Analyzing Intensive Longitudinal Data

Tilburg University

tesc

Overview

- General background on what all methods for analyzing ILD are doing.
- Discuss some common issues you'll likely encounter.
- Work on analyses yourself in R (likely won't have time to finish everything, BUT you'll have R code you can use in future studies from now on).

Tilburg University

The “rules” of analysis

- Analyzing your data is all about modeling.

- If your data is linear, model it as such!
  - If it isn’t…don’t.

The “rules” of analysis

- Analyzing your data is all about modeling.

- If your data is linear, model it as such!
  - If it isn’t…don’t.

- If the residual variance is the same “across” the board, model it as such!
  - If it isn’t…don’t

The “rules” of analysis

- Analyzing your data is all about modeling.
- If your data is linear, model it as such!
  - If it isn’t…don’t.
- If the residual variance is the same “across” the board, model it as such!
  - If it isn’t…don’t
- If scores are normally distributed around the mean estimates….you get the idea

TILBURG UNIVERSITY

tesc

Dependent Data

- Hierarchical data, or data with dependence between the observations, is no different that outliers etc.

- If dependence is a characteristic of your data, model it.

- Maybe do it with “multilevel analysis” if needed/appropriate(!!).

Example

- Let’s start with an example,
- I have 58 pupils from three different classes
- Interested whether being more extraverted makes you more popular, and if there is a gender difference
- List of all the variables:
  - *pupil*: pupil identification variable
  - *class*: class identification variable
  - student-level independent variables: *extraversion* (continuous; higher scores mean higher extraversion) and *gender* (dichotomous; 0=male, 1 =female)
  - *popular:* continuous outcome variable at the student-level (higher scores indicate higher popularity).

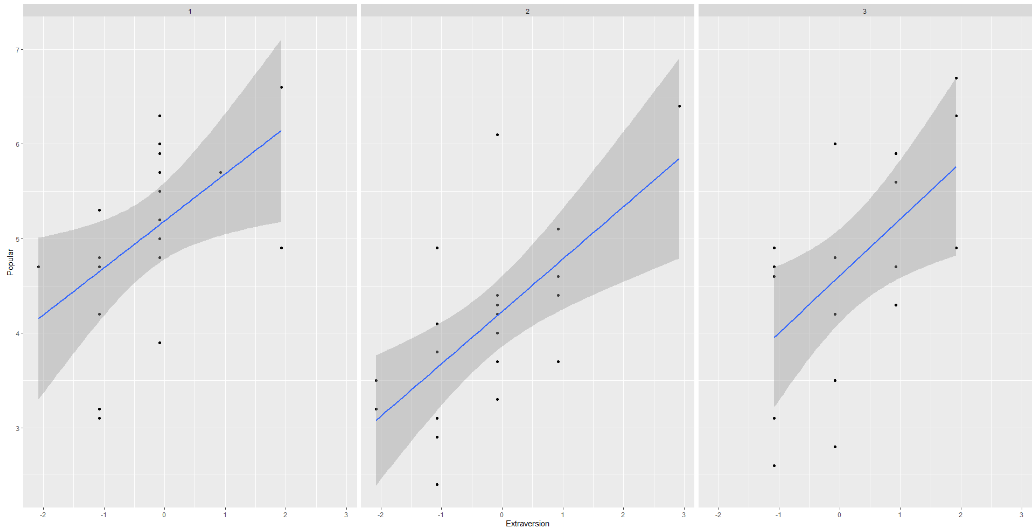
Example

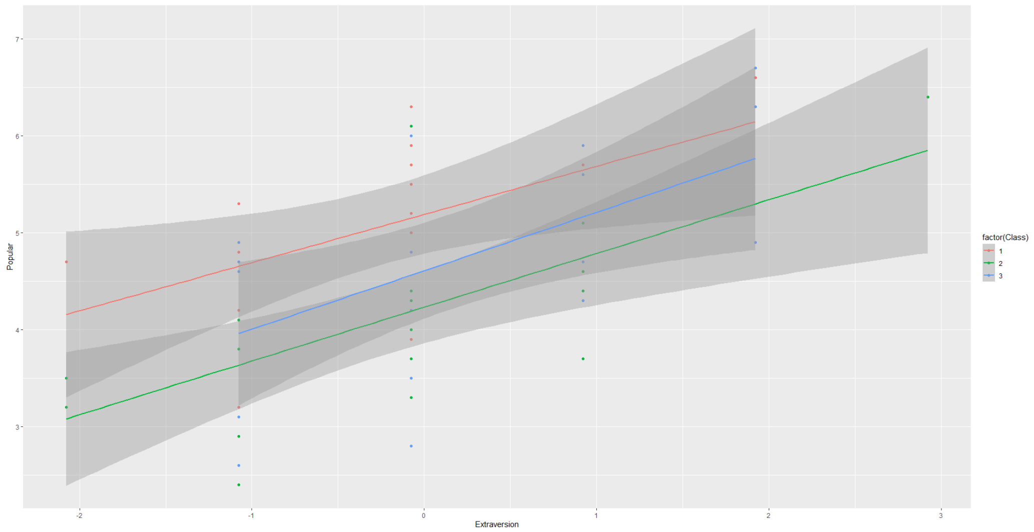
- How would you analyze this?
- Just mention all the different ways you can think of!
- No wrong answers!

Pupil	Class	Extraversion	Gender	Popular
1	1	-0,08	1	6,30
2	1	1,92	0	4,90
3	1	-1,08	1	5,30
4	1	-2,08	1	4,70
5	1	-0,08	1	6,00
6	1	-1,08	0	4,70
7	1	-0,08	0	5,90
8	1	-1,08	0	4,20
9	1	-0,08	0	5,20
10	1	-0,08	0	3,90

Some Options: Option 1

- Analyze all classes separately



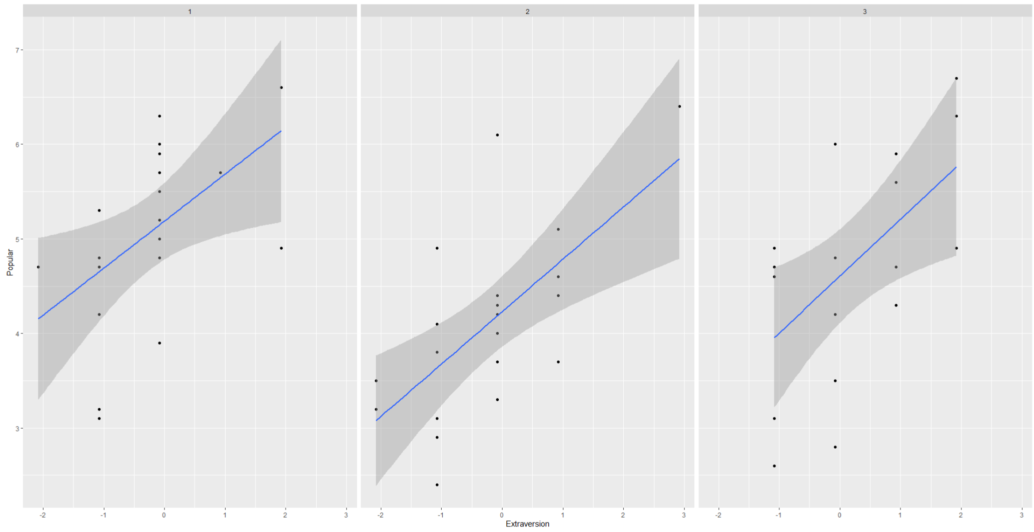


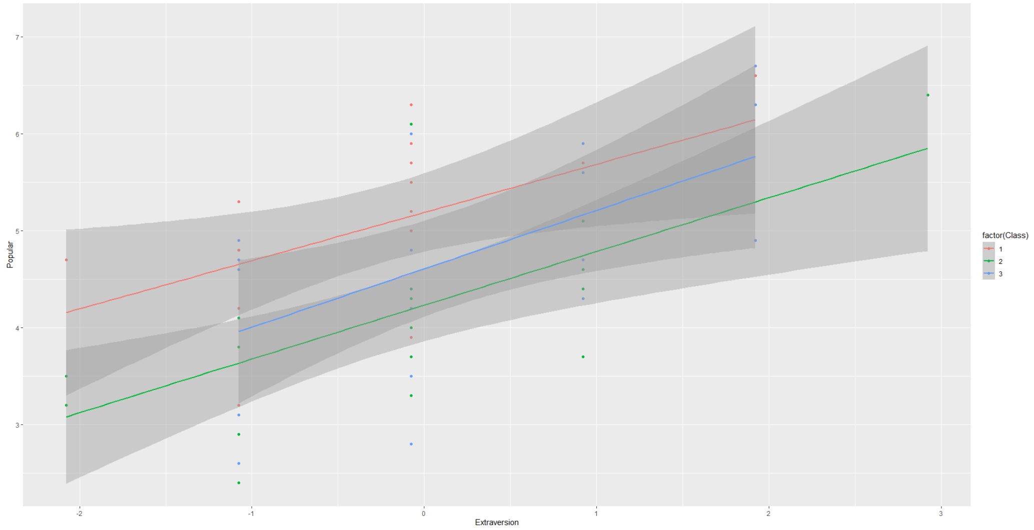
Some Options : Option 1

- Analyze all classes separately

- Benefits?

