VERGLEICHENDE SOZIALFORSCHUNG MIT MEHREBENENMODELLEN IN R

Forschungspraktikum I und II
Dr. Christian Czymara
Cross-level interactions

AGENDA

- Same level and cross-level interactions
- Grand-mean and group-mean centering
- Tutorial: Influence of political elites on attitudes toward Muslim immigrants

MODELLING INTERACTION EFFECTS

MARGINAL EFFECTS

- We are always interested in the association between a *change in* x and an associated *change in* y
- In econometrics, this is called a *marginal effect*
- Technically, the marginal effect is the slope, so the first derivate of regression equation w. r. t. x_1
- For linear effects: $y = \beta_0 + \beta_1 x_1$
- this is simply the corresponding coefficient: $\frac{\partial y}{\partial x} = \beta_1$
- But for some other effects, it is not (e. g., quadratic, logarithmic, or interaction terms)

INTERACTIONS

- Interactions are multiplicative terms:
- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \boxed{+ \beta_3 x_1 x_2}$
- First derivate with respect to x_1:
- $\frac{\partial y}{\partial x_1} = \beta_1 + \beta_3 x_2$

INTERPRETING INTERACTION EFFECTS

- $\frac{\partial y}{\partial x_1} = \beta_1 + \beta_3 x_2$
- Effect of x_1 on y depends on level of x_2 and vice versa (again: *conditional marginal effect)*
- β_1 thus has to be the effect of x_1 on y when x_2 is 0 (and vice versa)
- This is why it often makes sense to mean-center variables

EXAMPLE

- Research question: Do origin effects on trust in the police among immigrants depend on length of stay?
- Data: Pooled version of ESS 3-9
- Variables:
  - Institutional improvement (difference between rule of law in origin and destination country, `ruleoflaw_diff`)
  - Time since migration (five categories, `livecnty_comb`)
  - Both assumed continuous for simplicity
- Several controls
- Simplified version of Czymara & Mitchell (2021)

