Multilevel Modelling-course: DAY 2

University of Zurich
PD Dr. Conrad Ziller
conrad.ziller@uni-due.de
https://conradziller.com

14.11.2021

Schedule for Day 2

- What happened so far, and where are we heading?
- Fixed Effects models
- 3-Level models (and more)
- Cross-level interactions and centering
- Mediation

Schedule for Day 2

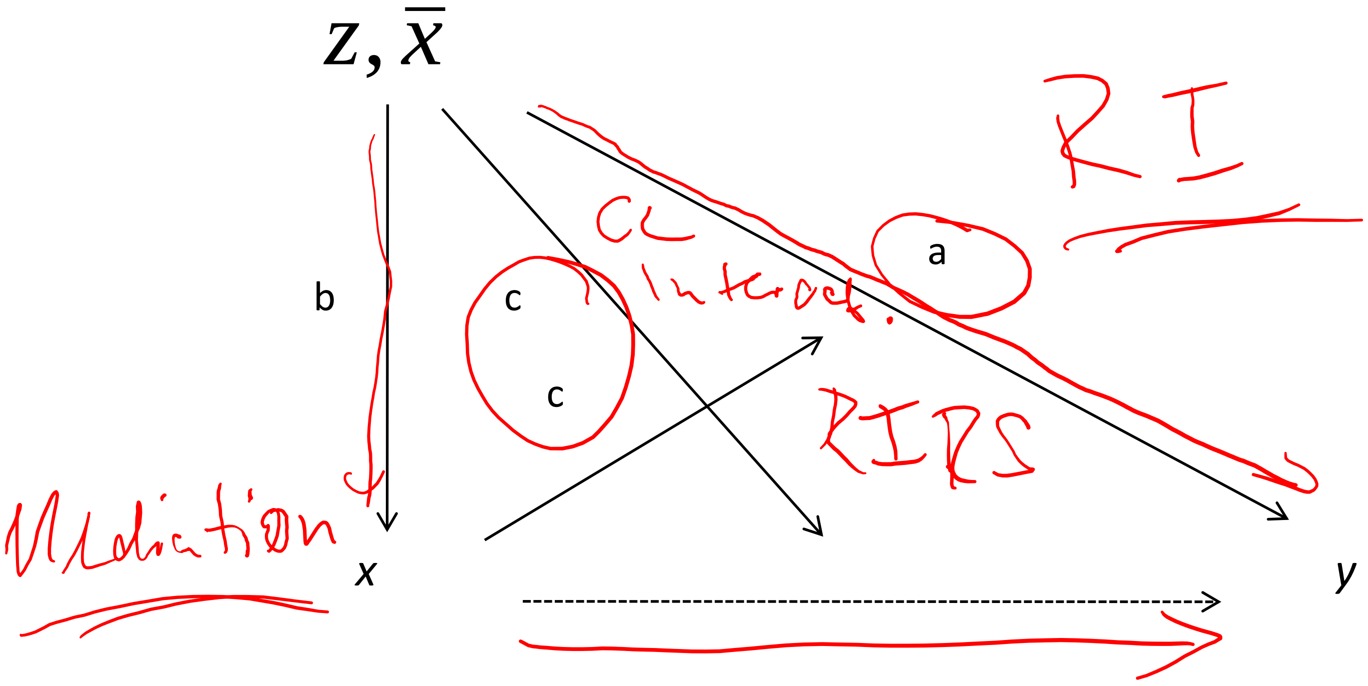
- **What happened so far, and where are we heading?**
- Fixed Effects models
- 3-Level models (and more)
- Cross-level interactions and centering
- Mediation

What happened so far? (1/4)

- We talked about multilevel data structures and why it is important to account for dependent observations
- With dependent observations, OLS…
  - …leads to biased SEs
  - …might lead to biased coefficient estimates
  - …is not able to estimate individual-level ***and*** contextual effects

What happened so far? (3/4)

- We talked about contextual effects and possible variables relationships



What happened so far? (3/4)

- We reviewed typical steps in multilevel analysis and went into their technicalities

1. (Empty model) ICC > 2.5 / 3 %
2. Random Intercept Model, individual variables only
3. Random Intercept Model, individual variables and context variables
4. Random Slopes Model → for lower level var.
5. Random Slopes Model, with cross-level interactions

What happened so far? (4/4)

- We talked about…
  - Model estimation (Maximum Likelihood)
  - Model comparison (Likelihood-Ratio-Tests, AIC, BIC, R^2)
  - Model assumptions (linearity, normally distributed errors, homoscedasticity, and *as always* exogenous predictor variables)

- Tried out some examples in R and Stata

How we proceed (1/2)

- **Important stuff**, we will focus on in the remaining time
  - Fixed effects versus random effects (aka multilevel) models
  - Interactions **(very relevant!)** and centering **(not a big issue)**
  - Practical guide to mediation **(complex topic, so we cover only the basics)**
  - 3- and more levels **(not difficult to implement)** and multilevel models with some fixed effects **(very useful and underated)**

How we proceed (2/2)

- **Relevant stuff** that, however, goes beyond a practical introduction:
  - Cross-classified models **(complex and not often used)**
  - Logistic multilevel regression for binary outcomes **(adds some complexity, often produces substantively similar results compared to a linear model)**
  - Growth curve models **(useful for time-related data)**

Schedule for Day 2

- What happened so far, and where are we heading?
- **Fixed Effects models**
- 3-Level models (and more)
- Cross-level interactions and centering
- Mediation

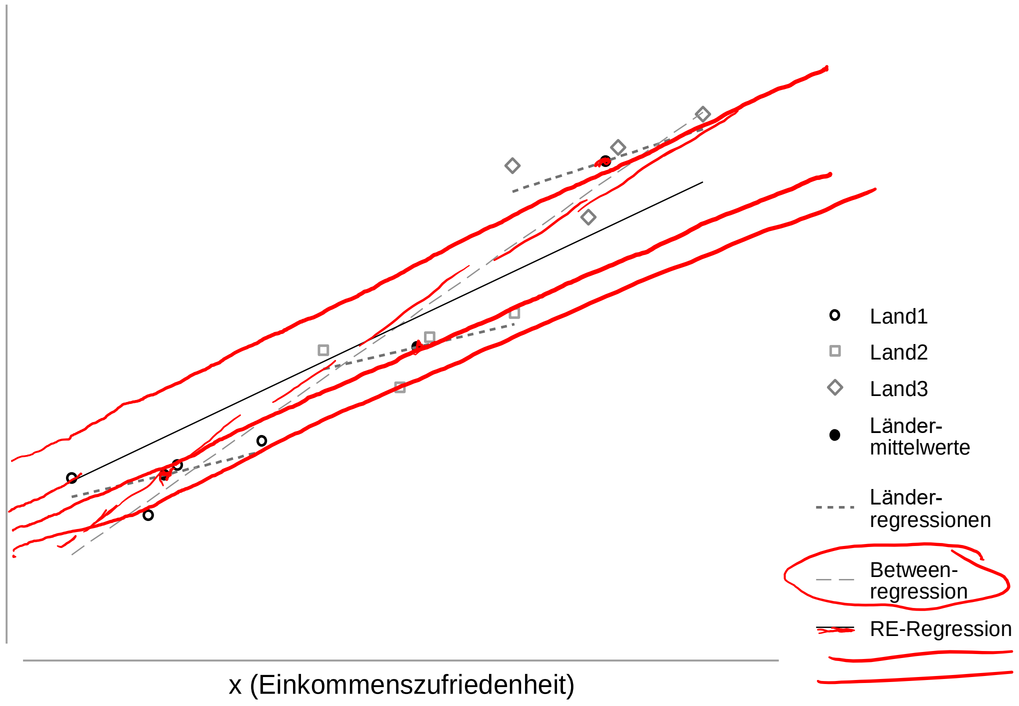
Remedies to statistical dependencies

- Control for factors that causally underlie the process of serial correlation
- Estimation of cluster - / panel-robust standard errors
- **Random effects a.k.a. multilevel models**
- **Fixed effect models**

Random effects / multilevel models

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 z_j + u_j + e_{ij}$$

- Similar to OLS models: estimation based on both between-group and within-group variance, but weighted toward within estimates
- Variance between groups is bound in the error term, which is assumed to be a normally distributed random variable u_j
- Contextual factors (e.g. GDP/pc) can be used to explain variance between groups



