Multilevel

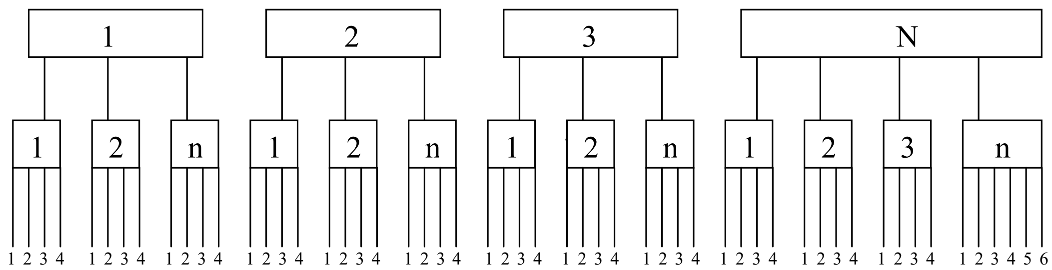
Hierarchical Analyses: The Equations

TILBURG UNIVERSITY

tesc

Hierarchical analysis:

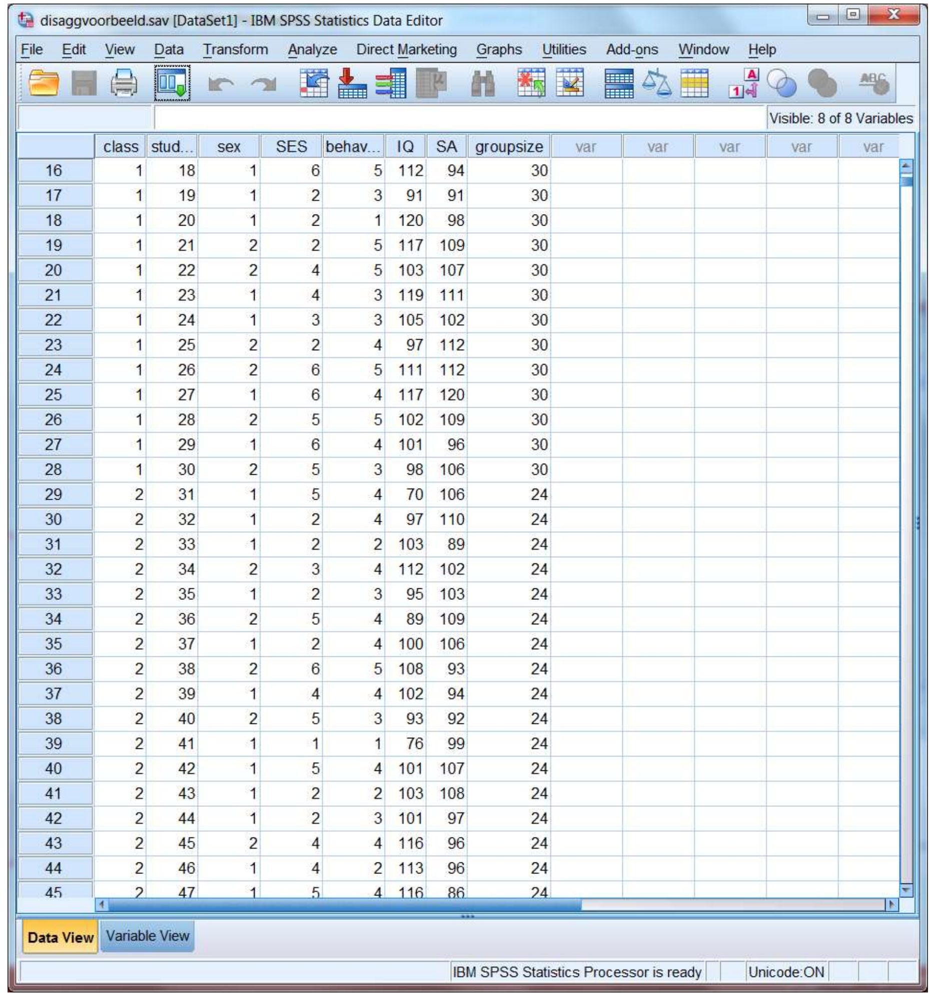
Figure 1: Example hierarchical data structure