EXAMPLE

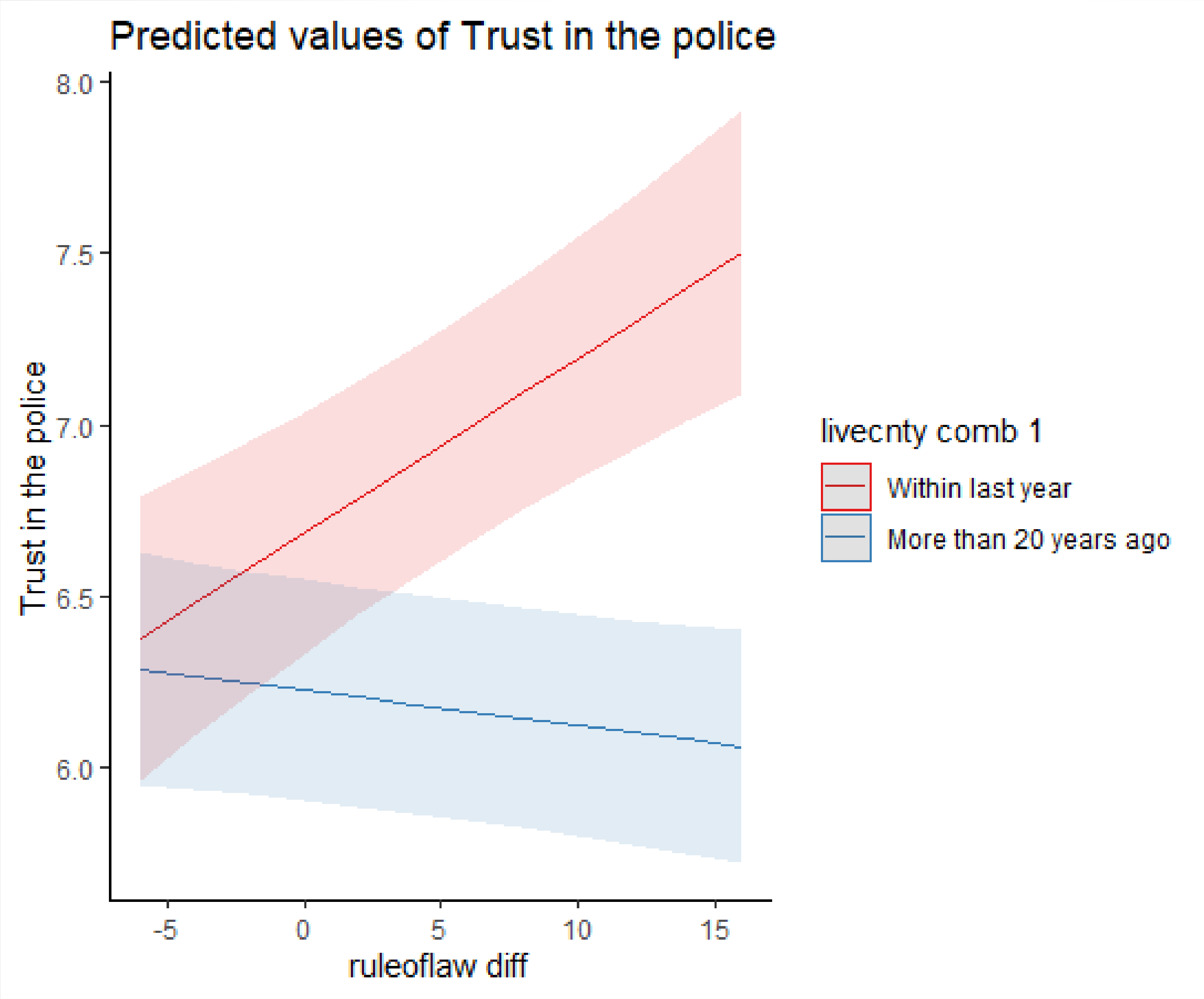
- $trstplc = \beta_0 + \beta_1 ruleoflaw\_diff + \beta_2 livecnty\_comb + \beta_3 ruleoflaw\_diff * livecnty\_comb$
- In this model, the effect of migrating into contexts with better working institutions (β_1) can vary by time since migration:
  - Main effect of institutional improvement (β_1)
  - Interaction effect of institutional improvement and time since migration (β_3)
- The effect of institutional improvement is, thus, the combination of β_1 and β_3

EXAMPLE

Variables	Coefficient
Institutional improvement	0.05 ***
	(0.01)
Time since migration	-0.11 ***
	(0.01)
Institutional improvement X Time since migration	-0.02 ***
	(0.01)

INTERPRETATION

- The relationship between institutional improvement and trust in the police depends on time since migration
- It is about 0.05 for those who migrated recently (`livecnty_comb`=0), slope of red line
- But the association between institutional improvement and trust decreases by -0.02 with each time since migration category
- For those in category 4 (“more than 20 years ago”), this association even becomes negative: 0.05-0.02*4=-0.03 (but not statistically significant)
- In other words: Immigration experiences matter mainly for those who migrated recently


Predicted values of Trust in the police
8.0
7.5
7.0
6.5
6.0
Trust in the police
-5
0
5
10
15
ruleoflaw diff
livecnty comb 1
Within last year
More than 20 years ago

INTERACTIONS WITH MULTI-NATIONAL DATA

INTERACTIONS FOR MULTI-NATIONAL RESEARCH

- In comparative research, the variables of an interaction can be on the same or different levels
- Example: y = income
  - Do the returns to education (x_1) differ for men and women (x_2)? → pure individual-level interaction
  - Does the welfare state type (z_1) have a stronger effect in richer countries (z_2)? → pure country-level interaction
  - Do the returns to education (x) depend on a country's economic wealth (z)? → *Cross-level* interactions

CROSS-LEVEL INTERACTIONS FOR RANDOM SLOPE MODELS

- What might explain why effects of individual-level variables differ across countries?
- Are there country-level characteristics that can explain this varying effect?
- Can be tested by adding an interaction between variable with random slope (the individual-level effect varying over countries) and a country-level variable
- In other words, which country characteristic accounts for the cross-country variation in the individual-level effect?