Fixed effects model

- **Basic version:** OLS regression with a dummy variable for each group j

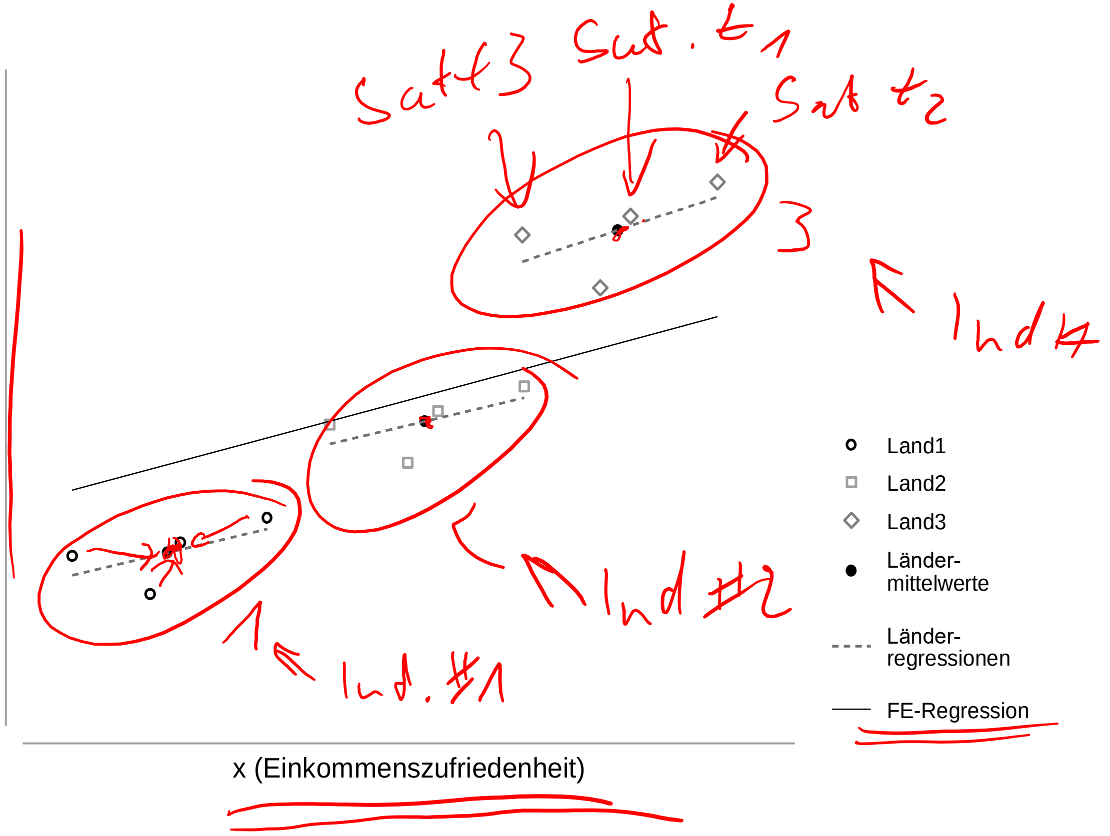
$$y_{ij} = \beta_0 + \sum_{j=1}^{J-1} \beta_j D_{j[i]} + \beta_1 x_{ij} + e_{ij}.$$

Dummy variable: absorbs the complete variance between groups

→ predictor no longer confounded with group differences in the outcome

- **~~Demeaning (same results)~~:** Centering the variables on the respective group average (i.e., the within estimator, as it is based on within variance only)

$$(y_{ij} - \bar{y}_j) = \beta_1 (x_{ij} - \bar{x}_j) + (e_{ij} - \bar{e}_j).$$



Fixed effects model

Pros:

- Avoids biased estimators due to unobserved group-level heterogeneity
- Adjust for serial correlation

Cons:

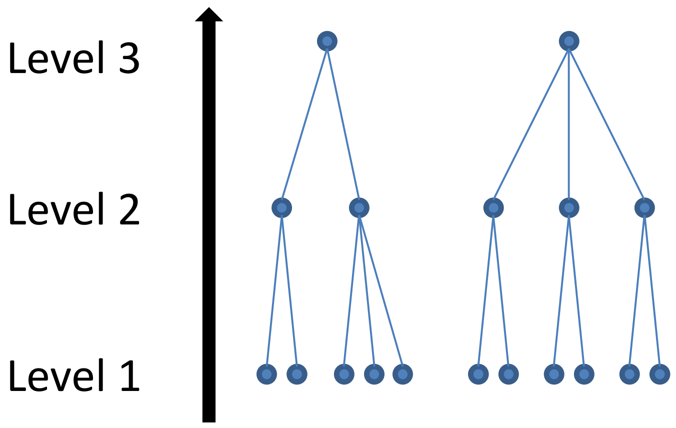
- Group-level variables cannot be included
- Low efficiency (= potentially SEs too high)

Schedule for Day 2

- What happened so far, and where are we heading?
- Fixed Effects models
- **3-Level models (and more)**
- Cross-level interactions and centering
- Mediation

3 – Level Model

- Individuals in regions (L-2) in countries (L-3)
- Pupils in classes (L-2) in schools (L-3)
- Workers in teams (L-2) in organisations (L-3)
- Repeated measurements in individuals (L-2) in regions (L-3)



3-Level Model

Extension of the 2-Level Model

–Random Intercept Modell

$+ \beta_2 z_{ijk}$

$$y_{ijk} = \beta_{0jk} + \beta_{1jk} x_{ijk} + e_{ijk}$$

–Region-in-Country-Intercept

$$\beta_{0jk} = \beta_{0k} + u_{0jk}$$

–Country-Intercept

$$\beta_{0k} = \gamma_{000} + v_{0k}$$

–Overall notation

Cntr Region Ind.

$$y_{ijk} = \underbrace{\gamma_{000} + \beta_1 x_{ijk}}_{\text{Fixed Part / Fixed Effects}} + \underbrace{v_{0k} + u_{0jk} + e_{ijk}}_{\text{Random Part / Random Effects}}$$

3-Level Model

- Variance components (ICC)

Reg

–Level 2 $$\rho = \frac{\sigma_{u_{0jk}}^{2}}{\sigma_e^{2} + \sigma_{u_{0jk}}^{2} + \sigma_{v_{0k}}^{2}}$$ or $$\rho = \frac{\sigma_{u_{0jk}}^{2} + \sigma_{v_{0k}}^{2}}{\sigma_e^{2} + \sigma_{u_{0jk}}^{2} + \sigma_{v_{0k}}^{2}}$$

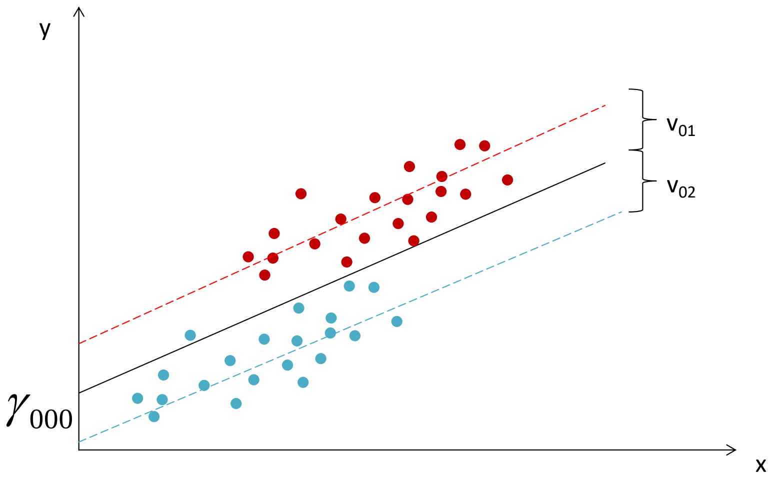
Cntr.

–Level 3 $$\rho = \frac{\sigma_{v_{0k}}^{2}}{\sigma_e^{2} + \sigma_{u_{0jk}}^{2} + \sigma_{v_{0k}}^{2}}$$

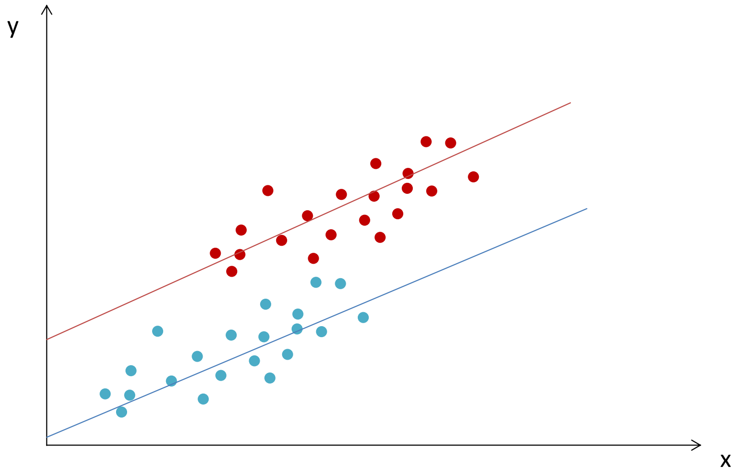
- Complexity of the models grows exponentially with the number of levels
  - Random slopes can occur at all higher levels (individual effect at level-2 and level-3; regional effect at level-3 ...)
  - Increasing number of possible cross-level interactions