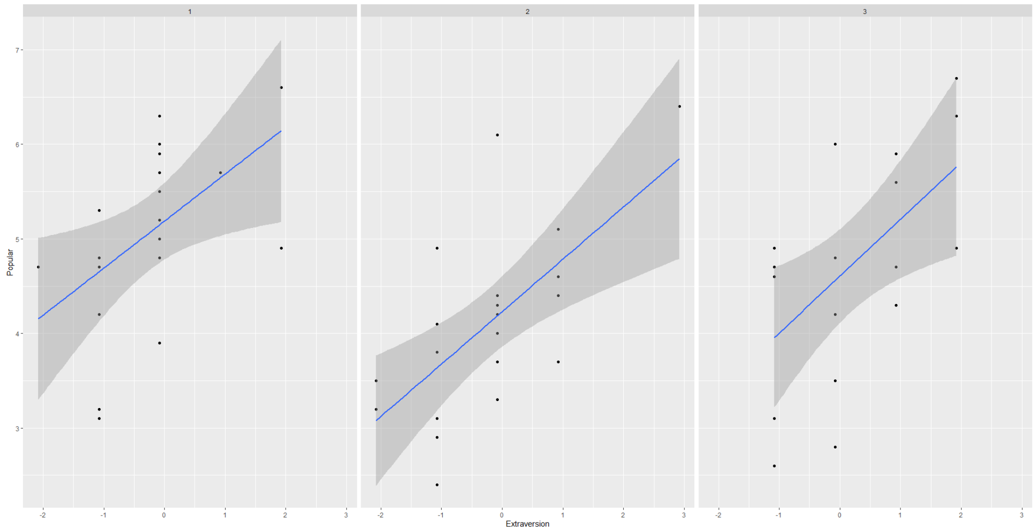
- Disadvantages?

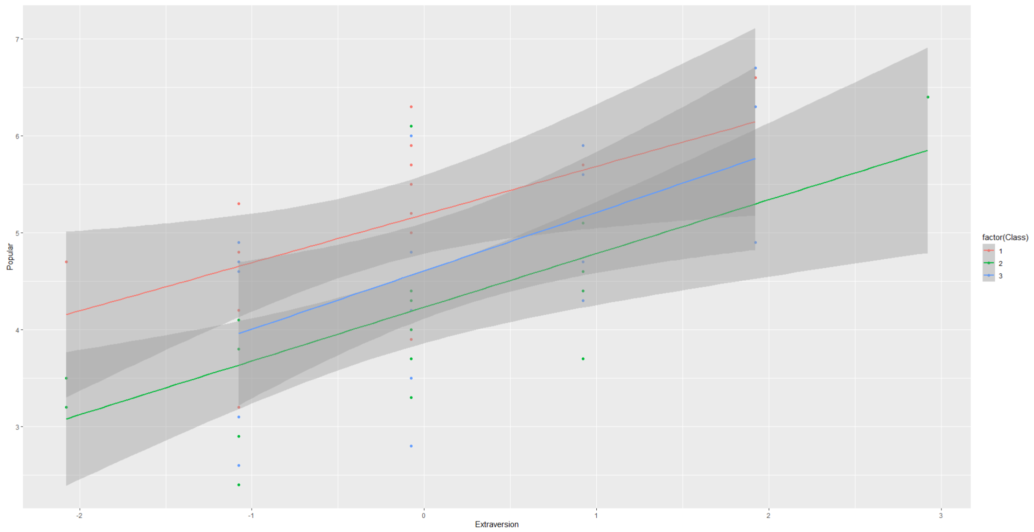




Some Options : Option 1

- Analyze all classes separately
- Benefits?
- Disadvantages?
- How did we deal with the dependence in our data?



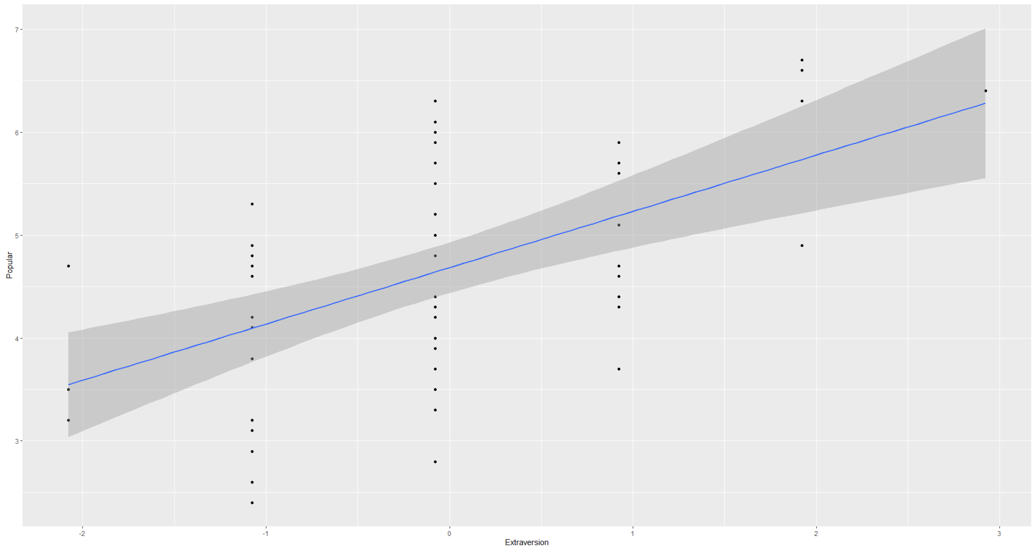


Richard McElreath

- *Many statistical models also have anterograde amnesia. As the models move from one cluster—individual, group, location—in the data to another, estimating parameters for each cluster, they forget everything about the previous clusters. They behave this way, because the assumptions force them to. Any of the models from previous chapters that used dummy variables to handle categories are programmed for amnesia. These models implicitly assume that nothing learned about any one category informs estimates for the other categories—the parameters are independent of one another and learn from completely separate portions of the data. This would be like forgetting you had ever been in a café, each time you go to a new café. Cafés do differ, but they are also alike*

Some Options: Option 2

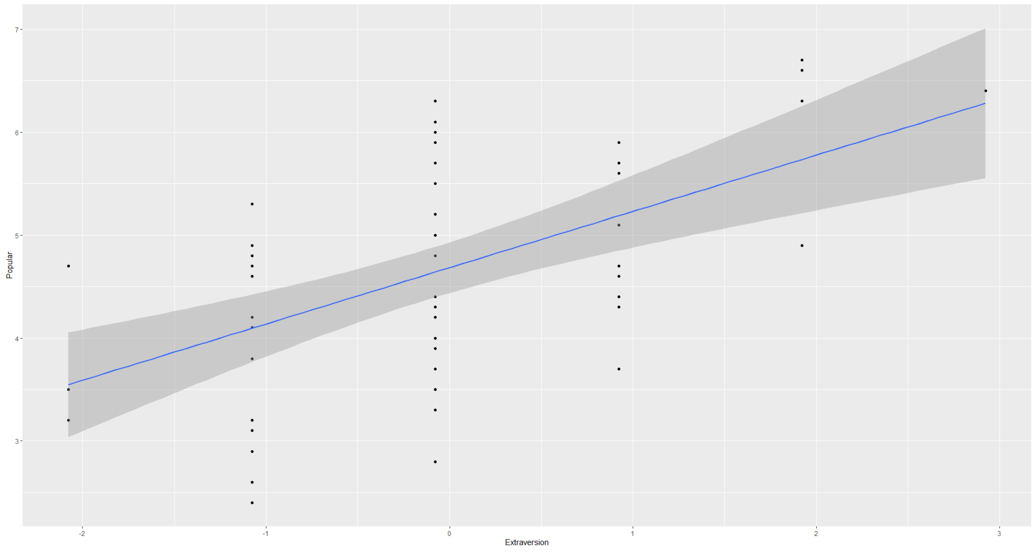
- Ok…so…analyze together then?