EXAMPLE

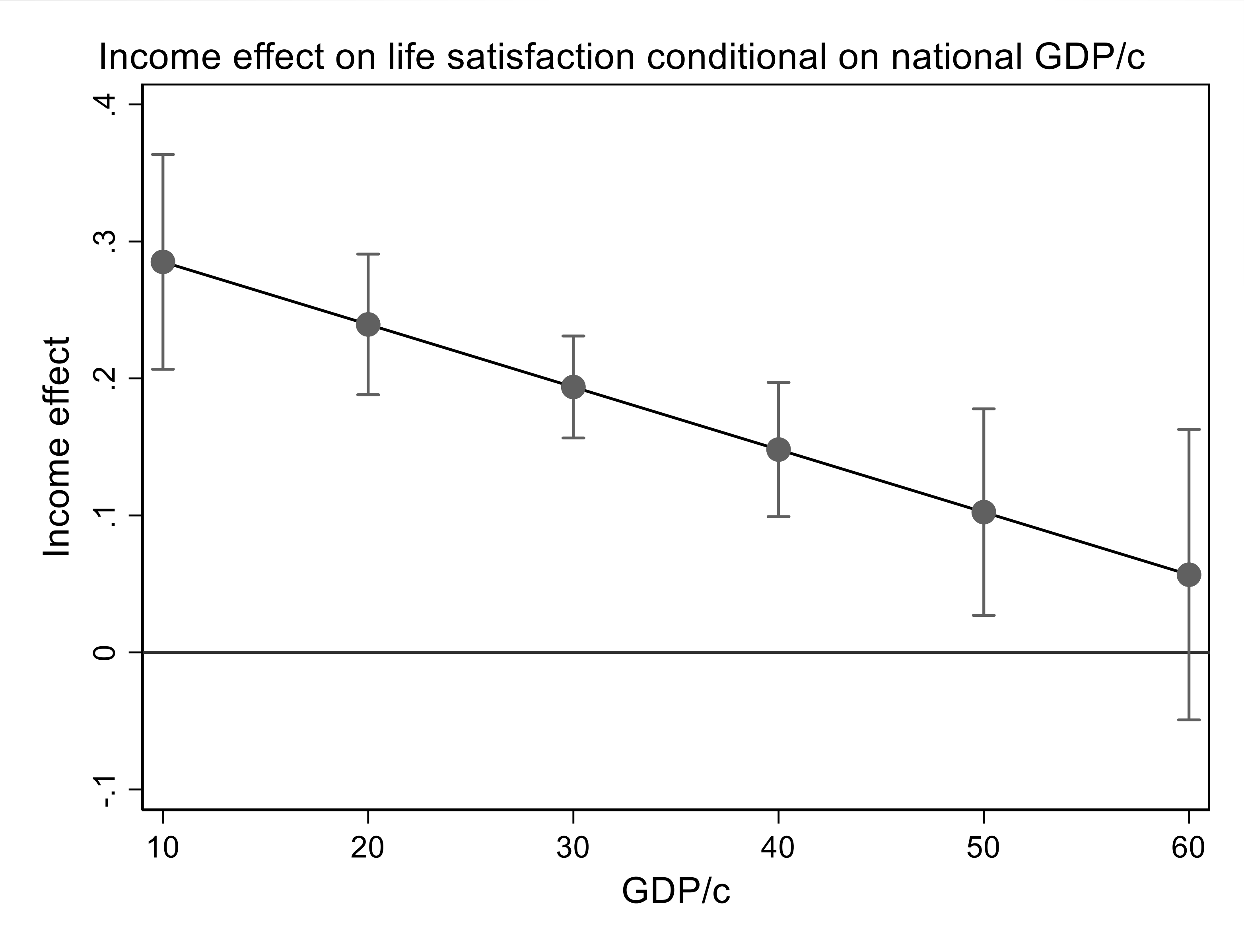
- Outcome : life satisfaction (`stflife`)
- Explanatory variables: income (`hinctnt`, individual level), GDP/c (`rgdpc`, country level)
- Question: does national wealth explain the variance of the income effect?

