@transform!(df4, :identity = categorical(:identity))
@transform!(df4, :intervention = categorical(:intervention))
```

# 9.2.3.3.3 Unconditional Model

```
m4A = fit(MixedModel, Oformula(outcome ~ t + warmth +
                                  physical_punishment +
                                   identity + intervention +
                                   HDI +
                                   (1 | UNregion) +
                                   (1 | country) +
                                   (1 | id)), df4)
```

```
Linear mixed model fit by maximum likelihood
 outcome ~ 1 + t + warmth + physical_punishment + identity + intervention + HDI + (1 | UNreg
    logLik -2 logLik
                          AIC
                                       AICc
 -28503.0394 57006.0787 57028.0787 57028.1081 57106.2335
Variance components:
           Column Variance Std.Dev.
```

```
(Intercept) 8.42110 2.90191
id
country (Intercept) 2.86347 1.69218
UNregion (Intercept) 4.72082 2.17274
Residual
                     26.02921 5.10188
```

Number of obs: 9000; levels of grouping factors: 3000, 30, 5

# Fixed-effects parameters:

```
Coef. Std. Error z Pr(>|z|)
(Intercept)
                             1.67514
                                        29.95
                                                <1e-99
                  50.1643
                  0.943379
                             0.0658668
                                       14.32
                                                <1e-45
warmth
                   0.91407
                             0.0379156 24.11
                                             <1e-99
physical_punishment -1.00861
                             0.0497772 -20.26 <1e-90
                  -0.133213
identity
                             0.151644 -0.88 0.3797
                                       5.65 <1e-07
intervention
                  0.858927
                             0.151962
HDI
                   0.0148553
                             0.0196604
                                        0.76
                                                0.4499
```

#### 9.2.3.3.4 Conditional Model

```
Linear mixed model fit by maximum likelihood
outcome ~ 1 + t + warmth + physical_punishment + identity + intervention + HDI + (1 | UNreg
logLik -2 logLik AIC AICc BIC
-28503.0394 57006.0787 57028.0787 57028.1081 57106.2335
```

# Variance components:

```
Column Variance Std.Dev.
id (Intercept) 8.42110 2.90191
country (Intercept) 2.86347 1.69218
UNregion (Intercept) 4.72082 2.17274
Residual 26.02921 5.10188
```

Number of obs: 9000; levels of grouping factors: 3000, 30, 5

# Fixed-effects parameters:

Coef. Std. Error z Pr(>|z|)

(Intercept)	50.1643	1.67514	29.95	<1e-99
t	0.943379	0.0658668	14.32	<1e-45
warmth	0.91407	0.0379156	24.11	<1e-99
physical_punishment	-1.00861	0.0497772	-20.26	<1e-90
identity	-0.133213	0.151644	-0.88	0.3797
intervention	0.858927	0.151962	5.65	<1e-07
HDI	0.0148553	0.0196604	0.76	0.4499

# 9.2.4 Interpretation

There is group level variation attributable to individual, country, and region.

As in other models, parental warmth, and participation in the intervention are associated with increases in the outcome. Parental use of physical punishment is associated with decreases in the outcome.

# 9.3 Cross-Classified Models

# 9.3.1 The Data

I use the *cross-sectional* data from *Multilevel Thinking* (Section 1.2) to which I have added an extra level of a hypothetical language.

# 9.3.2 The Equation

 $\text{outcome}_{ijm} = \beta_0 + \beta_1 \text{parental warmth}_{ijm} + \beta_2 \text{physical punishment}_{ijm} + \beta_3 \text{time}_{ijm} + \ (9.2)$ 

$$\beta_4 \mathrm{identity}_{ijm} + \beta_5 \mathrm{intervention}_{ijm} + \beta_6 \mathrm{HDI}_{ijm} +$$

$$u_{0j} + m_{0m} + e_{ijm}$$

Here  $u_{0j}$  (country) and  $m_{0m}$  (language) are not nested hierarchically, but are cross classified.

### 9.3.3 Run The Models

#### 9.3.3.1 Stata

### 9.3.3.1.1 Get The Data