Examples:	**Education**	**Organizational**	**Longitudinal**
Level 3	schools	organizations	classes
Level 2	classes	departments	pupils
Level 1	pupils	individuals	observations

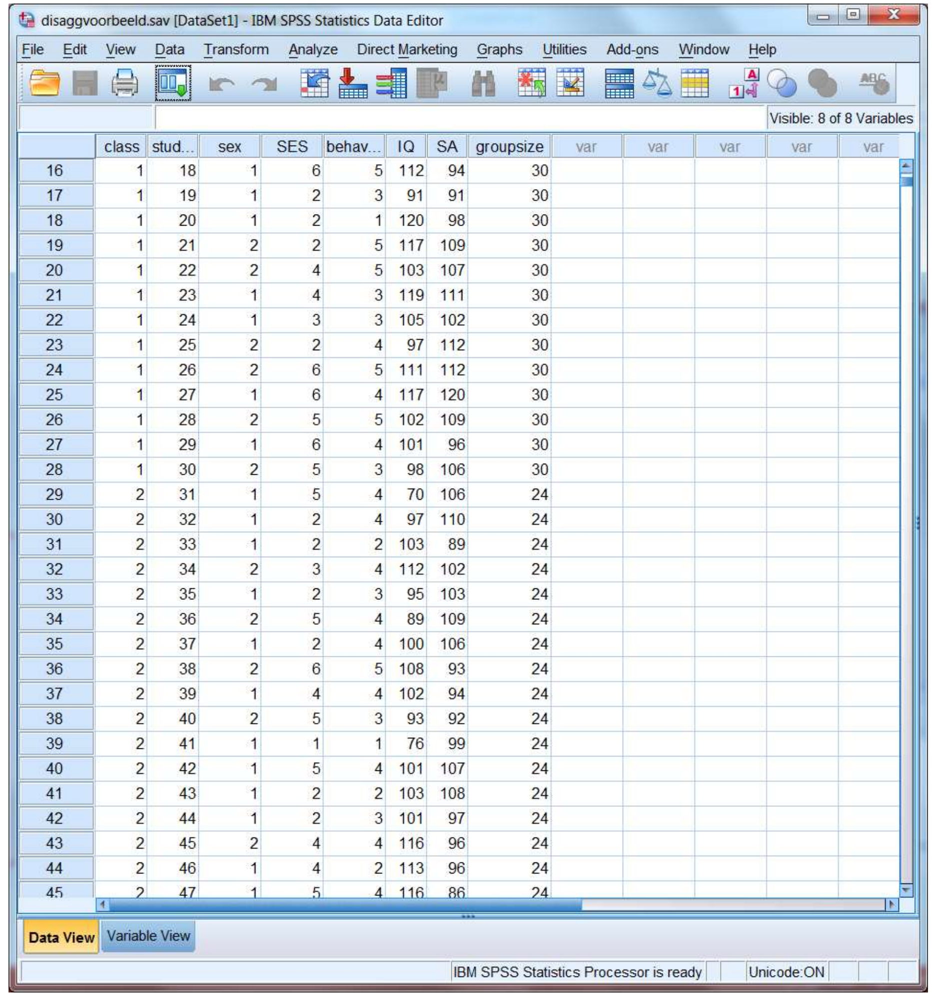
Why Multilevel?

- Variables at different levels/Dependency.
  - Level 1: Need to correct for dependency in data.



	class	stud...	sex	SES	behav...	IQ	SA	groupsize	var	var	var	var	var
16	1	18	1	6	5	112	94	30					
17	1	19	1	2	3	91	91	30					
18	1	20	1	2	1	120	98	30					
19	1	21	2	2	5	117	109	30					
20	1	22	2	4	5	103	107	30					
21	1	23	1	4	3	119	111	30					
22	1	24	1	3	3	105	102	30					
23	1	25	2	2	4	97	112	30					
24	1	26	2	6	5	111	112	30					
25	1	27	1	6	4	117	120	30					
26	1	28	2	5	5	102	109	30					
27	1	29	1	6	4	101	96	30					
28	1	30	2	5	3	98	106	30					
29	2	31	1	5	4	70	106	24					
30	2	32	1	2	4	97	110	24					
31	2	33	1	2	2	103	89	24					
32	2	34	2	3	4	112	102	24					
33	2	35	1	2	3	95	103	24					
34	2	36	2	5	4	89	109	24					
35	2	37	1	2	4	100	106	24					
36	2	38	2	6	5	108	93	24					
37	2	39	1	4	4	102	94	24					
38	2	40	2	5	3	93	92	24					
39	2	41	1	1	1	76	99	24					
40	2	42	1	5	4	101	107	24					
41	2	43	1	2	2	103	108	24					
42	2	44	1	2	3	101	97	24					
43	2	45	2	4	4	116	96	24					
44	2	46	1	4	2	113	96	24					
45	2	47	1	5	4	116	86	24					