Some Options: Option 2

- Ok…so…analyze together then?

- What is the main issue in doing so?

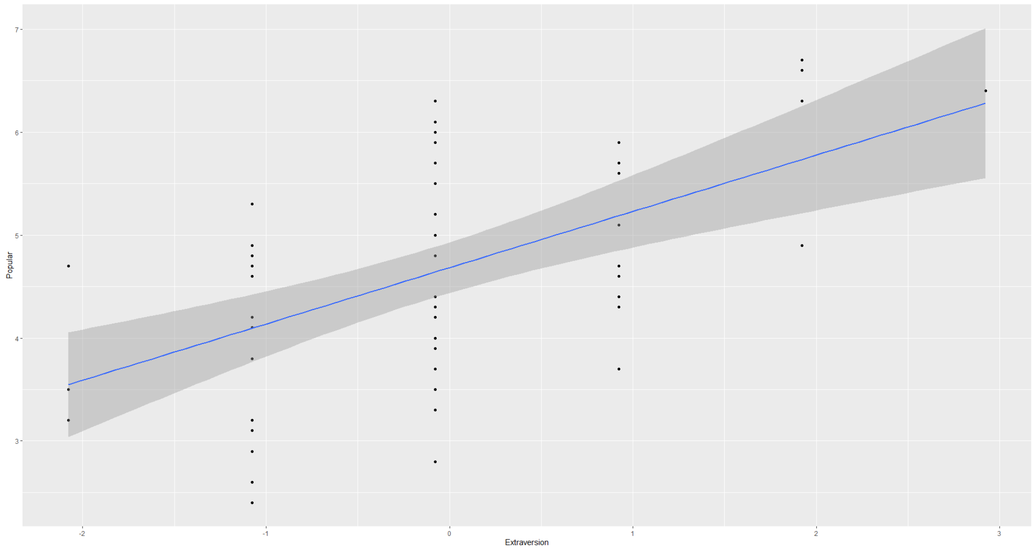


Some Options: Option 2

- Ok…so…analyze together then?

- What is the main issue in doing so?

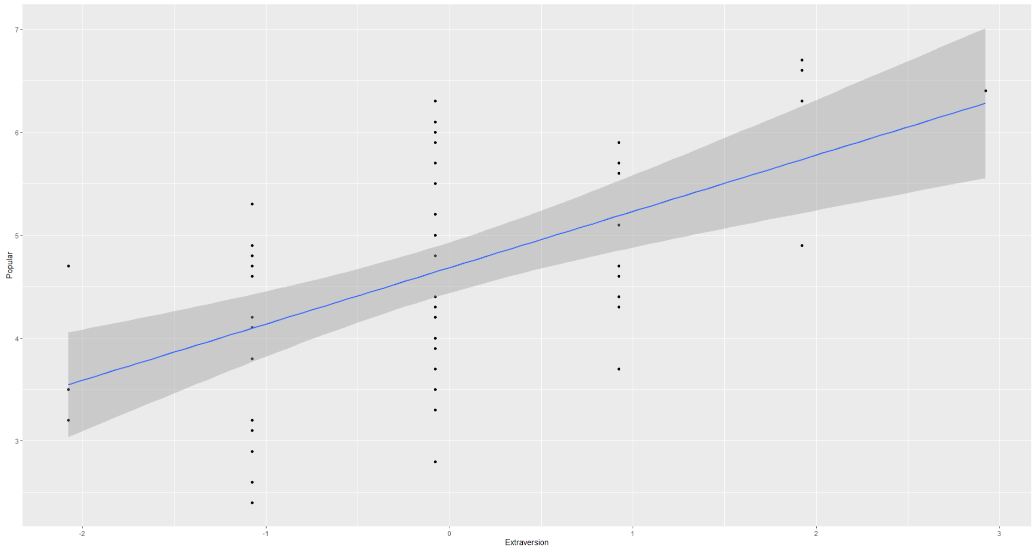
- Hint: How much data do we have?



Some Options: Option 2

- Hint: How much data do we have?

$$n_{eff} = \frac{n}{1 + (n_{clus} - 1)\rho}$$



- Ok, so we have less data than data points.

- Now what?

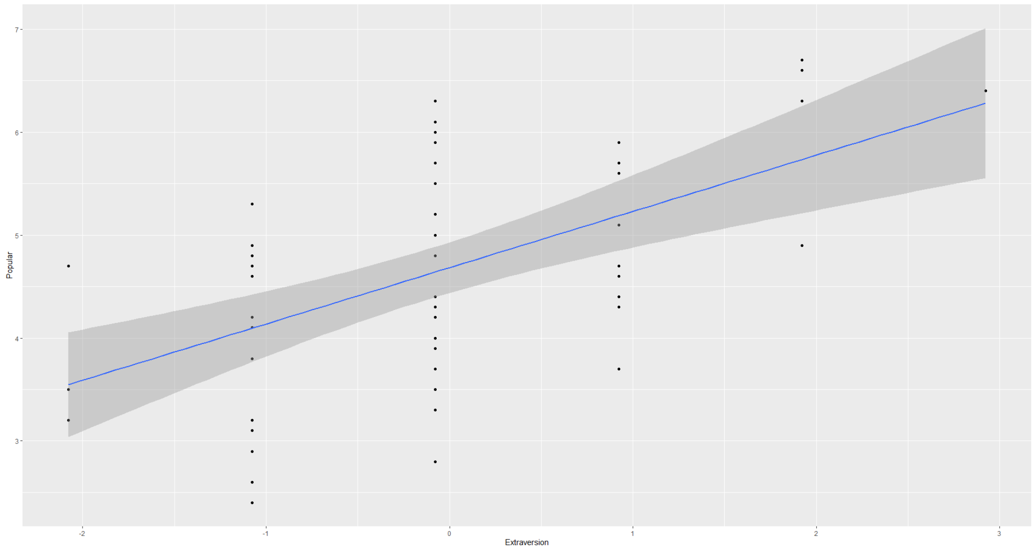
Some Options: Option 2

- Adjust se's

$$v_{eff} = v(1 + (n_{clus} - 1)\rho)$$

- Fortunately, smart ways to do this.

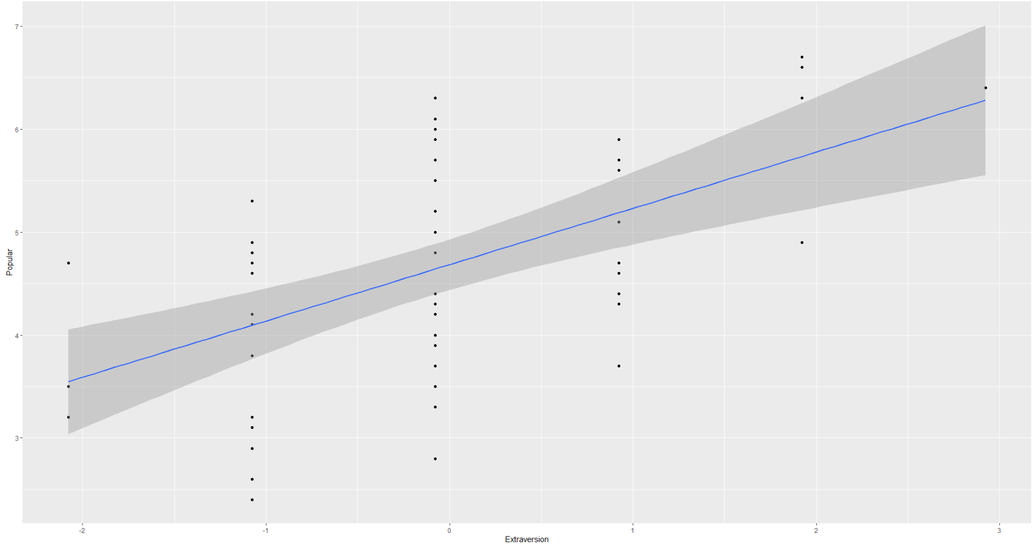
- Cluster robust s.e's



Some Options: Option 2

- Advantages?