```
use "crossclassified.dta", clear
```

#### 9.3.3.1.2 Unconditional Model

```
mixed outcome | | _all: R.country | | _all: R.language
Performing EM optimization ...
Performing gradient-based optimization:
Iteration 0: Log likelihood = -9835.8123
Iteration 1: Log likelihood = -9835.8111
Iteration 2: Log likelihood = -9835.8111
Computing standard errors ...
                                      Number of obs = 3,000
Mixed-effects ML regression
Group variable: _all
                                      Number of groups =
                                      Obs per group:
                                               min =
                                                     3,000
                                               avg = 3,000.0
                                               max = 3,000
                                      Wald chi2(0)
Log likelihood = -9835.8111
                                      Prob > chi2
     outcome | Coefficient Std. err. z P>|z| [95% conf. interval]
-----
       _cons | 52.43187 .3590214 146.04 0.000 51.7282 53.13554
         -----
 Random-effects parameters | Estimate Std. err. [95% conf. interval]
______
_all: Identity
```

var(R.count	•		.9244633	1.796798	5.620198
_all: Identity	İ				
var(R.langua	_		.3284087 	.4881235 	1.87482
var(Residu	al)   3	9.62877	1.045619	37.63148	41.73206
LR test vs. linear model:	chi2(2)	= 180.84		Prob > chi2	= 0.0000

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

### 9.3.3.1.3 Conditional Model

mixed outcome warmth physical\_punishment i.identity i.intervention HDI || \_all: R.country ||

Performing EM optimization ...

Performing gradient-based optimization: Iteration 0: Log likelihood = -9663.2195 Iteration 1: Log likelihood = -9663.2194

Computing standard errors ...

Mixed-effects ML regression

Group variable: \_all

Number of obs = 3,000

Number of groups = 1

Obs per group:

min = 3,000avg = 3,000.0

max = 3,000 Wald chi2(5) = 367.04

Log likelihood = -9663.2194 Prob > chi2 = 0.0000

z P>|z| [95% conf. interval] outcome | Coefficient Std. err. \_\_\_\_\_\_ warmth | .8331461 .0579811 14.37 0.000 .7195052 .946787 physical\_punishment | -.9979749 .080268 -12.43 0.000 -1.155297 -.8406525 1.identity | -.2922428 .2191421 -1.33 0.182 -.7217534 .1372678 1.intervention | .6097458 .2195139 0.005 2.78 .1795064 1.039985 HDI | -.0015879 .0204157 -0.08 0.938 -.0416021 .0384262 \_cons | 51.92255 1.411069 36.80 0.000 49.15691 54.6882 -----

Random-effects parameters			[95% conf.	_
_all: Identity var(R.country)	   3.361218	.9603072	1.920024	5.884192
_all: Identity var(R.language)	   1.121946	. 3269535	.6337502	1.986214
var(Residual)	•	.9263999	33.35002	36.98306
I.B test vs. linear model: chi	2(2) = 227.02		Prob > chi	2 = 0.0000

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

## 9.3.3.2 R

## 9.3.3.2.1 Get The Data

```
library(haven)

dfCC <- read_dta("crossclassified.dta")</pre>
```

## 9.3.3.2.2 Change Some Variables To Categorical

```
dfCC$identity <- factor(dfCC$identity)

dfCC$intervention <- factor(dfCC$intervention)</pre>
```

## 9.3.3.2.3 Unconditional Model

```
library(lme4)

library(lmerTest)

options(scipen = 999)

fitCC_A <- lmer(outcome ~</pre>
```

```
(1 | country) +
               (1 | language),
           data = dfCC)
summary(fitCC_A)
Linear mixed model fit by REML. t-tests use Satterthwaite's method [
lmerModLmerTest]
Formula: outcome ~ (1 | country) + (1 | language)
  Data: dfCC
REML criterion at convergence: 19671.8
Scaled residuals:
   Min 1Q Median
                        3Q
                               Max
-3.3899 -0.6602 -0.0104 0.6798 3.6924
Random effects:
Groups Name
                  Variance Std.Dev.
language (Intercept) 0.9604 0.980
country (Intercept) 3.2919 1.814
Residual
                   39.6276 6.295
Number of obs: 3000, groups: language, 100; country, 30
Fixed effects:
          Estimate Std. Error
                                df t value
                                                   Pr(>|t|)
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