TILBURG UNIVERSITY

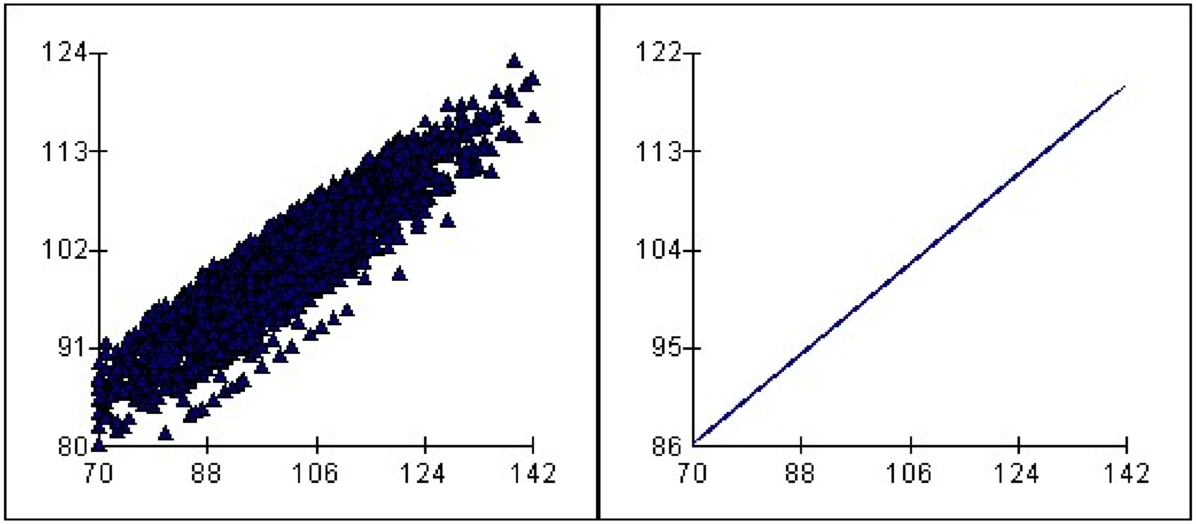


disaggvoorbeeld.sav [DataSet1] - IBM SPSS Statistics Data Editor

	class	stud...	sex	SES	behav...	IQ	SA	groupsize
16	1	18	1	6	5	112	94	30
17	1	19	1	2	3	91	91	30
18	1	20	1	2	1	120	98	30
19	1	21	2	2	5	117	109	30
20	1	22	2	4	5	103	107	30
21	1	23	1	4	3	119	111	30
22	1	24	1	3	3	105	102	30
23	1	25	2	2	4	97	112	30
24	1	26	2	6	5	111	112	30
25	1	27	1	6	4	117	120	30
26	1	28	2	5	5	102	109	30
27	1	29	1	6	4	101	96	30
28	1	30	2	5	3	98	106	30
29	2	31	1	5	4	70	106	24
30	2	32	1	2	4	97	110	24
31	2	33	1	2	2	103	89	24
32	2	34	2	3	4	112	102	24
33	2	35	1	2	3	95	103	24
34	2	36	2	5	4	89	109	24
35	2	37	1	2	4	100	106	24
36	2	38	2	6	5	108	93	24
37	2	39	1	4	4	102	94	24
38	2	40	2	5	3	93	92	24
39	2	41	1	1	1	76	99	24
40	2	42	1	5	4	101	107	24
41	2	43	1	2	2	103	108	24
42	2	44	1	2	3	101	97	24
43	2	45	2	4	4	116	96	24
44	2	46	1	4	2	113	96	24
45	2	47	1	5	4	116	86	24

TILBURG UNIVERSITY

Hierarchical analysis: Research questions



- Use multilevel because observations are likely to de dependent, however:
  - These type of questions could still be answered with normal regression.
  - Merely involves adjusting obtained s.e.