- Disadvantages?



Some Options: Option 2

- Pro-tip: Use the formula below in combination with G*power for power analyses!

$$n_{eff} = \frac{n}{1+(n_{clus}-1)\rho} \longrightarrow n = n_{eff} * (1 + (n_{clus} - 1)\rho)$$

- This way you can do a power analysis even if you don't have all the information needed for a simulation study.

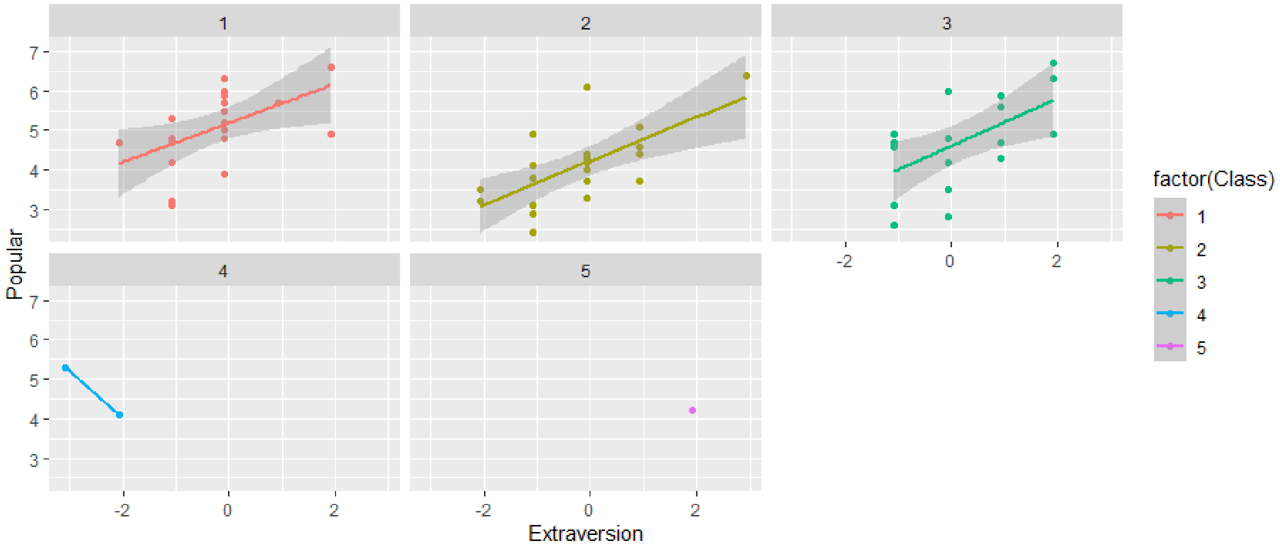
- G*power gives you the n_{eff} you need

Some Options: Option 3 ---- Pool(ing) Party

- Who ever heard of partial-pooling?

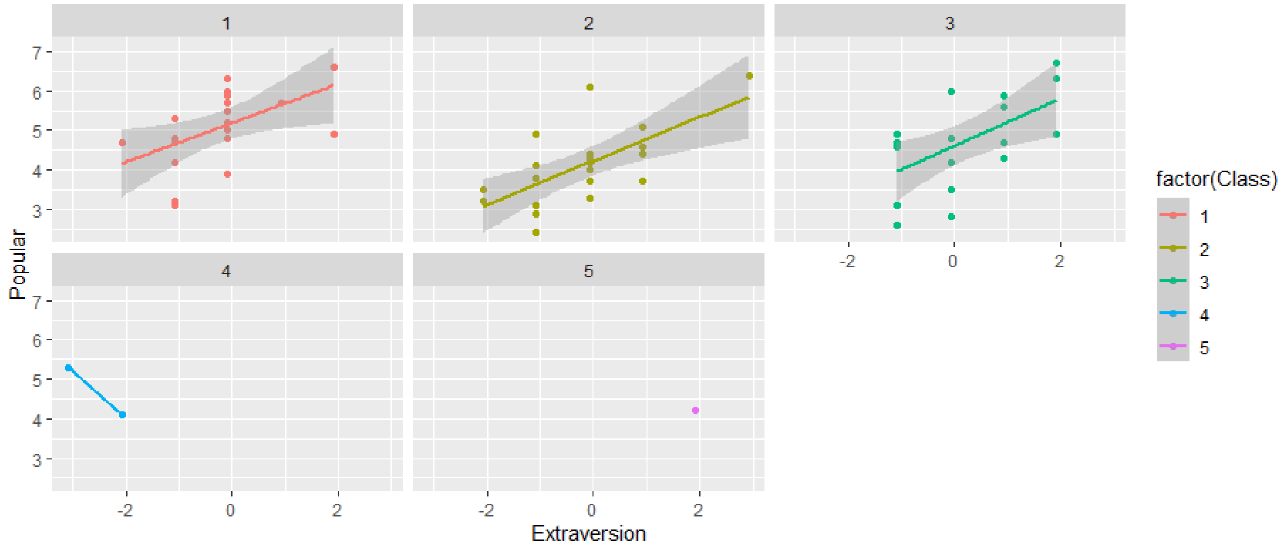
- I'll add two classes to our dataset to illustrate it better
  - One class with two pupils
  - One class with one pupil

Pool(ing) Party



- What would be a problem if I analyzed everyone as N=1?

Pool(ing) Party

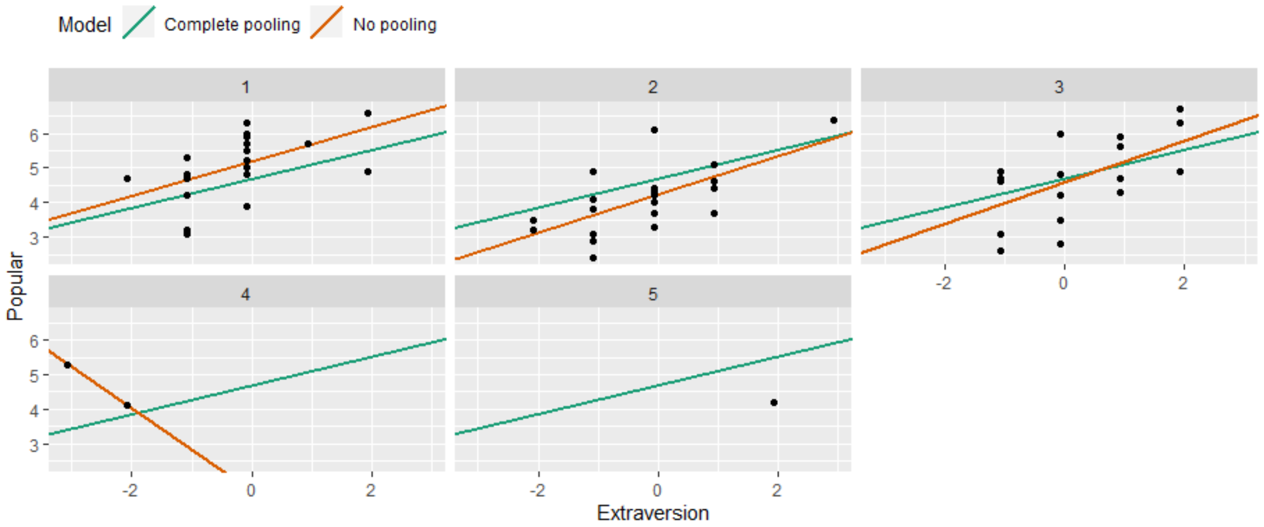


- Each panel shows an independently estimated regression line.
- This approach of fitting a separate line for each class is sometimes called the no pooling model
- Information from different classes is NOT combined or pooled together.