#### 9.3.3.2.4 Conditional Model

Error in model.frame.default(data = dfCC, drop.unused.levels = TRUE, formula = outcome ~ : i:

```
summary(fitCC_B)
```

 $\hbox{\it Error in h(simpleError(msg, call)): error in evaluating the argument 'object' in selecting a } \\$ 

#### 9.3.3.3 Julia

#### 9.3.3.3.1 Get The Data

```
using Tables, MixedModels, StatFiles, DataFrames, CategoricalArrays, DataFramesMeta
dfCC = DataFrame(load("crossclassified.dta"))
```

## 9.3.3.3.2 Change Some Variables To Categorical

```
@transform!(dfCC, :country = categorical(:country))
@transform!(dfCC, :language = categorical(:language))
@transform!(dfCC, :identity = categorical(:identity))
@transform!(dfCC, :intervention = categorical(:intervention))
```

### 9.3.3.3.3 Unconditional Model

```
Linear mixed model fit by maximum likelihood
outcome ~ 1 + (1 | country) + (1 | language)
logLik -2 logLik AIC AICC BIC
-9835.8111 19671.6222 19679.6222 19679.6356 19703.6477
```

Variance components:

```
Column Variance Std.Dev.
language (Intercept) 0.956631 0.978075
country (Intercept) 3.177768 1.782629
Residual 39.628773 6.295139
```

Number of obs: 3000; levels of grouping factors: 100, 30

Fixed-effects parameters:

Coef. Std. Error z Pr(>|z|)

(Intercept) 52.4319 0.35902 146.04 <1e-99

## 9.3.3.3.4 Conditional Model

```
Linear mixed model fit by maximum likelihood
outcome ~ 1 + warmth + physical_punishment + identity + intervention + HDI + (1 | country)
logLik -2 logLik AIC AICc BIC
```

-9663.2194 19326.4388 19344.4388 19344.4990 19398.4962

Variance components:

Column Variance Std.Dev. language (Intercept) 1.12193 1.05921 country (Intercept) 3.36119 1.83335

Residual 35.11960 5.92618

Number of obs: 3000; levels of grouping factors: 100, 30

## Fixed-effects parameters:

	Coef.	Std. Error	z	Pr(> z )
(Intercept)	51.9226	1.41106	36.80	<1e-99
warmth	0.833146	0.0579811	14.37	<1e-46
physical_punishment	-0.997975	0.080268	-12.43	<1e-34
identity	-0.292243	0.219142	-1.33	0.1823
intervention	0.609746	0.219514	2.78	0.0055
HDI	-0.00158794	0.0204156	-0.08	0.9380

## 9.3.4 Interpretation

There is group level variation attributable to both language and country.

As in other models, parental warmth, and participation in the intervention are associated with increases in the outcome. Parental use of physical punishment is associated with decreases in the outcome.

# 10 Reshaping Data

## 10.1 Introduction

Cross-sectional analyses (Section 6) make use of data in *wide* format: every row is a person, or family, and every person, or family, has a single row of data.

Longitudinal analyses (Section 7) make use of *long* data: every row is a person-timepoint, or family-timepoint, and every person, or family, has multiple rows of data.

Data therefore sometimes need to be *reshaped*, most often from *wide* format to *long* format. Almost any software that is capable of estimating multilevel models is capable of reshaping data.

Below, I detail the procedure for reshaping data in Stata and R.

## 10.2 Data in Wide Format

## Note

The data below are in wide format.

Every individual in the data set has a *single row of data*. Every row in the data set is an *individual* or *family*.