Hierarchical analysis: Research questions

- Could correct for this with the formula by Kish (1965):

- $v_{eff} = v(1 + (n_{clus} - 1)\rho)$, where
- $\rho = \frac{(MS_b - MS_w)}{MS_b + (n_{clus} - 1)MS_w)}$

- This obviously does not provide the nice missing data handling of multilevel

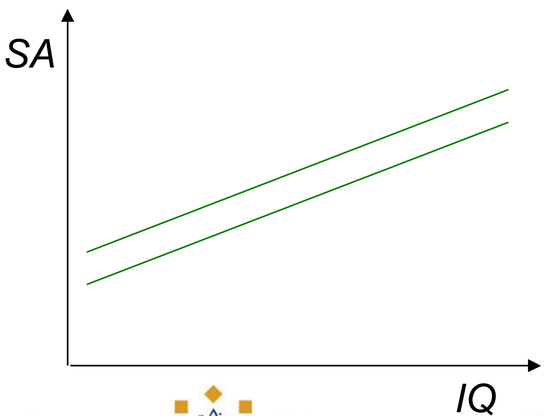
Why Multilevel?

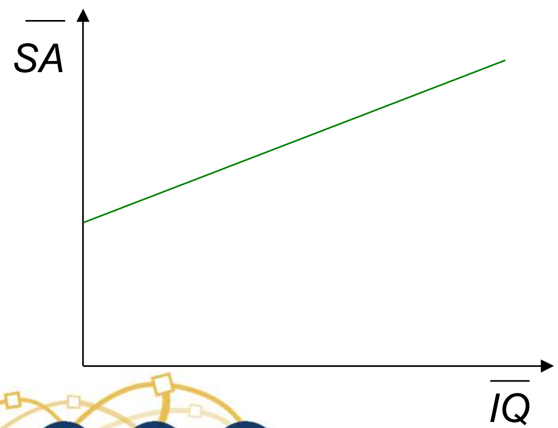
- Variables at different levels/Dependency.
  - Level 1: Need to correct for dependency in data.
  - Level 2: Disaggregation

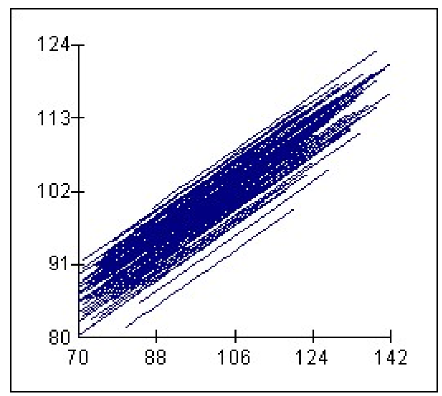
Multilevel research questions

- Questions with respect to the influence of variables at a higher level on the dependent variable at the lowest level:
  - *mean intelligence* of a class (*MIQ*) as predictor of mean *school achievement* (*SA*); (controlling for individual *IQ*)