TILBURG UNIVERSITY

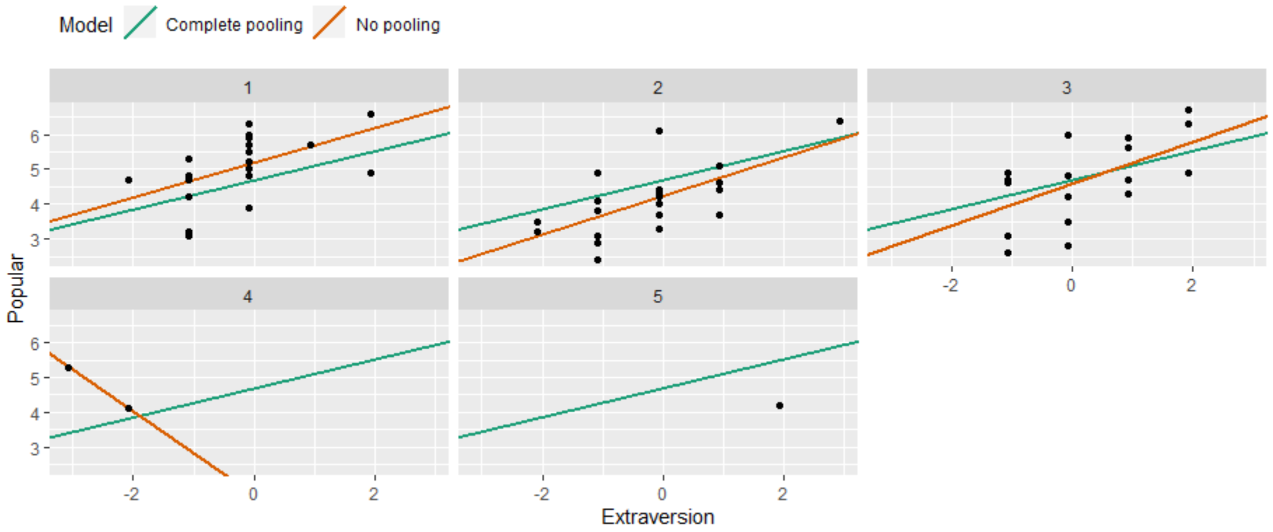
t e s c

Pool(ing) Party



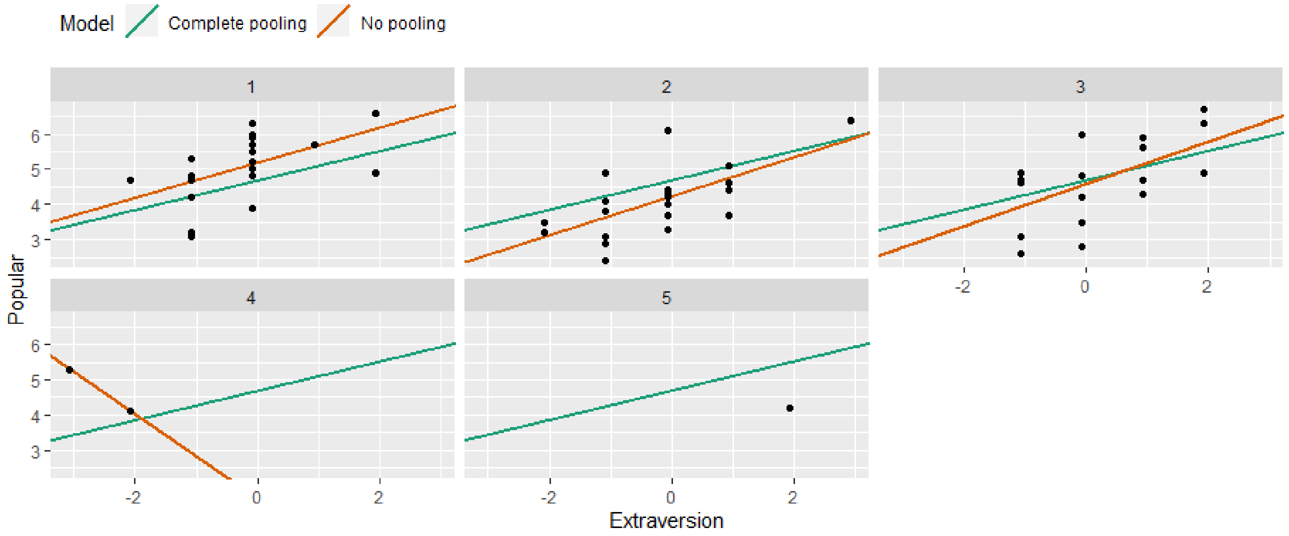
- Could also combine information from all classes (like we saw earlier).
- Fit a single line for the combined data set, unaware that the data came from different participants.

Pool(ing) Party



- Notice anything?
- Advantages?
- Disadvantages?

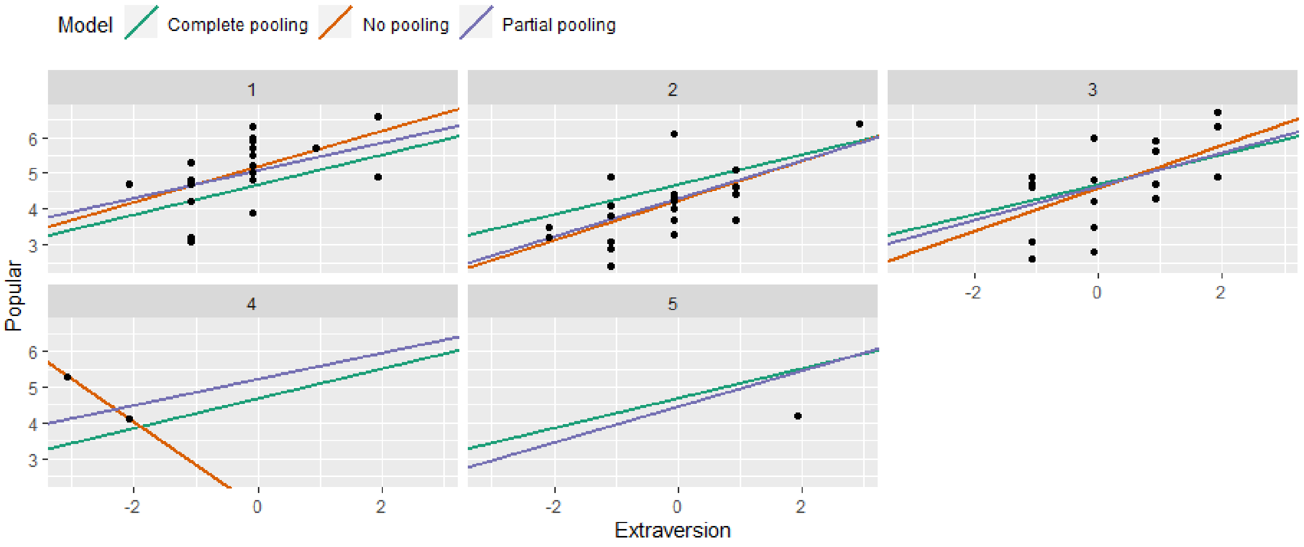
Pool(ing) Party



- We pool information from all the lines together to improve our estimates of each individual line.

- After seeing the 3 trend lines for the classes with complete data, we can make an informed guess about the trend lines for the two classes with incomplete data.