Table 10.1: Data in Wide Format

Table 10.1: Table continues below

id	$physical\_punishment 1$	warmth1	outcome1	$physical\_punishment 2$
1.1	3	3	57.47	3
1.10	2	0	62.9	3
1.100	2	5	62.71	1
1.11	4	4	55.61	2
1.12	5	4	41.15	5
1.13	4	5	63.66	3

Table 10.2: Data in Wide Format

Table 10.2: Table continues below

warmth2	outcome2	physical_punishment3	warmth3	outcome3	country	HDI
4	55.06	1	2	58.77	1	69
0	56.67	2	0	68.22	1	69
4	51.58	2	5	55.51	1	69
5	50.9	3	3	48.02	1	69
5	45.4	3	4	55.86	1	69
3	64.81	3	3	58.3	1	69

Table 10.3: Data in Wide Format

family	identity	intervention
1	1	0
10	1	0
100	1	1
11	1	1
12	0	0
13	0	1

## 10.3 Data Management

Because reshaping your data dramatically changes the structure of your data...

- 1. It is a good idea to have your raw data saved in a location where it will not be changed, and can be retrieved again if the reshape command does not work correctly, or if you simply want to modify your reshaping data workflow.(CF Section 2.3)
- 2. Usually we want to work with only a *subset* of your data, to keep only the data in which we are interested.
  - In Stata, the command to keep only variables of interest would be: keep y x z t.
  - In R, one option would be to use the subset function: mysubset <- subset(mydata, select = c(y, x, z, t))

## 10.4 Reshaping Data From Wide To Long

Usually, we are most interested in reshaping data from wide to long.

### 10.4.1 Stata

In Stata, I only list variables that vary over time, or are time varying. Stata assumes that variables that are not listed do not vary over time, or are time invariant.

Notice also that our *time varying* data are in the *stub-time* format, e.g. warmth1, warmth2, physical\_punishment1 physical\_punishment2, etc. Because the variables are named in this way, Stata knows to use the *stub* (e.g. warmth) as the variable name, and the numeric value, (e.g. 1, 2, 3) as the timepoint.

The id variable, whatever it is named, has to uniquely identify the observations. A useful Stata command here is isid, e.g. isid id. If your id variable is not unique, it is often due to missing values. drop if id == . usually solves the problem (assuming that your id variable is indeed named id, and not something else).

```
use simulated_multilevel_longitudinal_data_WIDE.dta, clear
describe
reshape long outcome physical_punishment warmth, i(id) j(wave)
```

Contains data from simulated\_multilevel\_longitudinal\_data\_WIDE.dta

Observations: 3,000

Variables: 15 3 Jul 2024 14:29

Variable name	Storage type	Display format	Value label	Variable label
id	str7	%9s		unique country family id
physical_puni	~1 float	%9.0g		1 physical_punishment
warmth1	float	%9.0g		1 warmth
outcome1	float	%9.0g		1 outcome
physical_puni	~2 float	%9.0g		2 physical_punishment
warmth2	float	%9.0g		2 warmth
outcome2	float	%9.0g		2 outcome
physical_puni	~3 float	%9.0g		3 physical_punishment
warmth3	float	%9.0g		3 warmth
outcome3	float	%9.0g		3 outcome
country	float	%9.0g		country id
HDI	float	%9.0g		Human Development Index
family	float	%9.0g		family id
identity	float	%9.0g		hypothetical identity group variable

```
intervention float %9.0g
                                           recieved intervention
Sorted by: id
(j = 1 2 3)
Data
                                 Wide -> Long
                               3,000
Number of observations
                                        ->
                                            9,000
Number of variables
                                  15
                                        -> 10
j variable (3 values)
                                        ->
                                            wave
xij variables:
            outcome1 outcome2 outcome3
                                        ->
                                             outcome
physical_punishment1 physical_punishment2 physical_punishment3->physical_punishment
               warmth1 warmth2 warmth3
                                            warmth
```

### 10.4.2 R

In R, I only list variables that vary over time, or are time varying.