Accounting for various levels in multilevel modeling

- Levels can be account for by random intercepts (=level) or unit fixed effects
- With random intercepts, the problem of unobserved heterogeneity remains
- With fixed effects, variance is not modelled, but absorbed
  - Advantage with multiple levels: Multilevel structure at lower levels is taken into account, no distortions due to unobserved heterogeneity (**useful for repeated cross-sectional survey data and an interest in time-varying macro-level predictors, e.g., changes in income inequality over time**)
  - Disadvantage: No macro variables can be estimated at the level of unit fixed effects



Multilevel (random)

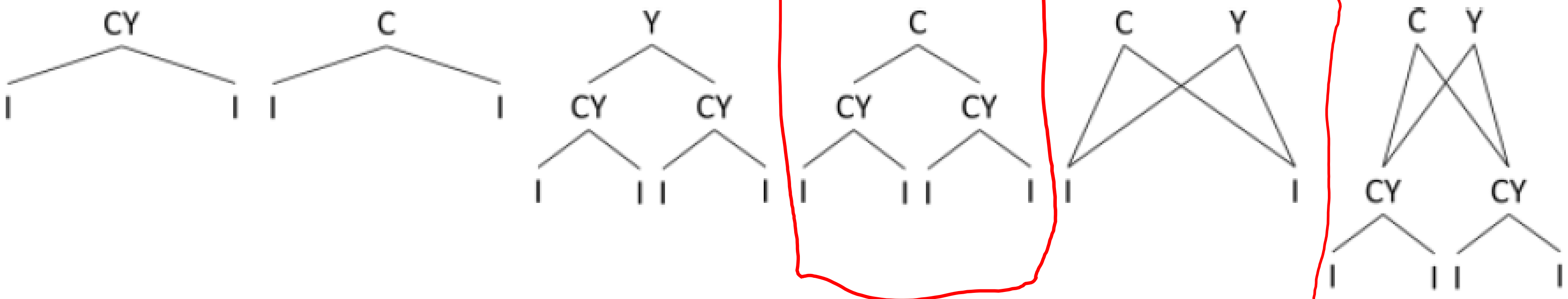


Dummy-Variables (fixed)

Accounting for various levels in multilevel modeling

Table 1. A typology of random effects structures for multilevel models of comparative longitudinal survey data

Random effects	Model A	Model B	Model C	Model D	Model E	Model F
Country		✓		✓	✓	✓
Year			✓		✓	✓
Country-year	✓		✓	✓		✓
Structure						

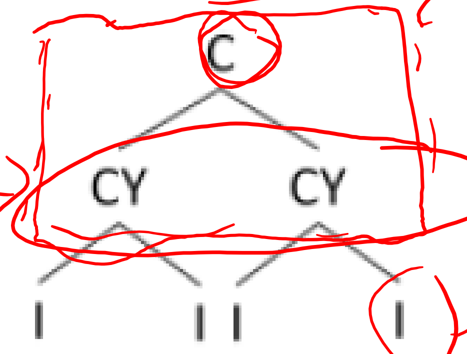


Note: C=country-level RE, Y = year-level RE, CY = country-year-level RE, I = individual level.

Source: Schmidt-Catran & Fairbrother 2016 (Eur Soc Rev): p.25

Accounting for various levels in multilevel modeling

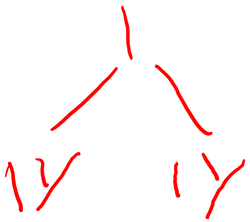
Model D
✓
✓



Plus time and country dummies

With repeated cross-sectional survey data (e.g., ESS, ISSP) and a substantial interest in estimating unbiased effects at the macro level:

- **Random intercept at the level of country-years** (and countries) to account for serial correlation due to clustering at the level of respondents (nested in country sampling points)
- **Time dummies** (=time FEs) to account for serial correlation due to time and common trends in the DV
- **Country dummies** (=unit FEs) to account for unobserved heterogeneity related to country differences (e.g., institutional or historical differences) → makes estimates more reliable



Group session #1

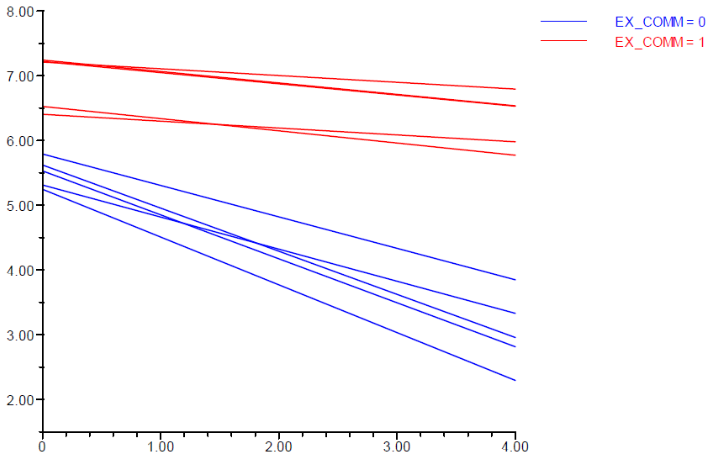
- You will be assigned to breakout rooms in three groups
- Each group scans through a text with complex multilevel structures, esp. the methods section in order to answer the following questions
  1. What is the outcome, and at which level is it measured
  2. What is/are the core predictor variable(s), and at which level is it/ are they measured
  3. How many levels exist and how is does the empirical design account for (different versions of) clustering? Please draw a diagram.
  4. Is unobserved heterogeneity addressed? If yes, how?

Schedule for Day 2

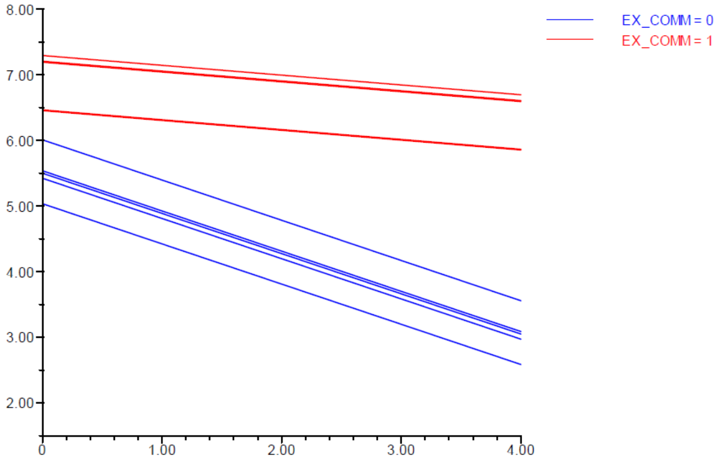
- What happened so far, and where are we heading?
- Fixed Effects models
- 3-Level models (and more)
- **Cross-level interactions and centering**
- Mediation

Cross-Level-Interactions

- Interaction between context and individual variable
- Idea: Context feature explains different effect of individual feature across contexts (random slope)
- A priori (usually) identify significant random slope



Interaction with RS



Interaction without RS

Cross-Level-Interactions

- Specification of country-specific intercepts and slopes

$$y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}z_j + u_{0j}$$

$$\cancel{\beta_{1j} = \gamma_{10} + u_{1j}}$$

- With interaction:

$$\beta_{1j} = \gamma_{10} + \boxed{\gamma_{11}z_j} + u_{1j}$$

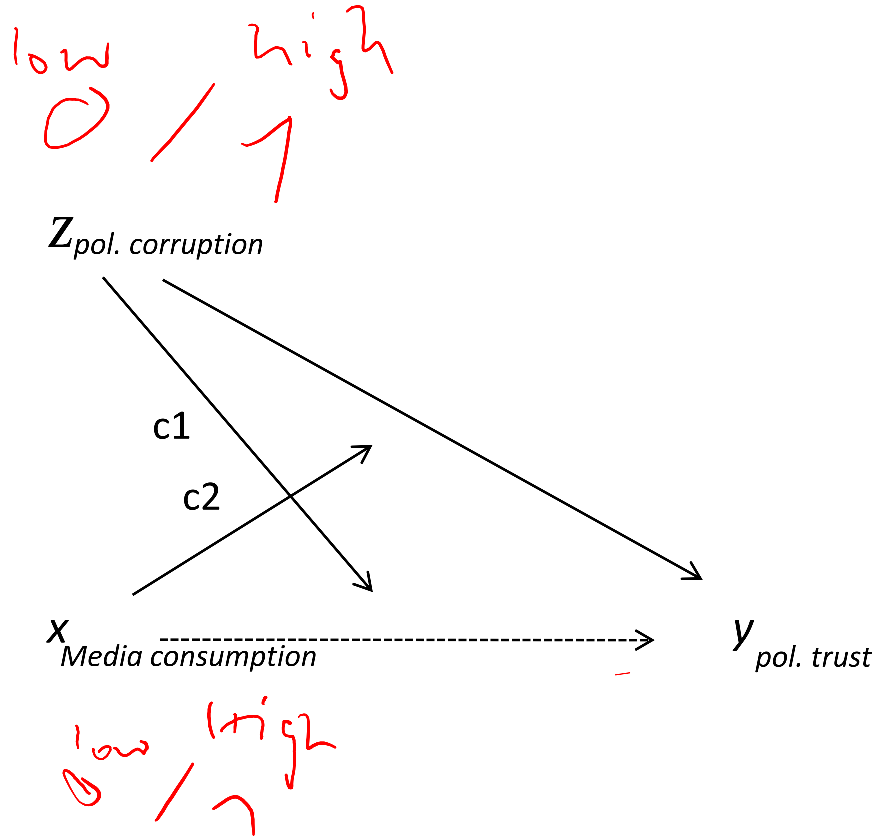
- Overall notation

RS RI

$$y_{ij} = \gamma_{00} + \underbrace{\gamma_{10}x_{ij} + \cancel{\gamma_{01}}z_j + \gamma_{11}x_{ij}z_j}_{\text{Fixed Part / Fixed Effects}} + \underbrace{u_{1j}x_{ij} + u_{0j} + e_{ij}}_{\text{Random Part / Random Effects}}$$