Pool(ing) Party



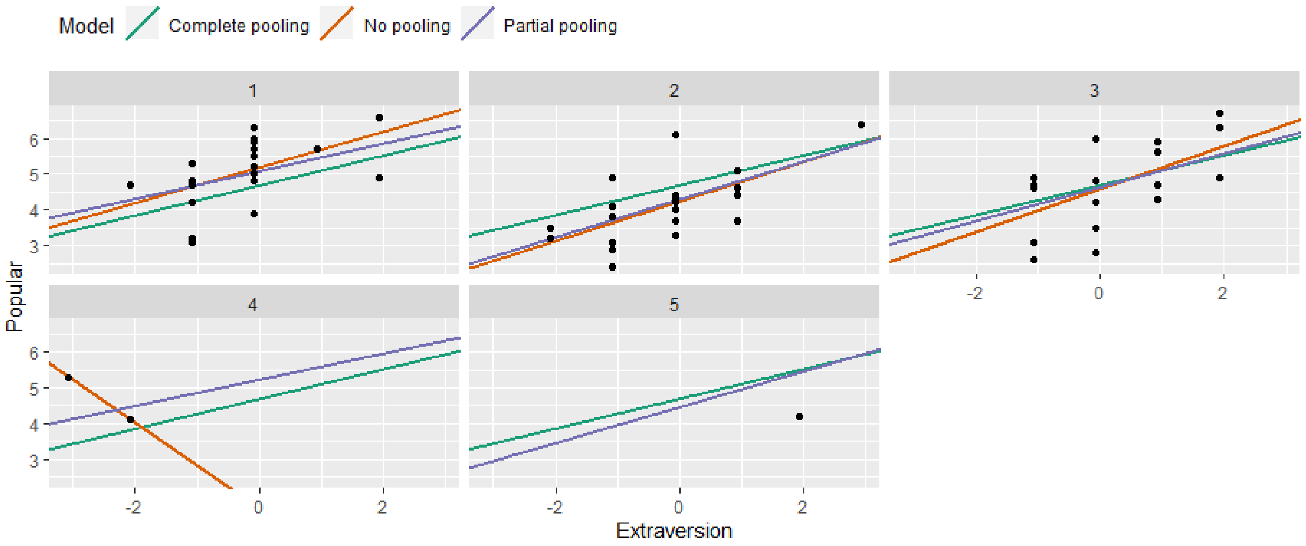
- We pool information from all the lines together to improve our estimates of each individual line.

- After seeing the 3 trend lines for the classes with complete data, we can make an informed guess about the trend lines for the two classes with incomplete data.

Tilburg University

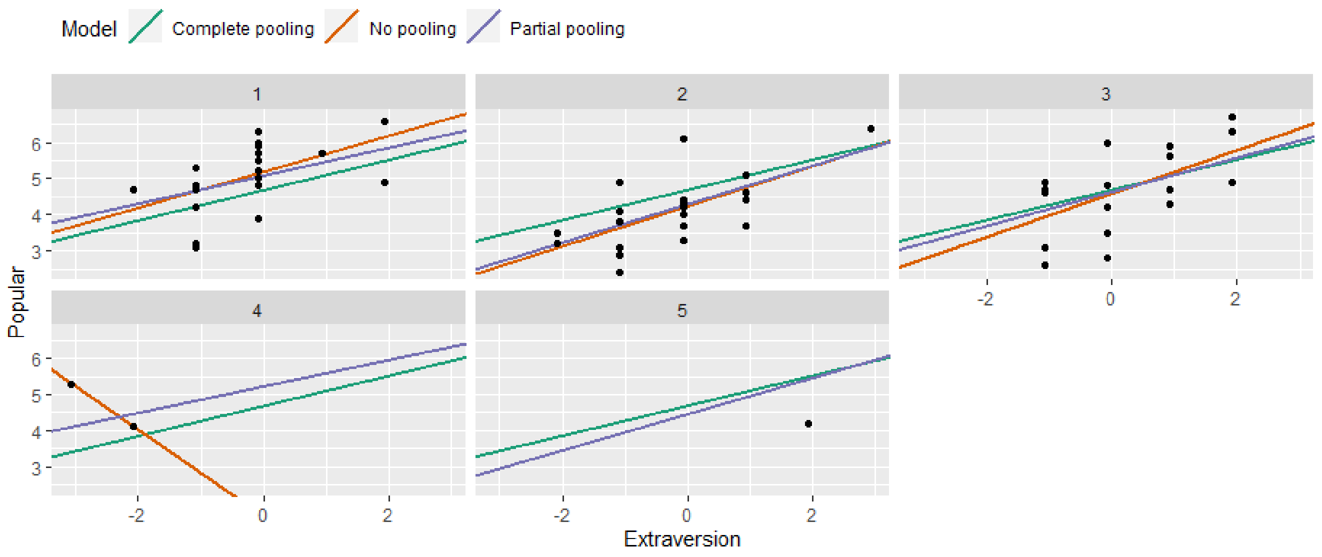
tesc

Pool(ing) Party



- Most of the time, the no pooling and partial pooling lines are the same.
- When the two differ, it's because the partial pooling model's line is pulled slightly towards the complete-pooling line.

Pool(ing) Party



- Amount of pull toward no-pooling line depends on amount of data.

- Also depends on how extreme a person is.

Tilburg University

tesc

Partial Pooling = Multilevel

- So that’s the beauty of Multilevel:
  - We get information about individual units (like in separate analyses).
  - But, we don’t have “amnesia”.

- And can statistically compare for differences between units.
  - Is extraversion equally important in all classes?
  - Does experience of the teacher also play a role?

Residuals

Longitudinal Multilevel Regression

- Remember that you always need to think about what your model says about the data.

- And a multilevel regression (with a fixed effect of time) is saying something that might be a bit weird with longitudinal data.

- Remember:

$$\rho = \frac{\sigma_{u_0}^2}{\sigma_{u_0}^2 + \sigma_e^2}$$

Multilevel Repeated Measures

$$\rho = \frac{\sigma_{u_0}^2}{\sigma_{u_0}^2 + \sigma_e^2}$$

$$\boldsymbol{\Sigma}(\mathbf{Y}) = \begin{pmatrix} \sigma_e^2 + \sigma_{u_0}^2 & \sigma_{u_0}^2 & \sigma_{u_0}^2 & \cdots & \sigma_{u_0}^2 \\ \sigma_{u_0}^2 & \sigma_e^2 + \sigma_{u_0}^2 & \sigma_{u_0}^2 & \cdots & \sigma_{u_0}^2 \\ \sigma_{u_0}^2 & \sigma_{u_0}^2 & \sigma_e^2 + \sigma_{u_0}^2 & \cdots & \sigma_{u_0}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{u_0}^2 & \sigma_{u_0}^2 & \sigma_{u_0}^2 & \cdots & \sigma_e^2 + \sigma_{u_0}^2 \end{pmatrix}.$$

- Makes certain assumptions about variance-covariance matrix
  - Related to Sphericity in RM ANOVA
  - Compound Symmetry
  - All variances and covariances equal (likely?)
  - Inflated Type I errors if violated

- What do we do if violated?

Multilevel Repeated Measures

$$GPA_{it} = \pi_{1i}T1_{it} + \pi_{2i}T2_{it} + \pi_{3i}T3_{it} + \pi_{4i}T4_{it} + \pi_{5i}T5_{it} + \pi_{6i}T6_{it}$$

$$\pi_{1i} = \beta_{10} + u_{1t}$$

$$\pi_{2i} = \beta_{20} + u_{2t}$$

$$\vdots$$

$$\pi_{6i} = \beta_{60} + u_{6t}$$

- Use dummies for all measurement occasions
  - Remove Intercept
  - Each dummie has random slope

- This is basically a MANOVA

SPSS MANOVA analysis

- Disadvantages:
  - no time variable
  - no time varying covariates
  - listwise deletion (can solve this by using multilevel approach on previous slide or multiple imputation)

- No direct test for "change", but can use a contrast

Alternative: Mixed Effects Location-Scale Models

- Remember, analyzing is all about modeling!
- If the variance changes over time -> Model it!!
- Typical patterns in residual (autoregression) can be included in most packages.
- Mixed-Effects Location-Scale models, don’t just specify regression equation for predicted mean, but also for predicted variance.
- Easy to do using brms package in R

Methodological Considerations: Small Level 2 N

Fixed Effects Models

- Multilevel is great, but we are modeling on level 2 as well!

- Would you run a regression on N=4?

- Would you calculate a variance on N=2?

- Same with multilevel, if N on level 2 is small (let's say less than 10), modeling distributions there is probably not a great idea.

- Fortunately, there is a solution! ☺.

Fixed Effects Models

- Hint, the solution is not cluster robust se’s as is sometimes suggested
  - Guess what, that correction also need to be estimated ;).
- How have you corrected for things in regression in the past?

Fixed Effects Models

- Hint, the solution is not cluster robust se's as is sometimes suggested
  - Guess what, that correction also need to be estimated ;).