$$SA_{ij} = b_0 + b_1 IQ_{ij} + b_2 MIQ_j + u_j + e_{ij}$$

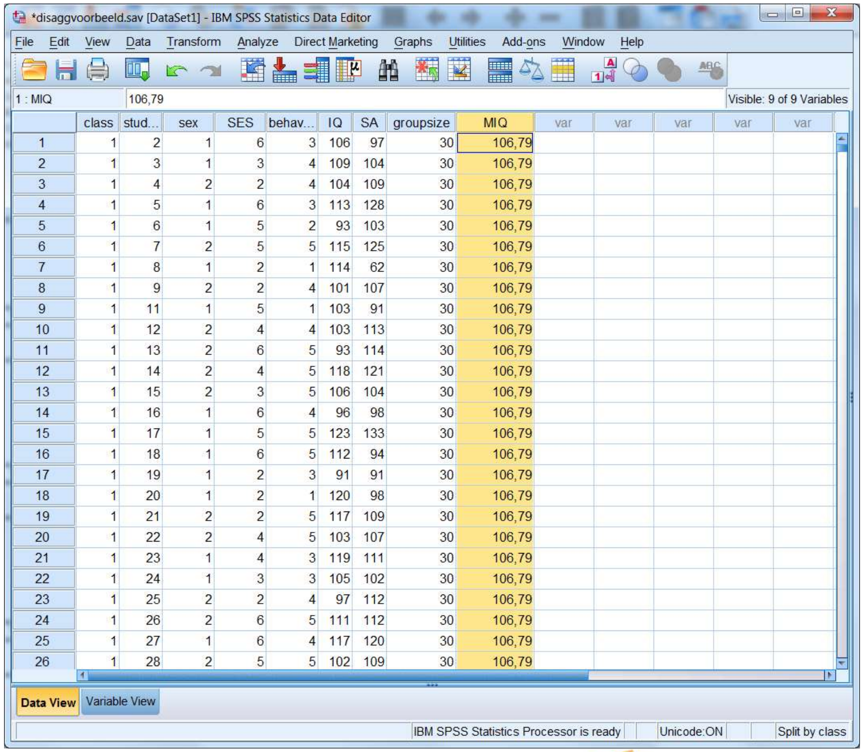




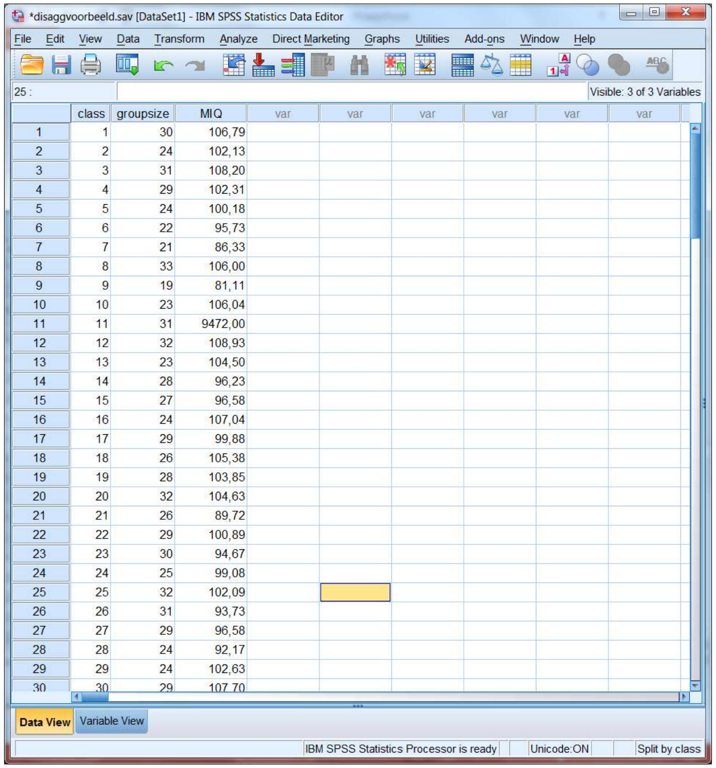


Why multilevel analysis?

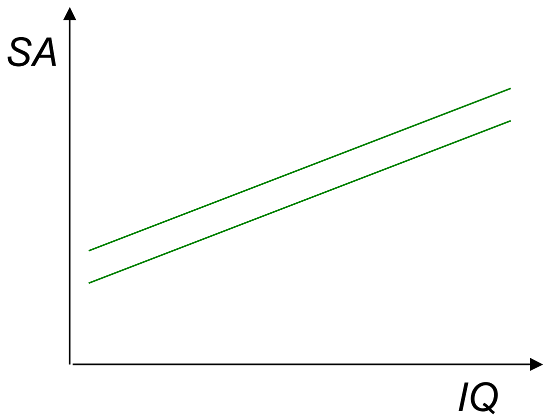
Example: Disaggregation versus multilevel analysis