→In R: `lmer(stflife ~ hinctnt + rgdpc + hinctnt*rgdpc + (1 + hinctnt | cntry), = ESS02)`

EXAMPLE

Life satisfaction	Model 0	Model 1	Model 2	Model 3	**Model 4**
Income		0.18 ***	0.18 ***	0.19 ***	**0.33 *** **
GDP/c			0.02 ***	0.02 ***	**0.05 ** **
Income × GDP/c					**-0.005 ** **
Intercept	7.02***	5.99 ***	5.10 ***	5.48 ***	**4.42 *** **
Random effects					
Intercept	0.617	0.319	0.210	0.771	**0.587**
Income				0.008	**0.005**
Covar(Intercept-Income)				-0.07	**-0.046**
Residual	4.559	4.262	4.262	4.194	**4.23**

* p<0.05, ** p<0.01, *** p<0.001


Income effect on life satisfaction conditional on national GDP/c
Income effect
.4
.3
.2
.1
0
-.1
10
20
30
40
50
60
GDP/c

EXAMPLE

Life satisfaction	Model 0	Model 1	Model 2	**Model 3**	**Model 4**
				No cross-level interaction	Includes cross-level interaction
Random effect of Intercept				0.008	0.005

→ Reduction of slope variance by adding cross-level interaction: 1 - (0.005 / 0.008) = 0.375

→ 37.5 percent of variance of income effect explained by the interaction with GDP/c

CROSS-LEVEL INTERACTIONS AND RANDOM SLOPES

- Strictly speaking, models with cross-level interactions should include a random slope for the individual-level moderator *even this random slope is not of theoretical interest*
- This is because (potential) correlation of first level errors within countries (e_{ij}) affects standard errors (heteroscedasticity)
- When omitting the random slope (u_{1j}), individuals from the same country are treated as contributing independent information to the cross-level interaction (see Heisig & Schaeffer 2019: 263)
- This is similar to the violation of the independence assumption for models without random intercept (u_{0j}) (see session on HLMs)
- Typically, standard errors will be anti-conservative and effects will look "too significant"