- How have you corrected for things in regression in the past?

$$SA_{ij} = b_{01}Class1 + b_{02}Class2 + \ \dots + b_{0x}ClassX + b_1 IQ_{ij} + e_{ij}$$

- Add dummies! One for each class (and remove the intercept)

Fixed Effects Models

$$SA_{ij} = b_{01}Class1 + b_{02}Class2 + \ ... + b_{0x}ClassX + b_1 IQ_{ij} + e_{ij}$$

- This is called a Fixed Effects model and is used often in economics.
- It works reaaaaaly well, as the dummies take care of all the level 2 differences.
- Estimates of level 1 predictors unbiased.
- Also deal with “unmodeled” level 2 influences, so is very robust.

Fixed Effects Models

$$SA_{ij} = b_{01}Class1 + b_{02}Class2 + \ \dots + b_{0x}ClassX + b_1 IQ_{ij} + e_{ij}$$

- But! No free lunch.

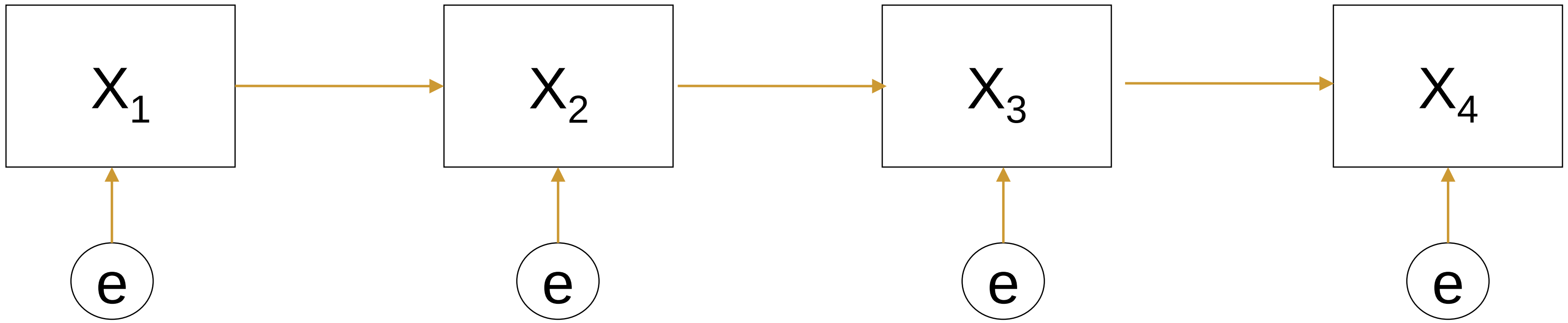
- You can’t model level 2 variables!
  - Since all level 2 variance is in the dummies they are perfectly colinear with level 2 predictors.
  - Not a big problem, if level 2 N is small what do you hope to find there anyway?

- Can model interactions between level 1 and level 2 predictors though.

Reversible Change

Reversible Change

- One type of model used is the AR model



Reversible Change

How could you run an AR-model using tools you already know?

Reversible Change

- Can run an AR model with multilevel regression too!

$$Y_t = b_0 + b_1 Y_{t-1} + \epsilon$$

Reversible Change

- Can run an AR model with multilevel regression too!

$$Y_t = b_0 + b_1 Y_{t-1} + \epsilon$$

$$Y_{ij} = b_0 + b_1 X_{ij} + \epsilon_{ij}$$

Reversible Change

$$Y_t = b_0 + \ b_1 Y_{t-1} + \ \epsilon_t$$

b_0 = Long run tendency → Think "mean".

b_1 = Autoregressive parameter → inertia.

ϵ_t = Residual/Innovation → All variation that can not be predicted by previous measurement.

Reversible Change

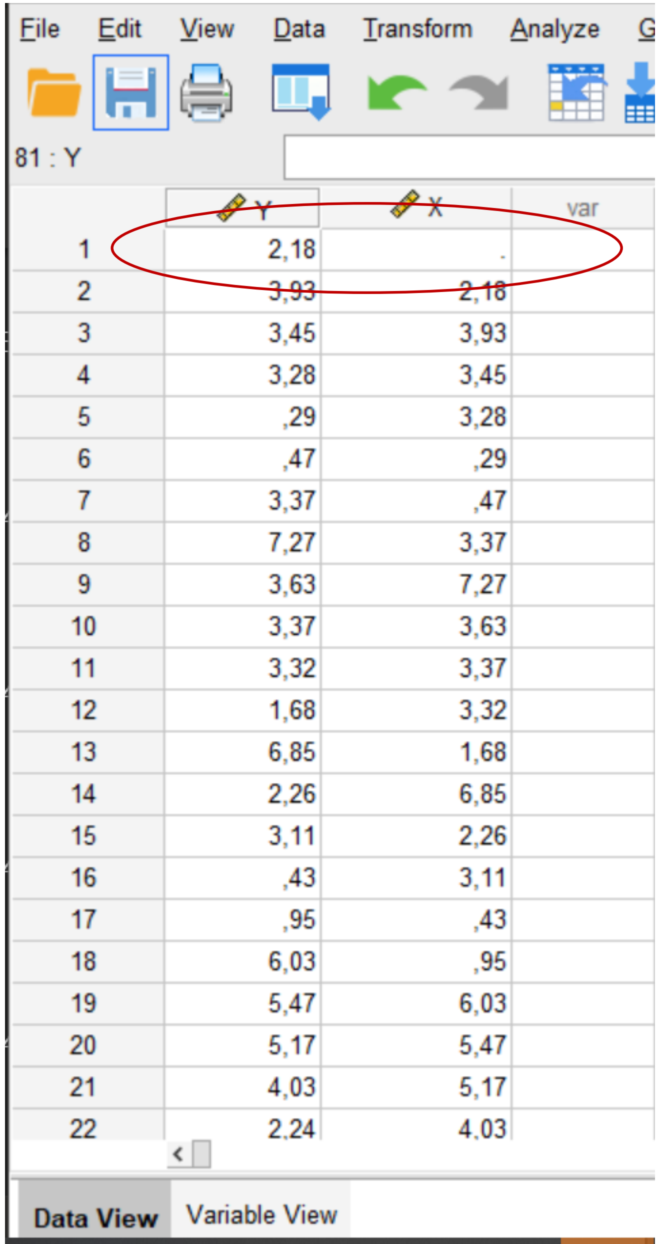
File Edit View Data Transform Analyze G

81 : Y

	Y	X	var
1	2,18	.	
2	3,93	2,18	
3	3,45	3,93	
4	3,28	3,45	
5	,29	3,28	
6	,47	,29	
7	3,37	,47	
8	7,27	3,37	
9	3,63	7,27	
10	3,37	3,63	
11	3,32	3,37	
12	1,68	3,32	
13	6,85	1,68	
14	2,26	6,85	
15	3,11	2,26	
16	,43	3,11	
17	,95	,43	
18	6,03	,95	
19	5,47	6,03	
20	5,17	5,47	
21	4,03	5,17	
22	2,24	4,03	

Data View Variable View

Reversible Change



	Y	X	var
1	2,18	.	
2	3,93	2,18	
3	3,45	3,93	
4	3,28	3,45	
5	,29	3,28	
6	,47	,29	
7	3,37	,47	
8	7,27	3,37	
9	3,63	7,27	
10	3,37	3,63	
11	3,32	3,37	
12	1,68	3,32	
13	6,85	1,68	
14	2,26	6,85	
15	3,11	2,26	
16	,43	3,11	
17	,95	,43	
18	6,03	,95	
19	5,47	6,03	
20	5,17	5,47	
21	4,03	5,17	
22	2,24	4,03	

Reversible Change

$$Y_{it} = b_{0i} + \; b_{1i}Y_{i,t-1} + \; \epsilon_{it}$$

- Couple of things:
  - The intercept is not the mean!

$$\mu = \frac{b_o}{1 - \; b_1^2}$$

  - To get the mean you need to group-mean center the lagged predictor $Y_{i,t-1}$
  - This model assumes there is no trend! If there is, remove it first!!