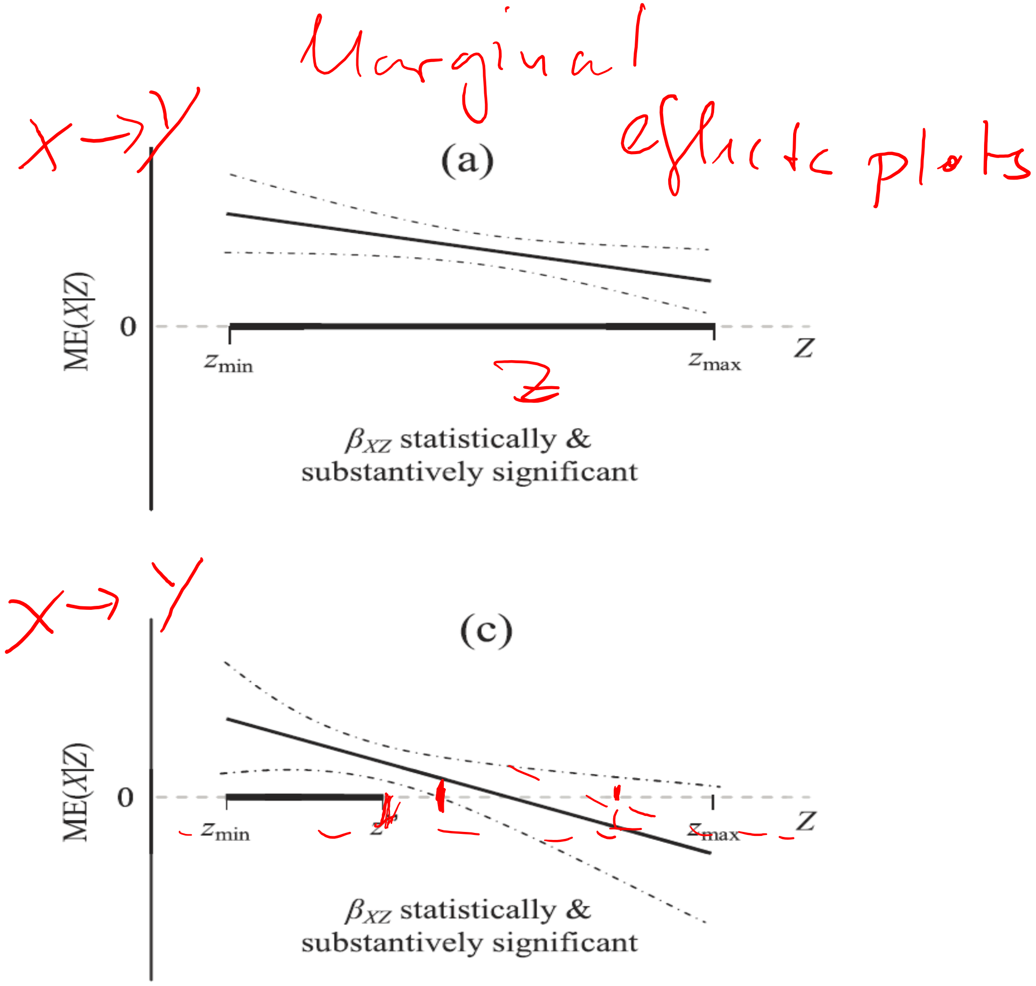
Cross-Level-Interactions

- Hypotheses should be symmetrical (if theoretically reasonable)
  - H1: Political corruption reduces political trust, especially when media consumption is high. (c2)
  - H2: Media consumption reduces political trust, especially in contexts characterized by high political corruption. (c1))
- Statement on high and low values of the moderator often useful
  - H2a: In contexts with ***high*** political corruption, the relationship between media consumption and political trust is ***negative***.
  - H2b: In contexts with ***low*** political corruption, the relationship between media consumption and political trust is ***positive***.



Cross-Level-Interactions

- Always estimate both interaction term and unconditional effects in the model
- Symmetrical interpretation of the results along the hypotheses formulated
- Is the correlation significant for all values of the moderator? If not, what does this mean for the hypotheses?



Berry et al. 2012: 661

Stata

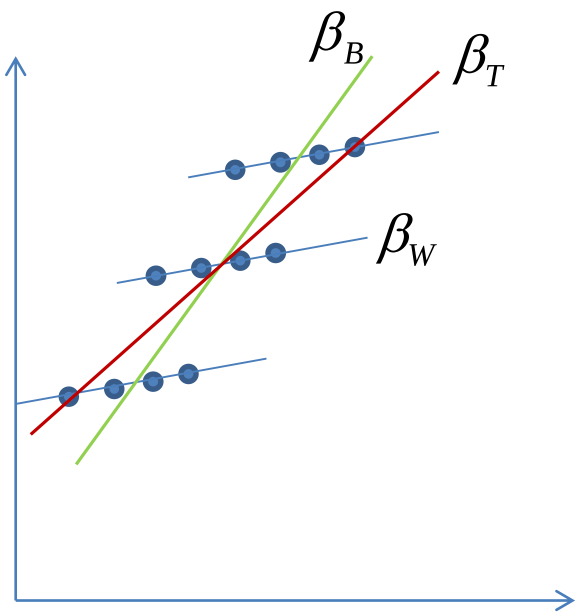
```
//Form interaction term beforehand
gen x_z=x*z
mixed y x z x_z ||id:


//better interact within the model per # (simplifies post-
estimation)
mixed y x z c.x#c.z ||id:

margins, dydx(x) at(z = (0 (0.1) 1))
marginsplot,  yline(0)
margins, dydx(z) at(x = (0 (0.1) 1))
marginsplot,  yline(0)
```

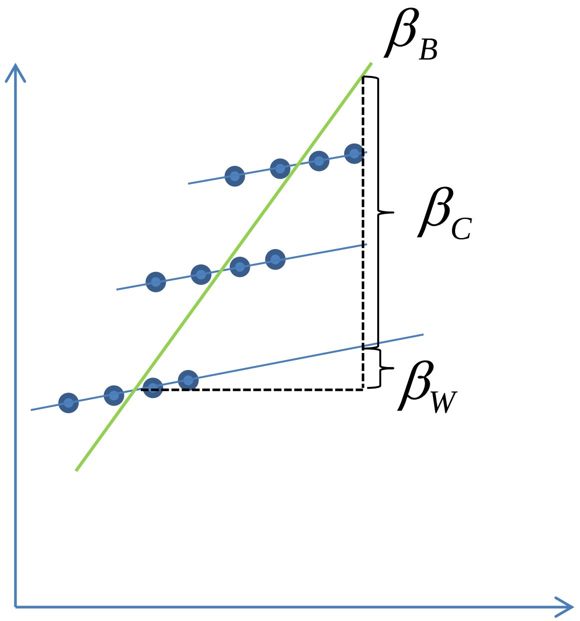
Centering

- Version 1: At the overall mean or grand mean (also CGM)
  - Continuous variables receive a meaningful zero point
  - Constant can be interpreted as an estimated value of the outcome for persons who have a mean expression on all characteristics
  - All other parameter estimates remain identical (as in the uncentred case)
  - Possibly reduces multicollinearity problems

- Version 2: At the group mean (also CWC)
  - Only for Level-1 variables
  - Consequence: Between-group variance is removed from variable
  - Constant is an estimate of the outcome for individuals who have all characteristics corresponding to their group mean
  - Changes correlation structure of the data; interpretation of coefficient estimates differs for level-2 variables

Context Effect of Group Mean





blue = Within-Effect
green = Between-Effect
red = Total Effect (mixed)