TRANSFORMING VARIABLES: CENTERING

MEAN CENTERING

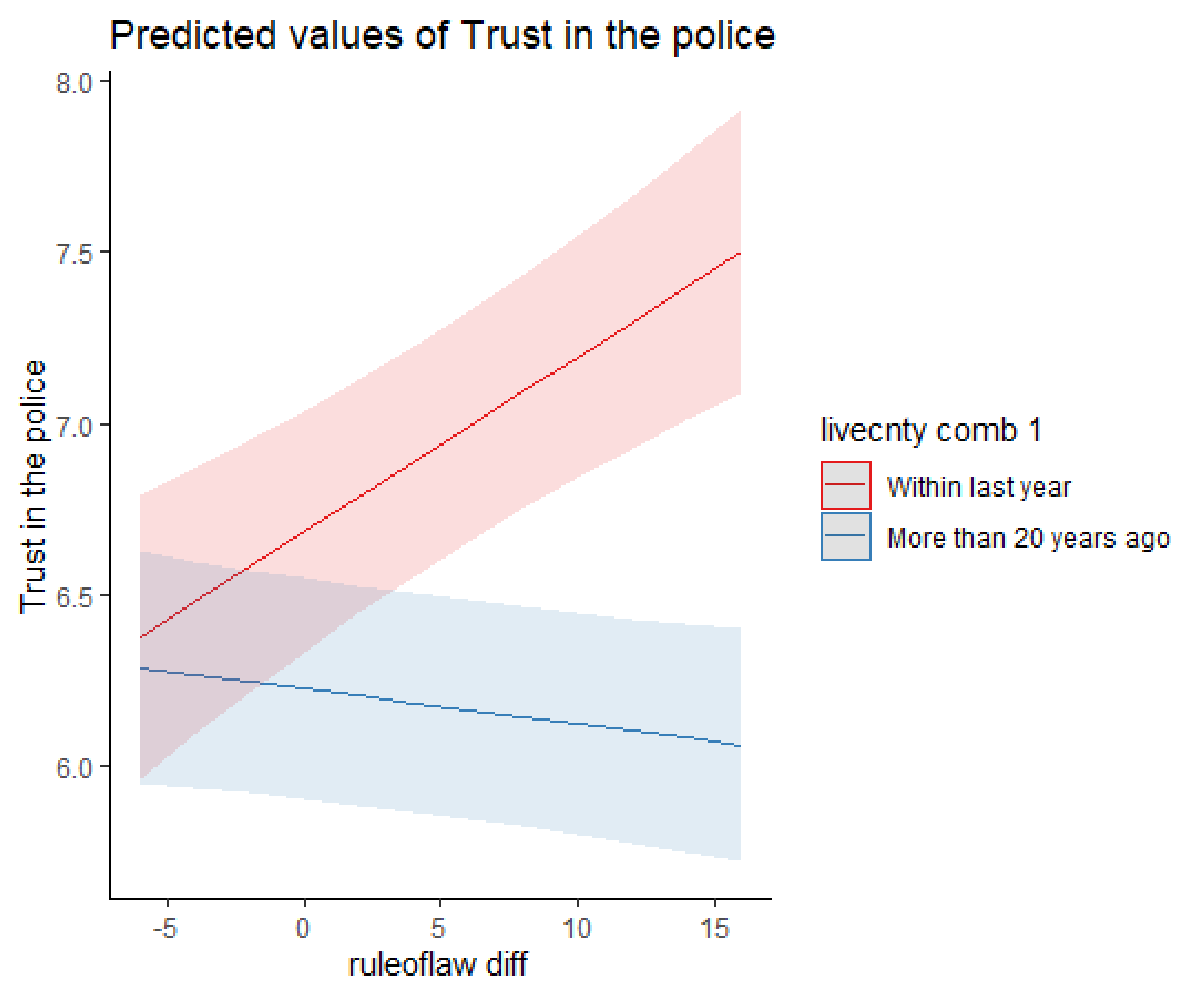
- To mean center a variable means to redefine the zero point of a variable
- Useful for continuous x
- Mathematically, it is the difference from a variable's value to its mean (variable – mean(variable))
- Centered variable will have a mean of 0
- Centering is a linear transformation (subtraction) and, thus, does not change substantive results
- … but potentially the p-values → always plot interactions
- With multi-level data, there are two possibilities: *grand mean centering* and *group mean centering*