Reversible Change

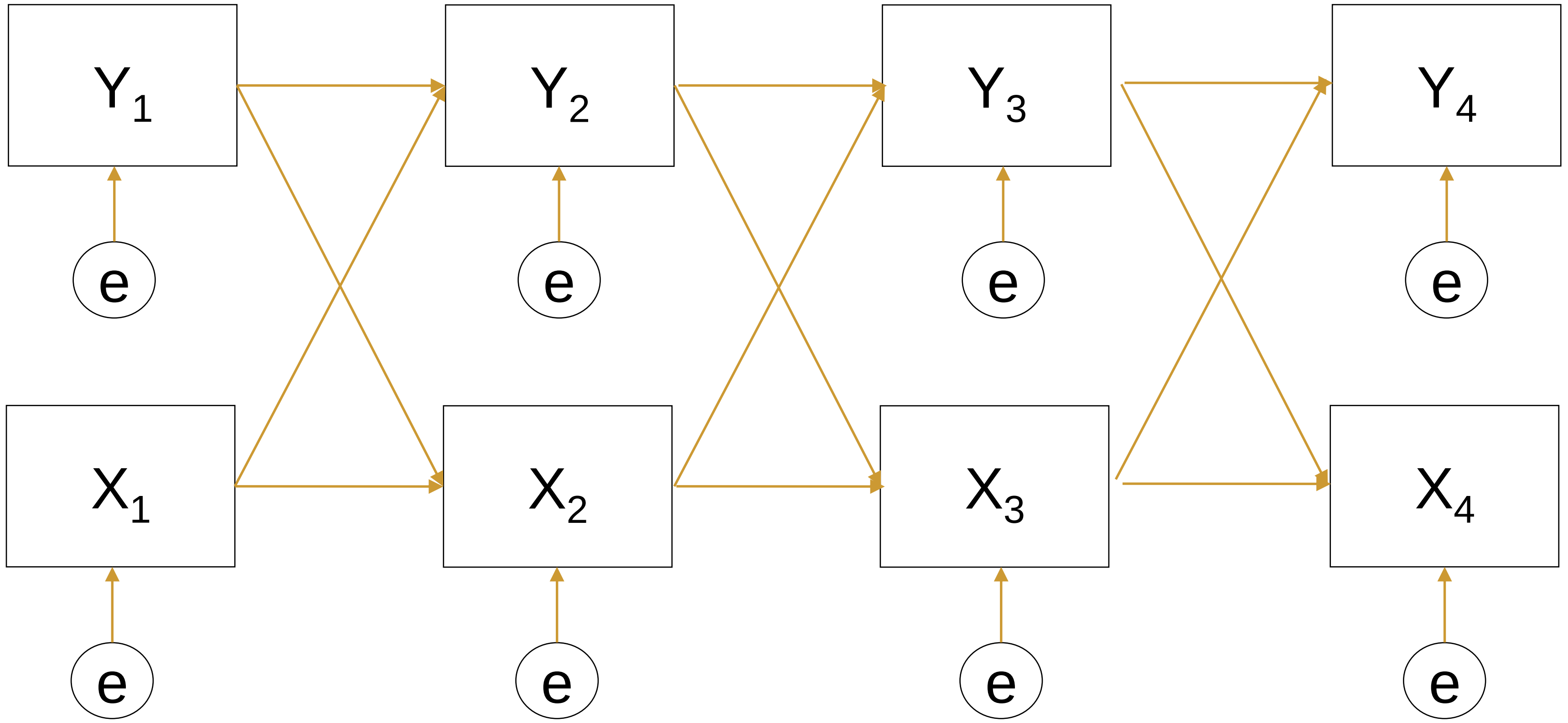
How do you remove the trend?

$$Y_{it} = b_0 + \ b_1 Time_{it} + \delta_{it}$$

$$\delta_{it} = \ b_1 \delta_{i,t-1} + \ \epsilon_{it}$$

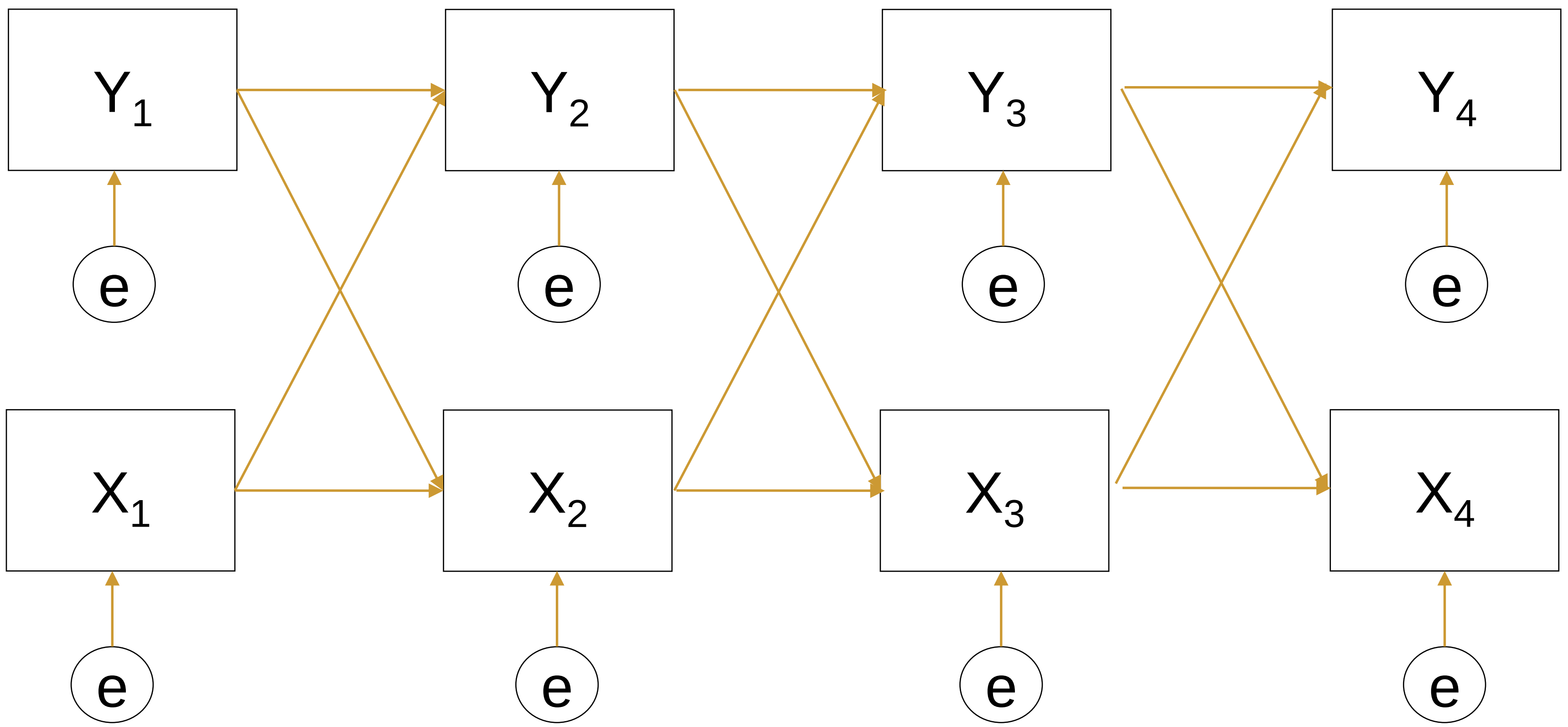
Reversible Change

- Will often be interested in (longitudinal) relation between two or more variables



Reversible Change

- This is called a VAR model



Reversible Change

$$\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} b_{Y0} \\ b_{X0} \end{bmatrix} + \begin{bmatrix} b_{Y1} & b_{XY} \\ b_{YX} & b_{X1} \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{Yt} \\ \epsilon_{Xt} \end{bmatrix}$$

b_{Y0}/b_{X0}= Long run tendency → Think “mean”.

b_{Y1}/b_{X1}= Autoregressive parameter → inertia.

b_{YX}/b_{XY}= Cross-lagged effects.

$\epsilon_{Yt}/\epsilon_{Xt}$= Residual/Innovation → All variation that can not be predicted by previous measurement.

Reversible Change

$$\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} b_{Y0} \\ b_{X0} \end{bmatrix} + \begin{bmatrix} b_{Y1} & b_{XY} \\ b_{YX} & b_{X1} \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{Yt} \\ \epsilon_{Xt} \end{bmatrix}$$

This model is very close to the longitudinal network models you see in the literature!