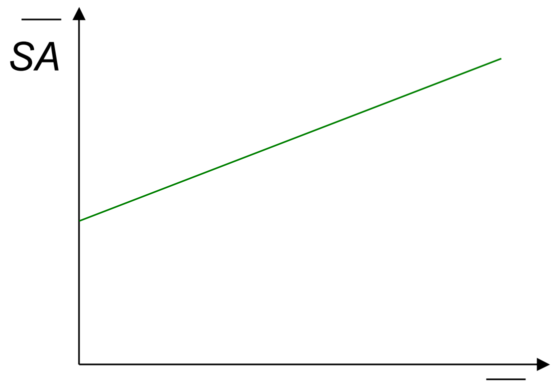


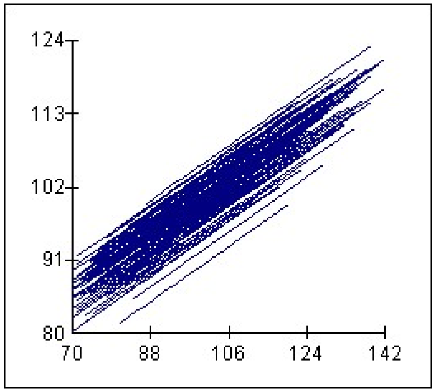
	class	stud...	sex	SES	behav...	IQ	SA	groupsize	MIQ
1	1	2	1	6	3	106	97	30	106,79
2	1	3	1	3	4	109	104	30	106,79
3	1	4	2	2	4	104	109	30	106,79
4	1	5	1	6	3	113	128	30	106,79
5	1	6	1	5	2	93	103	30	106,79
6	1	7	2	5	5	115	125	30	106,79
7	1	8	1	2	1	114	62	30	106,79
8	1	9	2	2	4	101	107	30	106,79
9	1	11	1	5	1	103	91	30	106,79
10	1	12	2	4	4	103	113	30	106,79
11	1	13	2	6	5	93	114	30	106,79
12	1	14	2	4	5	118	121	30	106,79
13	1	15	2	3	5	106	104	30	106,79
14	1	16	1	6	4	96	98	30	106,79
15	1	17	1	5	5	123	133	30	106,79
16	1	18	1	6	5	112	94	30	106,79
17	1	19	1	2	3	91	91	30	106,79
18	1	20	1	2	1	120	98	30	106,79
19	1	21	2	2	5	117	109	30	106,79
20	1	22	2	4	5	103	107	30	106,79
21	1	23	1	4	3	119	111	30	106,79
22	1	24	1	3	3	105	102	30	106,79
23	1	25	2	2	4	97	112	30	106,79
24	1	26	2	6	5	111	112	30	106,79
25	1	27	1	6	4	117	120	30	106,79
26	1	28	2	5	5	102	109	30	106,79



	class	groupsize	MIQ
1	1	30	106,79
2	2	24	102,13
3	3	31	108,20
4	4	29	102,31
5	5	24	100,18
6	6	22	95,73
7	7	21	86,33
8	8	33	106,00
9	9	19	81,11
10	10	23	106,04
11	11	31	9472,00
12	12	32	108,93
13	13	23	104,50
14	14	28	96,23
15	15	27	96,58
16	16	24	107,04
17	17	29	99,88
18	18	26	105,38
19	19	28	103,85
20	20	32	104,63
21	21	26	89,72
22	22	29	100,89
23	23	30	94,67
24	24	25	99,08
25	25	32	102,09
26	26	31	93,73
27	27	29	96,58
28	28	24	92,17
29	29	24	102,63
30	30	29	107,70







- This type of question could also be answered with normal regression by running a regression on mean scores:
  - Power obviously an issue.
  - Should means from groups of different size be treated the same?

Why Multilevel?

- Variables at different levels.
  - Level 1: Need to correct for dependency in data.
  - Level 2: Disaggregation
  - Level 1&2: Cross-level interaction
    - We want to MODEL differences between level 2 units in the effects of level 1 variables.

Hierarchical Analyses: The Steps

Analysis Strategy

Analysis strategy

- Multilevel model with:
  - *p* explanatory variables at the lowest level
  - *q* explanatory variables at the highest level

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + \gamma_{0q} Z_{qj} + \gamma_{pq} Z_{qj} X_{pij} + u_{pj} X_{pij} + u_{0j} + e_{ij}$$

Steps of a multilevel analysis

1. Check whether multilevel is necessary
2. Add all level 1 main effects and interactions
3. Add all level 2 main effects and interactions
4. Check level 1 effects for random slopes
5. If random slopes are present: add cross-level interactions

Step 1: two level intercept-only model

- Analyze a two-level model with no explanatory variables: the *intercept-only* model:

$$Y_{ij} = \beta_{0j} + e_{ij}$$

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

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij}$$

- γ_{00} : the regression intercept
- u_{0j} : residual error at the second level
- e_{ij} : residual error at the first level

- Decomposition of the total variance in two terms
- Should we perform a multilevel analysis?