GRAND MEAN CENTERING

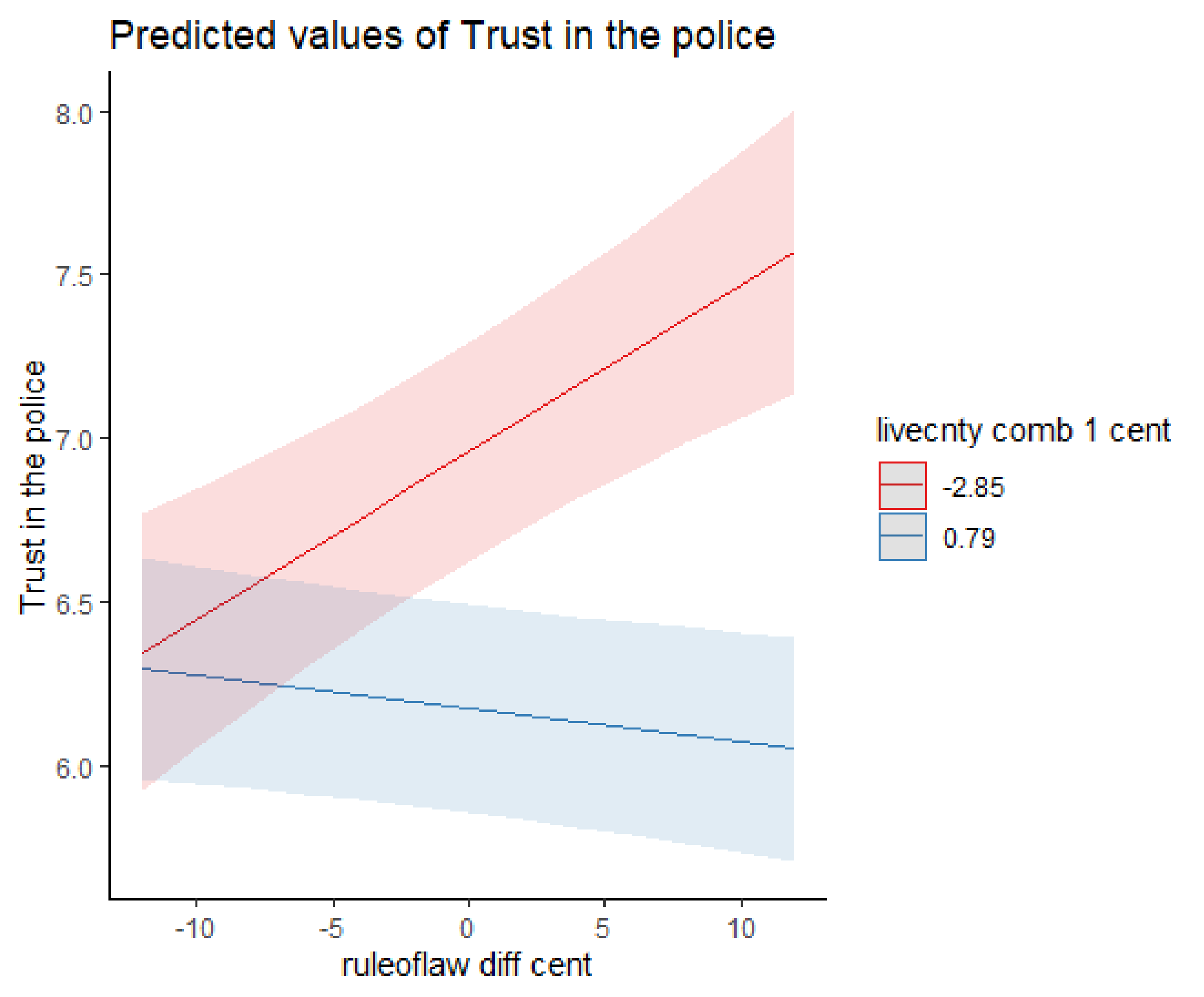
- Centering on the overall mean across all countries and individuals (grand mean)
- Interpretation for cross-level interaction: "*Effect of x in a country with average z*" or "*Effect of z for a person with an average x*"
- Model with mean centered variables substantively equivalent to model with original variables

- Why do the effects of `ruleoflaw_diff` and `livecnty_comb1` differ across both models?
- Why is the interaction the same?

	Trust in the police				Trust in the police			
Predictors	*Estimates*	*std. Error*	*CI*	*p*	*Estimates*	*std. Error*	*CI*	*p*
(Intercept)	6.68	0.18	6.33 – 7.04	**<0.001**	6.34	0.16	6.03 – 6.66	**<0.001**
ruleoflaw diff	0.05	0.01	0.03 – 0.07	**<0.001**				
livecnty comb 1	-0.11	0.02	-0.16 – -0.07	**<0.001**				
ruleoflaw_diff:livecnty_comb1	-0.02	0.00	-0.02 – -0.01	**<0.001**				
ruleoflaw diff cent					0.00	0.00	-0.00 – 0.01	0.455
livecnty comb 1 cent					-0.20	0.02	-0.23 – -0.16	**<0.001**
ruleoflaw_diff_cent:livecnty_comb1_cent					-0.02	0.00	-0.02 – -0.01	**<0.001**
Random Effects								
σ^2	5.57				5.57			
τ_{00}	0.15 essround:cntry				0.15 essround:cntry			
	0.53 cntry				0.53 cntry			
ICC	0.11				0.11			
N	7 essround				7 essround			
	22 cntry				22 cntry			
Observations	19655				19655			
Marginal R^2 / Conditional R^2	0.008 / 0.116				0.008 / 0.116			