Between-Effect = Within-Effect
+ context effect

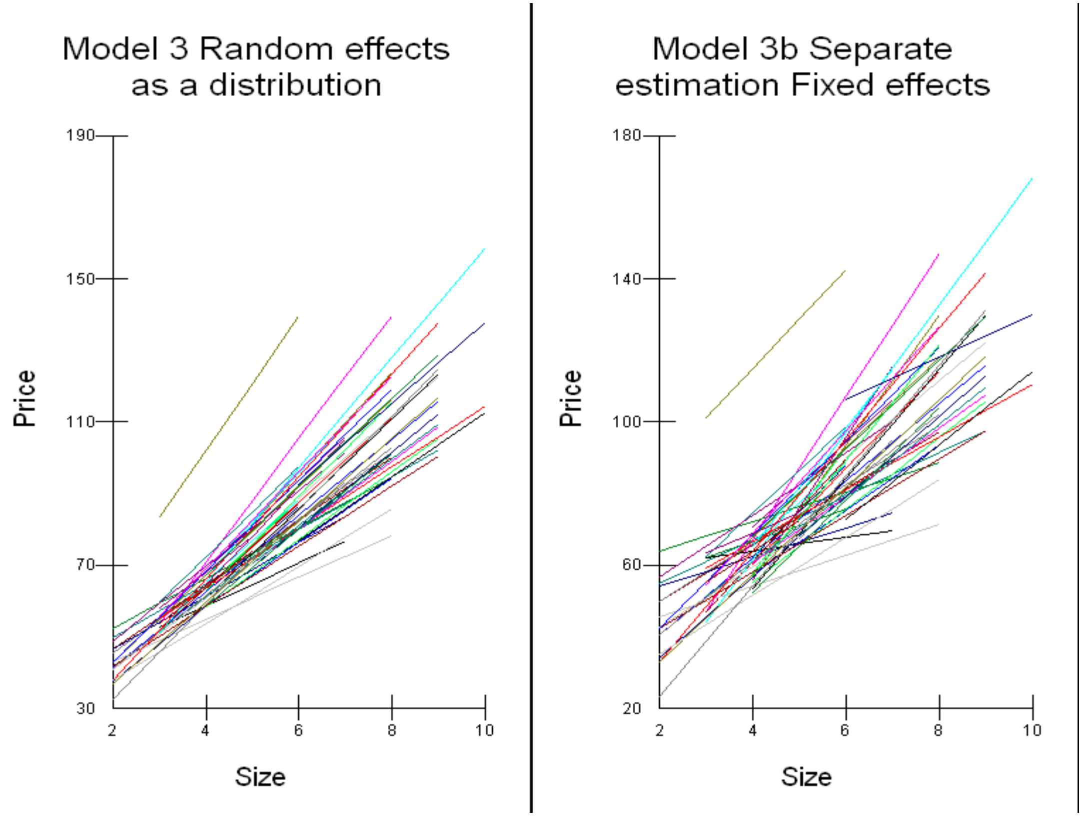
Centering and Effect Interpretation

- Centering at the Grand Mean (CGM)
  - –Effect of x = β(within)
  - –Effect of $\bar{x}$ (i.e., the group mean) = β(context)
  - –β(between) = β(within) + β(context)

- Centering at the Group Mean (CWC)
  - –Effect of x = β(within) *[without „shrinkage"]*
  - –Effect of $\bar{x}$ (i.e., the group mean) = β(between)
  - –β(context) = β(between) - β(within)

Shrinkage

- Regression lines across individual contexts "shrink" towards the Grand Mean
- The fewer the observations at level 1, the stronger the shrinkage
- Certainty of estimation is to some extent borrowed from large clusters



from: Jones & Subramanian 2009, MlWin Traning Manual

Centering

- **When to center on the overall mean?**
  - Mostly useful, as it facilitates interpretation of the concept and context effects.
  - If substantial interest in level 2 variable (or interaction at level 2), as it co-controls for composition effects of level 1.

- **When to center on the group mean?**
  - Substantive reasons: In the case of poorly comparable group means (e.g. center income at the country mean).
  - If there is substantial interest in level 1 variables, as there is no shrinkage
  - In the case of cross-level interaction → practically usually no difference compared to no centering
  - For level-1 interactions

Stata

```
ssc install center

//grand-mean
center UV1 UV2 , pre(cgm_) mean(mgm_)

//group-mean
bys id: center UV1 UV2 , pre(cwc_) mean(mwc_)


xtmixed AV cgm_UV1 cgm_UV2 … ||id:
xtmixed AV cwc_UV1 cwc_UV2 … ||id:
```

Group session #2

Last session, I asked you to prepare a description an interaction you are interested in. Introduce your interaction in pairs and discuss:

- Graphical representation
- What are the underlying hypotheses?
- What is an appropriate test, how would you proceed?

- Which problems have you encountered (share them with others and me)?

Schedule for Day 2

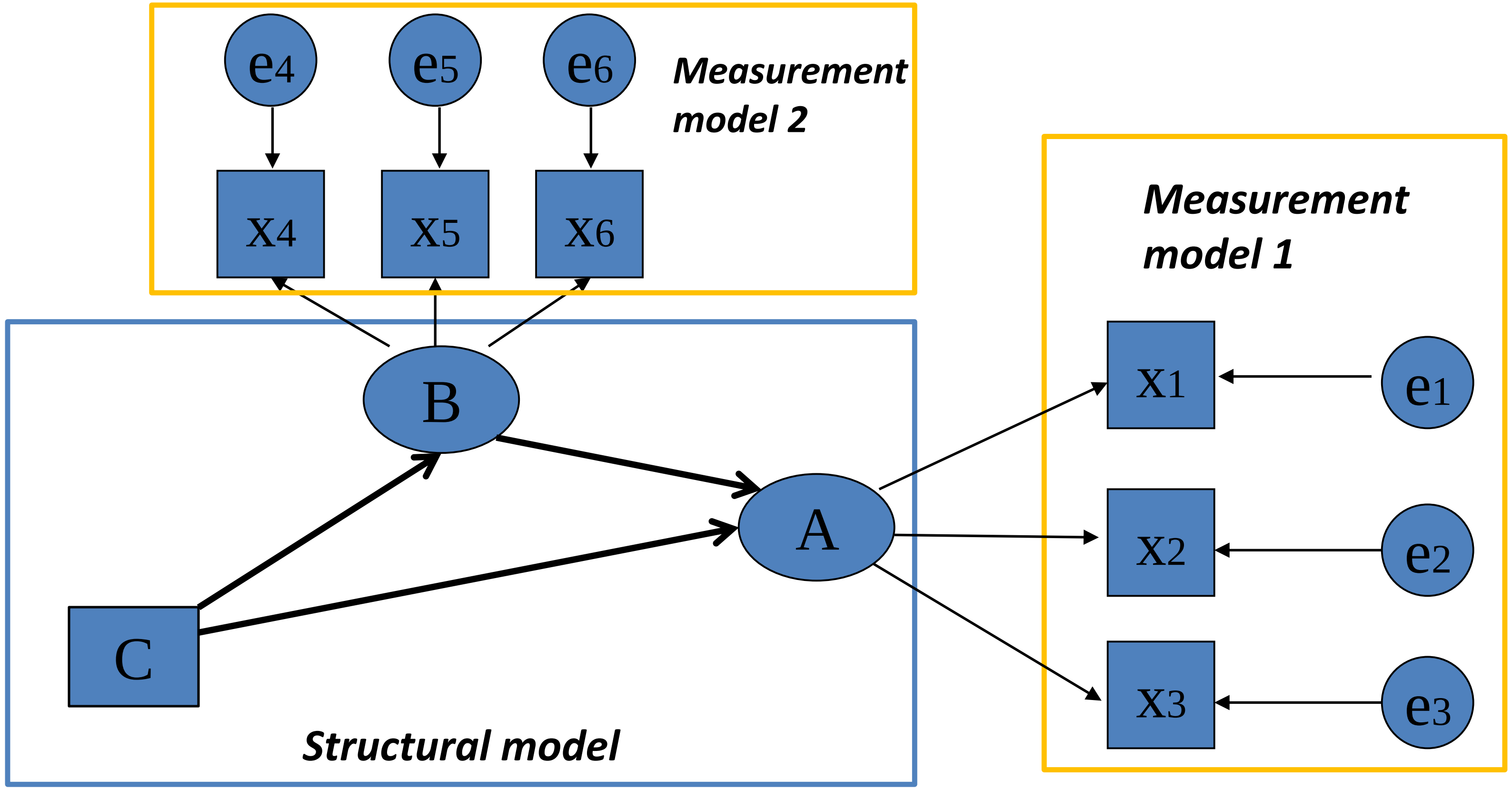
- What happened so far, and where are we heading?
- Fixed Effects models
- 3-Level models (and more)
- Cross-level interactions and centering
- **Mediation**

Structural Equation Models

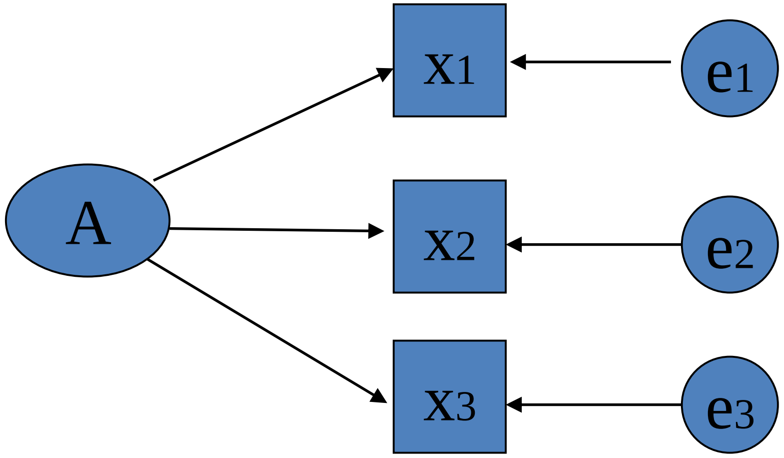
Structural Equation Models (= SEM)