Notice also that our *time varying* data are in the *stub-time* format, e.g. warmth1, warmth2, physical\_punishment1 physical\_punishment2, etc. In order to facilitate reshaping the data, it is helpful in R to insert an underscore (\_) to separate the *stub* from the *time* variable.

```
library(dplyr) # data wrangling
library(tidyr) # tidy (reshape data)
```

```
# rename variables with "_" separator

df <- simulated_multilevel_longitudinal_data_WIDE %>%
    rename(outcome_1 = outcome1,
        outcome_2 = outcome2,
        outcome_3 = outcome3,
        physical_punishment_1 = physical_punishment1,
        physical_punishment_2 = physical_punishment2,
        physical_punishment_3 = physical_punishment3,
        warmth_1 = warmth1,
        warmth_2 = warmth2,
        warmth_3 = warmth3)
```

## 10.5 Data in Long Format

## Note

The data below are in long format.

Every individual/family in the data set has a multiple rows of data. Every row in the data set is an individual-timepoint or family-timepoint.

Table 10.4: Data in Long Format

Table 10.4: Table continues below

country	HDI	family	id	identity	intervention	t
1	69	1	1.1	1	0	1
1	69	1	1.1	1	0	2
1	69	1	1.1	1	0	3
1	69	2	1.2	1	1	1
1	69	2	1.2	1	1	2
1	69	2	1.2	1	1	3

Table 10.5: Data in Long Format

physical_punishment	warmth	outcome
3	3	57.47
3	4	55.06
1	2	58.77
2	1	50.1
3	0	53.31
3	1	49.79

# 11 Aggregating Data

In many instances, we may wish to aggregate data. For example, we may wish to create *contextual variables* representing the average level of an indicator across a group. In the examples I am using in this book, the group under consideration is the country. Aggregating data is also an important part of discussions of *within* and *between* variation, and is an important part of the correlated random effects model.

In the examples below, I create a group level variable for warmth, representing the average level of parental warmth in each country. If warmth is denoted by warmth<sub>ij</sub> then the country level variable is denoted by  $\overline{warmth}_{.j}$ .

Below, I detail the procedure for aggregating data in Stata and R.

### 11.0.1 Stata

#### 11.0.1.1 Get The Data

```
use simulated_multilevel_data.dta
```

### 11.0.1.2 Create A Group Level Variable

```
bysort country: egen mean_warmth = mean(warmth)
```

### 11.0.2 R

#### 11.0.2.1 Get The Data

```
library(haven)

df <- read_dta("simulated_multilevel_data.dta")</pre>
```

## 11.0.2.2 Create A Group Level Variable

```
library(dplyr)

df <- df %>%
   group_by(country) %>%
   mutate(mean_warmth = mean(warmth))
```

## References

- Allaire, J. J., Teague, C., Scheidegger, C., Xie, Y., & Dervieux, C. (2024). Quarto (Version 1.4). https://doi.org/10.5281/zenodo.5960048
- Arel-Bundock, V., Enevoldsen, N., & Yetman, C. (2018). Countrycode: An r package to convert country names and country codes. *Journal of Open Source Software*, 3(28), 848. https://doi.org/10.21105/joss.00848
- Bates, D. (2024). MixedModels.jl Documentation. https://juliastats.org/MixedModels.jl/stable/
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01
- Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. SIAM Review, 59(1), 65–98. https://doi.org/10.1137/141000671
- Hemken, D. (2023). Statamarkdown: 'Stata' markdown. https://CRAN.R-project.org/package=Statamarkdown
- Li, C. (2019). JuliaCall: An R package for seamless integration between R and Julia. The Journal of Open Source Software, 4(35), 1284. https://doi.org/10.21105/joss.01284
- R Core Team. (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/
- Schanen, J. (2021). Math person (Strogatz Prize entry). National Museum of Mathematics. StataCorp. (2023). Stata 18 mixed effects reference manual. Stata Press.
- Thoreau, H. D. (1975). The commercial spirit of modern times [1837]. In J. J. Moldenhauer, E. Moser, & A. C. Kern (Eds.), Early essays and miscellanies. Princeton University Press.