Predicted values of Trust in the police
Trust in the police
ruleoflaw diff
livecnty comb 1
Within last year
More than 20 years ago



Predicted values of Trust in the police
Trust in the police
ruleoflaw diff cent
livecnty comb 1 cent
-2.85
0.79

BENEFITS OF GRAND MEAN CENTERING IN HLM

- Better interpretability of intercept‘s fixed effect (the intercept): all x and z at their means (expected value of average subject)
- Better interpretability random effect(s) → Expected variances when all all x and z at their means
- Better interpretability of main effects for interactions (effect of x_1 conditional on x_2 being on its mean)
- (Potentially) faster calculations, easier convergence of models

GROUP MEAN CENTERING

- Center at the country mean
- Only possible for individual-level variables (for country variables the country mean is simply its value)
- In R:
  - ```
    data %<>%
      group_by(cntry) %>%
      mutate(var_g_cent = var - mean(var))
    ```
- Not a simple reparameterization of a model but a completely *different model* (cf. Hox 2010: 68 f.)
    ```

WHEN TO USE GROUP MEAN CENTERING

- Usually, individual-level effects include a mixture of effects *within* and *between* countries
- Group mean centering *removes all between-group variance* (highly used with panel data, not so much with cross-national data, for a comparison see Ziller 2018)
- Only variance *within* countries (between individuals) is analyzed, all differences *between* countries are automatically controlled for
- This is equivalent to using country fixed effects (adding dummy variables for all countries but one)
- Accordingly, group mean centering can not be used for country level variables
- However, estimation of cross-level interactions possible (see example)

	Trust in the police				Trust in the police				Trust in the police			
Predictors	*Estimates*	*std. Error*	*CI*	*p*	*Estimates*	*std. Error*	*CI*	*p*	*Estimates*	*std. Error*	*CI*	*p*
(Intercept)	6.68	0.18	6.33 – 7.04	**<0.001**	6.34	0.16	6.03 – 6.66	**<0.001**	6.34	0.17	6.00 – 6.67	**<0.001**
ruleoflaw diff	0.05	0.01	0.03 – 0.07	**<0.001**								
livecnty comb 1	-0.11	0.02	-0.16 – -0.07	**<0.001**								
ruleoflaw_diff:livecnty_comb1	-0.02	0.00	-0.02 – -0.01	**<0.001**								
ruleoflaw diff cent					0.00	0.00	-0.00 – 0.01	0.455				
livecnty comb 1 cent					-0.20	0.02	-0.23 – -0.16	**<0.001**				
ruleoflaw_diff_cent:livecnty_comb1_cent					-0.02	0.00	-0.02 – -0.01	**<0.001**				
ruleoflaw diff g cent									0.00	0.00	-0.00 – 0.01	0.331
livecnty comb 1 g cent									-0.19	0.02	-0.22 – -0.16	**<0.001**