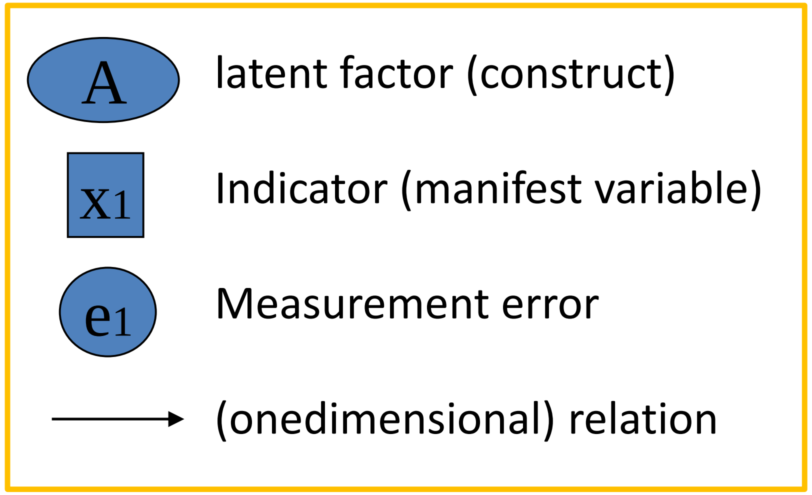
- consist of systems of equations
- can be divided into measurement models and structural models
- direct and indirect effects can be distinguished (mediation)
- distinguish between latent (non-measured) and manifest (measured) variables
- typical graphical representation of the models

Structural Equation Models

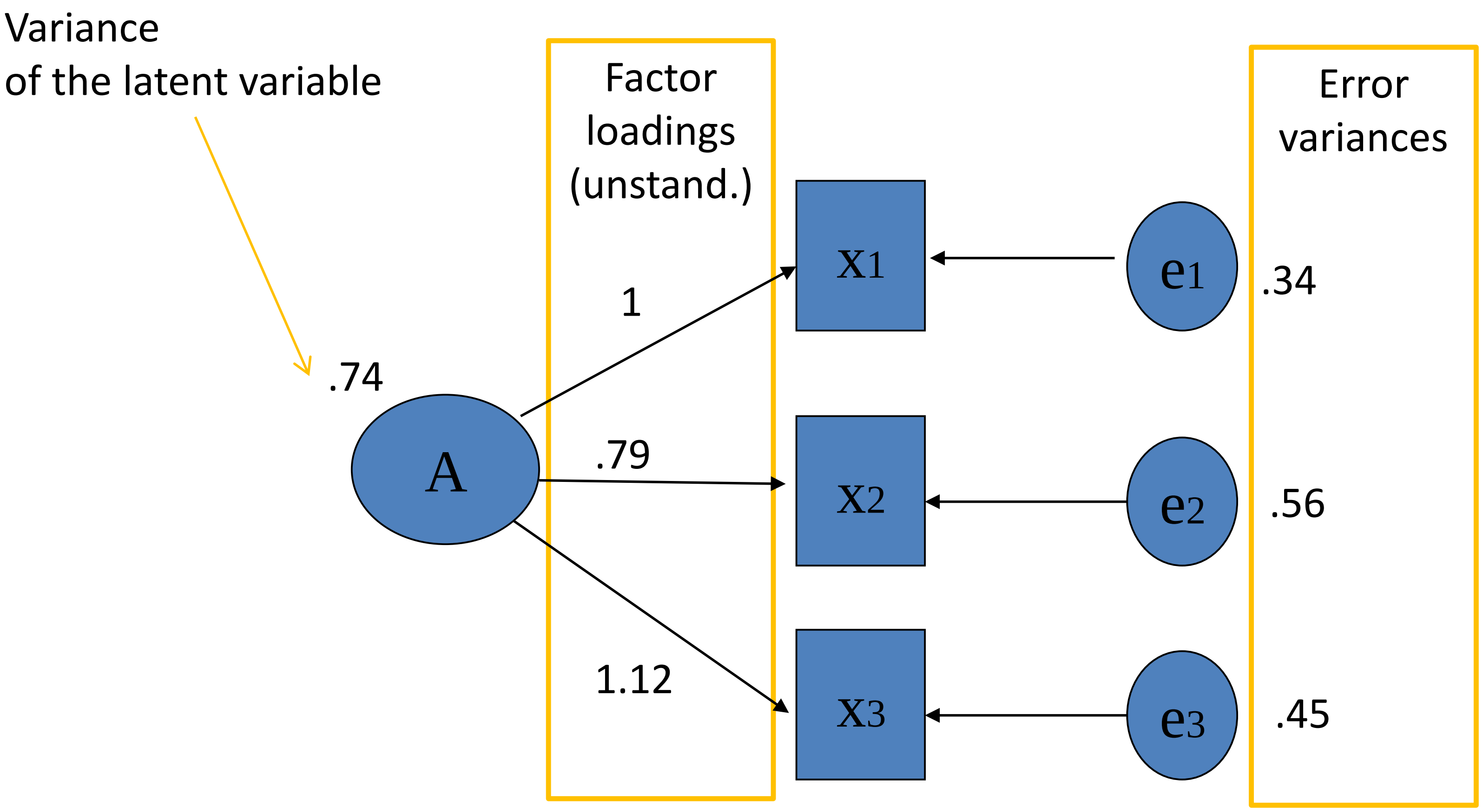


Measurement Model with one Factor

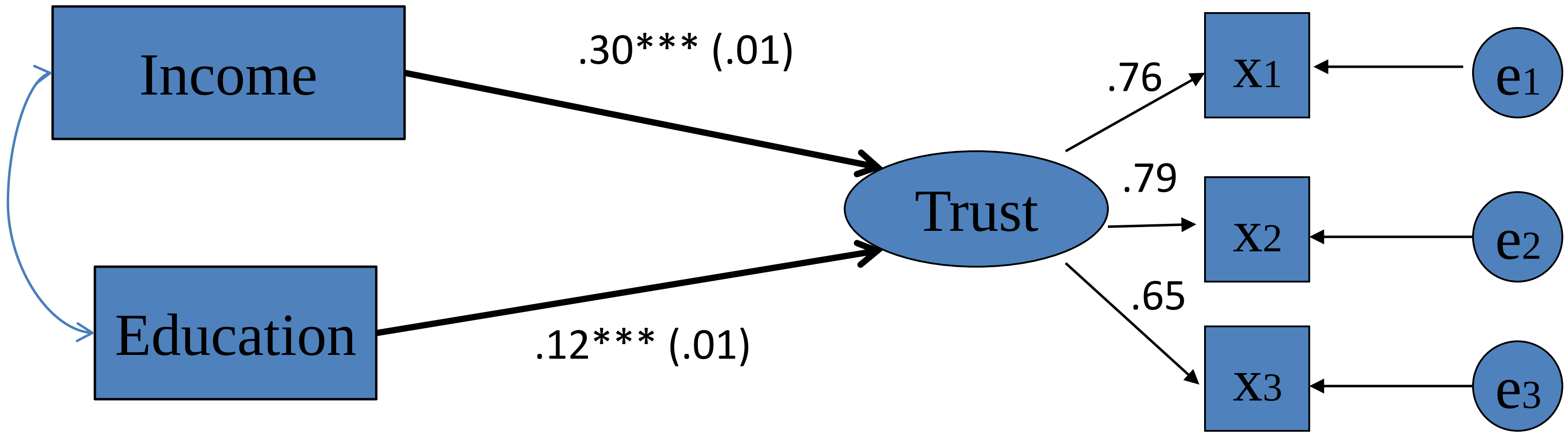




Measurement Model: Parameter (Loadings and variances)