ruleoflaw_diff_g_cent:livecnty_comb1_g_cent									-0.01	0.00	-0.02 – -0.00	**0.006**
Random Effects												
σ^2	5.57				5.57				5.57			
τ_{00}	0.15 $_{essround:cntry}$				0.15 $_{essround:cntry}$				0.15 $_{essround:cntry}$			
	0.53 $_{cntry}$				0.53 $_{cntry}$				0.61 $_{cntry}$			
ICC	0.11				0.11				0.12			
N	7 $_{essround}$				7 $_{essround}$				7 $_{essround}$			
	22 $_{cntry}$				22 $_{cntry}$				22 $_{cntry}$			
Observations	19655				19655				19655			
Marginal R^2 / Conditional R^2	0.008 / 0.116				0.008 / 0.116				0.006 / 0.125			

INTERPRETATION

- Focus on `ruleoflaw_diff` (individual level)
  - Uncentered model: Effect of `ruleoflaw_diff` for those who migrated *most recently*
  - Grand mean centered model: Effect of `ruleoflaw_diff` for those whith an *average* time since migration
  - Group mean centered model: Effect of `ruleoflaw_diff` net of country-level differences of police trust (purely based on *within* country variation)
- In this example, estimates of group mean centered model is similar to the others because most variance of trust in police is on the individual-level (between country differences less important)
- Effects of group mean centered variables will differ…
  - The more effects within countries and between countries differ
  - The larger the share of between country variance

LINEAR TRANSFORMATIONS AND RANDOM SLOPES

- Linear models are invariant for linear transformations (substantive results do not change)
- … but not estimates of random intercept when random slopes are modelled
- This is because the spread of intercepts changes when x is rescaled
- This holds for random slope models independent of any interactions

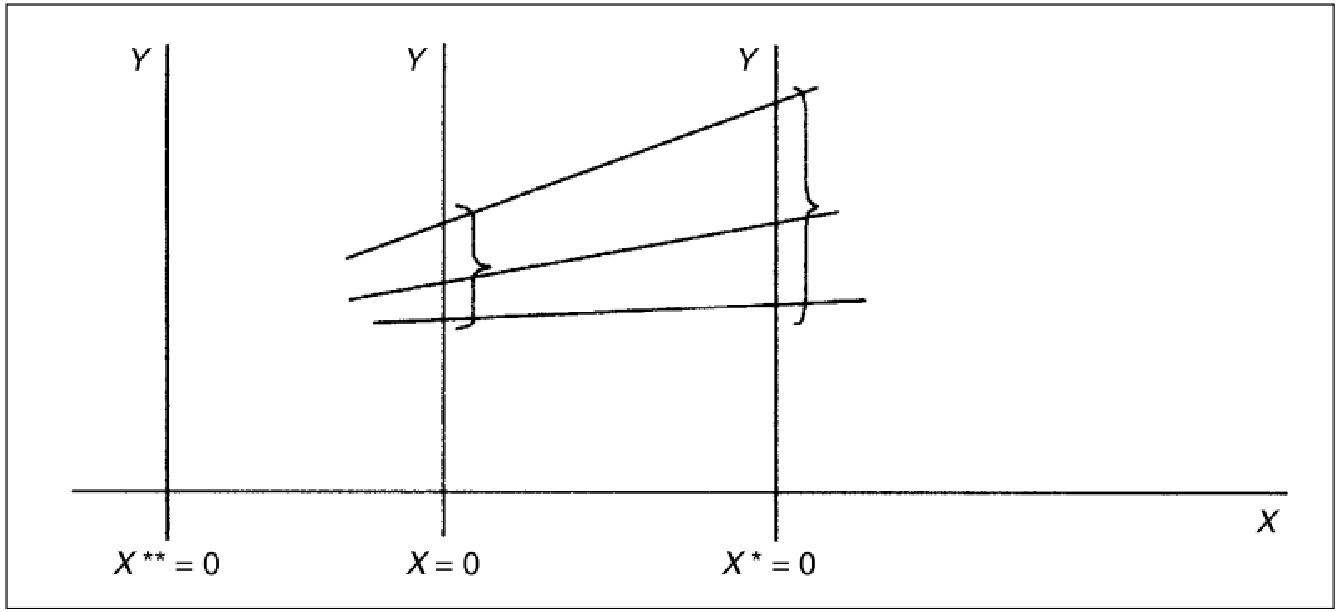


Figure 4.2 Varying regression lines, with shifts on X.

Hox 2010: 61

IN BRIEF

- Grand mean centering changes intercept but not slope
- Group mean centering changes both intercept and slope

WRAPPING IT UP

WHICH HLM IS THE RIGHT ONE?

- This depends *solely on the particular research question*
- Is it plausible that individual level effects are constant over the whole population? → No random slope needed
- Is it interesting to see whether effects differ between social groups? → Model interaction
- Etc.

→The use of a certain model should be justified theoretically (and empirically)

LITERATURE

- Heisig & Schaeffer (2019). Why You Should *Always* Include a Random Slope for the Lower-Level Variable Involved in a Cross-Level Interaction. European Sociological Review, 35 (2): 258-279