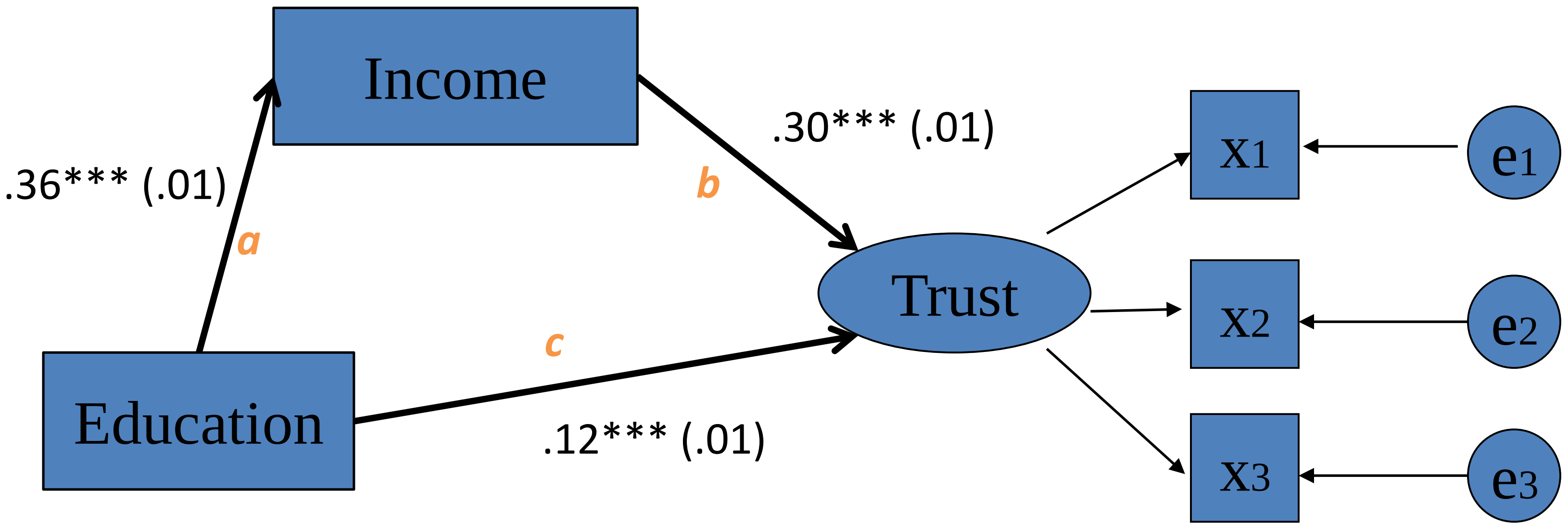


Example 1: Social Trust



Example 2: Social Trust

Indirect = .11*** (.01)
Total = .23*** (.00)

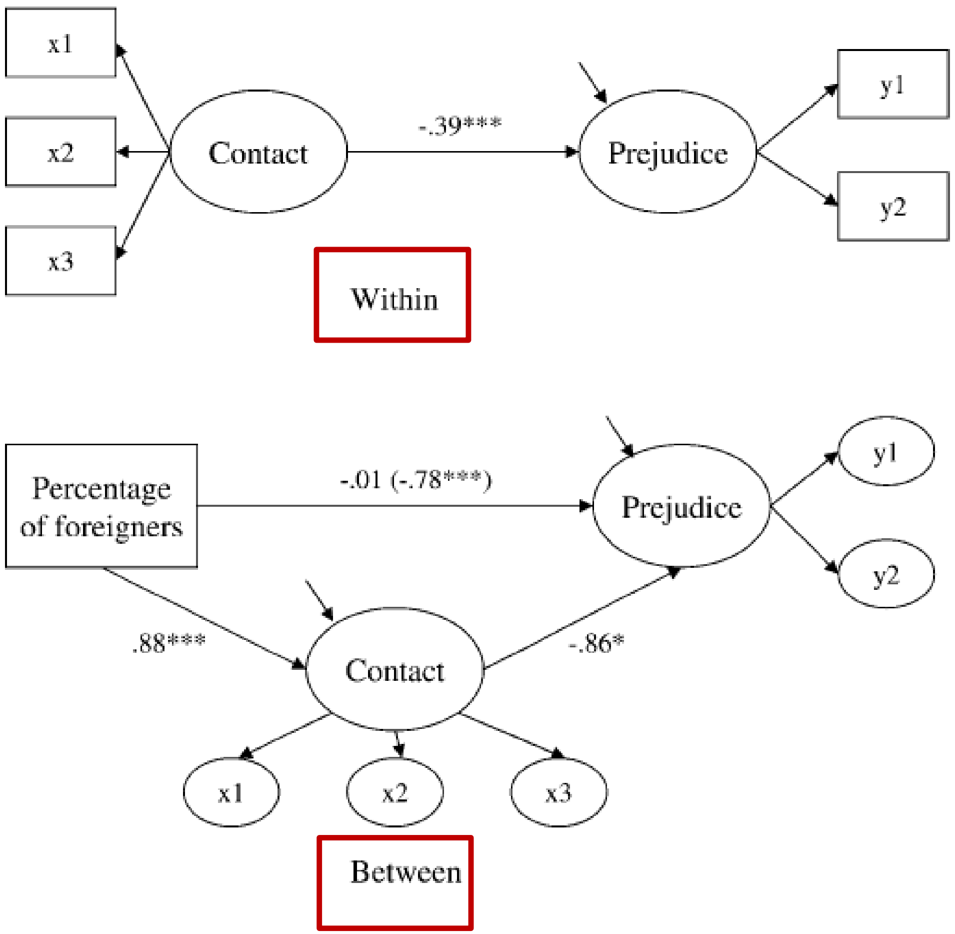


*Indirect effect of education = a * b*
*Total effect of education= indirect + direct effect or a * b + c*

Stata

```
* Example 1
sem (ppltrst pplfair pplhlp <- Trust) ///
(Trust <- income educyears) , latent(Trust) stand
estat gof , stats(all) // shows fit-indices


* Example 2
sem (ppltrst pplfair pplhlp <- Trust) ///
(Trust <- income educyears) ///
(income <- educyears) , latent(Trust) stand
estat teffects, stand //shows indirect effect
estat gof , stats(all)
```

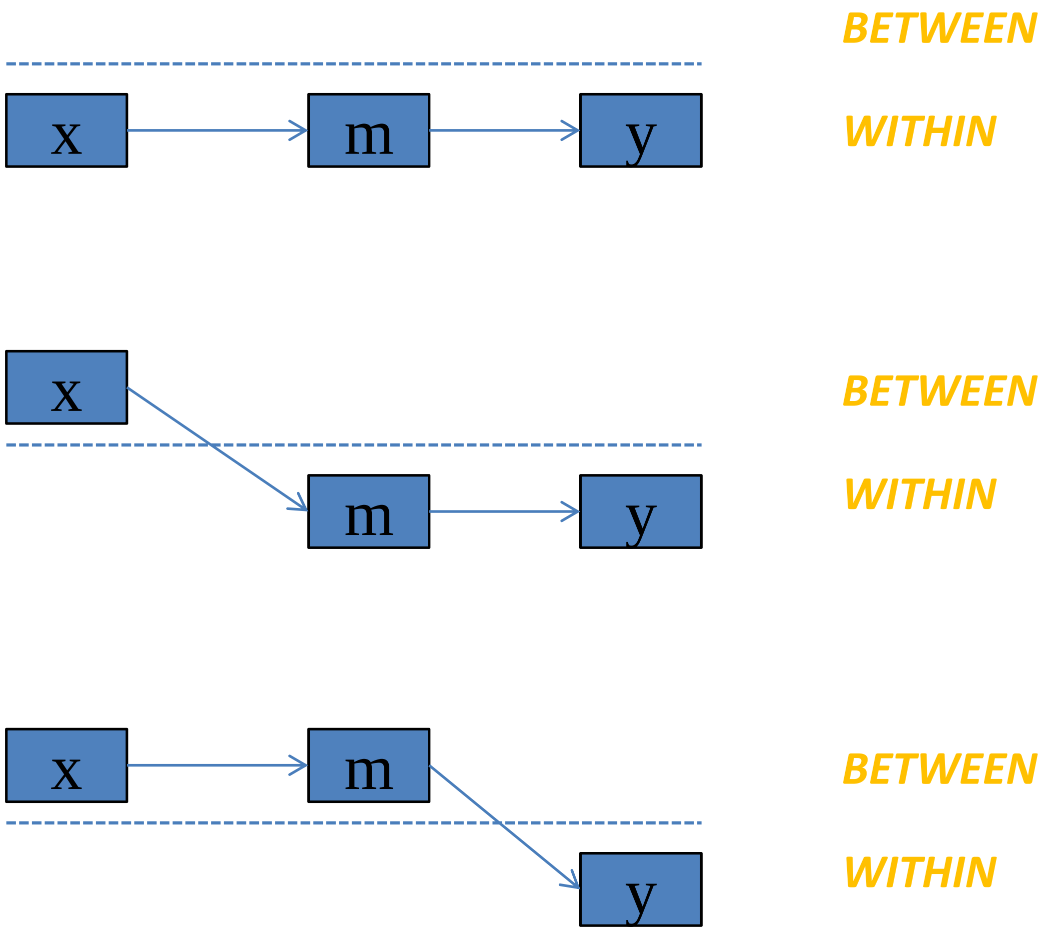
Example Multi-Level SEM



Wagner, U., Christ, O., Pettigrew, T.F., Stellmacher, J., & Wolf, C. (2006). Prejudice and minority proportion: Contact instead of threat effects. *Social Psychology Quarterly*, *69*, 380-390.

Typical Multi-Level SEM-Mediationmodel

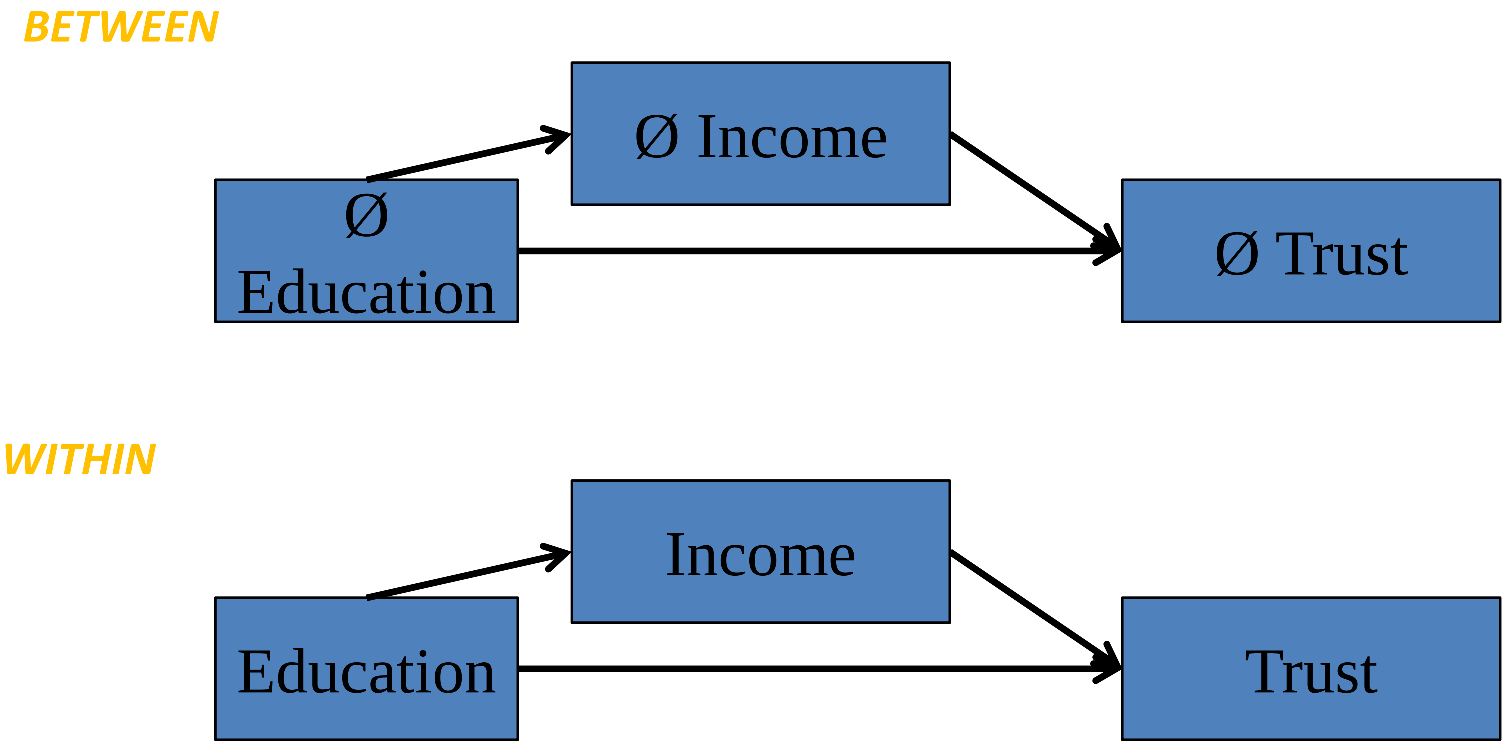
- 1-1-1
- 2-1-1
- 2-2-1



Multi-Level SEM

- Without latent variables or mediation, multilevel regression and multilevel SEM produce equivalent results
- Latent measurement models (CFA) only reliable with large number of clusters (> 60)
- With multilevel mediation and level-2 involved, the between-level is one that is interpreted (Preacher et al. 2010)
- Mplus has enormous advantages over Stata in specifying ML-SEM models; Mplus syntax see http://www.quantpsy.org/pubs/syntax_appendix_081311.pdf
- In R: lavaan package mimicks Mplus; Stata: gsem (slow and often does not converge)

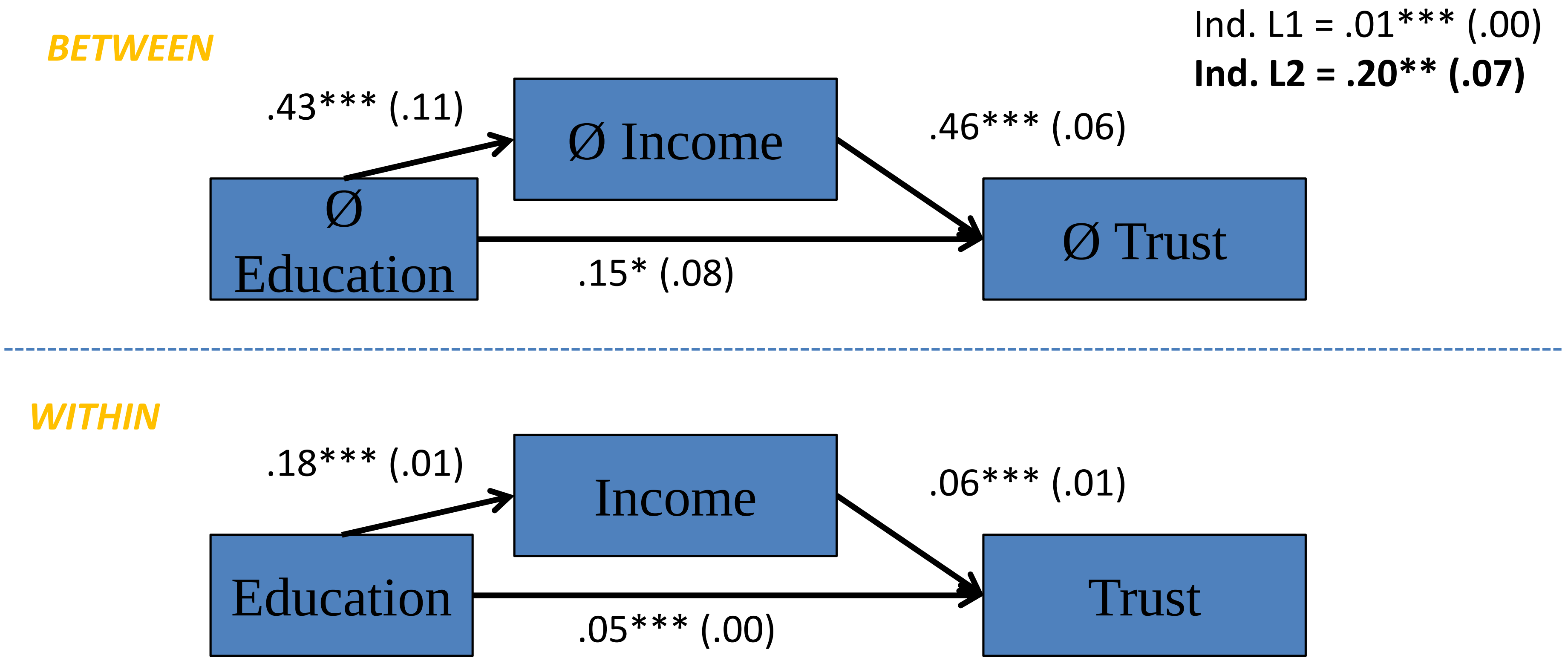
Example 3: Social Trust on Individual and Country-Level



Mplus – Output

```
MODEL RESULTS

                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Within Level

 TRUST      ON
    INCOME             0.056      0.011      5.286      0.000
    EDUCYEARS          0.046      0.004     11.346      0.000

 INCOME     ON
    EDUCYEARS          0.177      0.009     20.568      0.000

 Residual Variances
    TRUST              3.063      0.157     19.554      0.000
    INCOME             3.487      0.193     18.066      0.000

Between Level

 TRUST      ON
    INCOME             0.458      0.062      7.443      0.000
    EDUCYEARS          0.148      0.076      1.960      0.050

 INCOME     ON
    EDUCYEARS          0.429      0.113      3.801      0.000

 Intercepts
    TRUST              0.669      0.870      0.769      0.442
    INCOME             1.022      1.342      0.761      0.447

 Residual Variances
    TRUST              0.212      0.048      4.462      0.000
    INCOME             1.646      0.614      2.681      0.007

 New/Additional Parameters
    ITRUST             0.010      0.002      5.752      0.000
    ITRUST2            0.196      0.070      2.814      0.005
```

Example 3: Social Trust on Individual- and Country – Level



Group session #3

Come together in groups and find examples for all three types of multilevel mediation:

Predictor-Mediator-Outcome

L1-L1-L1

L2-L1-L1

L2-L2-L1

Thank you for your Attention!