Tilburg University

Step 1: two level intercept-only model

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij}$$

- H_0: individuals in the same group aren't more alike than individuals in different groups

  H_0: $\sigma^2_{u_0} = 0$

- H_1: individuals in the same group are more alike than individuals in different groups

  H_1: $\sigma^2_{u_0} > 0$

Intraclass correlation

$$\rho = \frac{\sigma^2_{u_0}}{\sigma^2_{u_0} + \sigma^2_e}$$

- variance of the second level errors u_{0j} $\sigma^2_{u_0}$
- variance of the first level errors e_{ij} σ^2_e

- The ‘percentage’ variance at the second level

- The expected correlation between the observations on the dependent variable of two randomly chosen units that are in the same group.

Step 2: first level variables

Add all first level variables (+ their interactions) with fixed effects:

$$Y_{ij} = \beta_{0j} + \beta_{pj} X_{pij} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{pj} = \gamma_{p0}$$

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + u_{0j} + e_{ij}$$

Step 2: first level variables

Add all first level variables (+ their interactions) with fixed effects:

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + u_{0j} + e_{ij}$$

- X_{pij} : *p* explanatory variables at the first level (e.g.,gender, extraversion, and gender*extraversion)
- H_0: There is no relation between the dependent variable *Y* and the explanatory variable *X*
- H_1: There is a relation between the dependent variable *Y* and the explanatory variable *X* (+ direction)

- Calculate the explained variance at level 1 and at level 2

Step 2: first level variables

- Calculate the explained variance at level 1 and level 2.

- Why?

Tilburg University

Step 2: first level variables

- Calculate the explained variance at level 1 and level 2.

- Variables at the lowest level can explain variance at the first level
  - Girls are more popular than boys.

- Variables at the lowest levels can explain variance at the higher levels.
  - Girls are more popular than boys.
  - In a class with a lot of girls the mean popularity is higher.

Step 3: Second level variables

Add all second level variables (+ their interactions) with fixed effects:

$$Y_{ij} = \beta_{0j} + \beta_{pj} X_{pij} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{0q} Z_{qj} + u_{0j}$$

$$\beta_{pj} = \gamma_{p0}$$

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + \gamma_{0q} Z_{qj} + u_{0j} + e_{ij}$$

Step 3: Second level variables

Add all second level variables (+ their interactions) with fixed effects:

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + \gamma_{0q} Z_{qj} + u_{0j} + e_{ij}$$

- Z_{qj} : q explanatory variables at the second level (e.g., teacher experience).
- H_0: there is no relation between the explanatory variable Z and the mean score of the dependent variable *Y*
- H_1: there is a relation between the explanatory variable *Z* and the mean score of the dependent variable *Y* (+ direction?)
- Calculate the explained variance at level 2

Step 4: random slopes

Assess whether any of the slopes of the explanatory variables of the first level has a significant variance between the second level units:

$$Y_{ij} = \beta_{0j} + \beta_{pj} X_{pij} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{0q} Z_{qj} + u_{0j}$$

$$\beta_{pj} = \gamma_{p0} + u_{pj}$$

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + \gamma_{0q} Z_{qj} + u_{pj} X_{pij} + u_{0j} + e_{ij}$$

Tilburg University

tesc

Step 4: random slopes

Assess whether any of the slopes of the explanatory variables of the first level has a significant variance between the second level units:

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + \gamma_{0q} Z_{qj} + u_{pj} X_{pij} + u_{0j} + e_{ij}$$

- u_{pj} : second level residuals of the slopes of the first level explanatory variables X_{pij}

- H_0: the relation between the explanatory variable X and the dependent variable Y is the same within all level two units (H_0: $\sigma^2_{u_{1j}}$ = 0)

- H_1: the relation between the explanatory variable X and the dependent variable Y is not the same within all level two units (H_1: $\sigma^2_{u_{1j}}$ > 0)

TILBURG UNIVERSITY

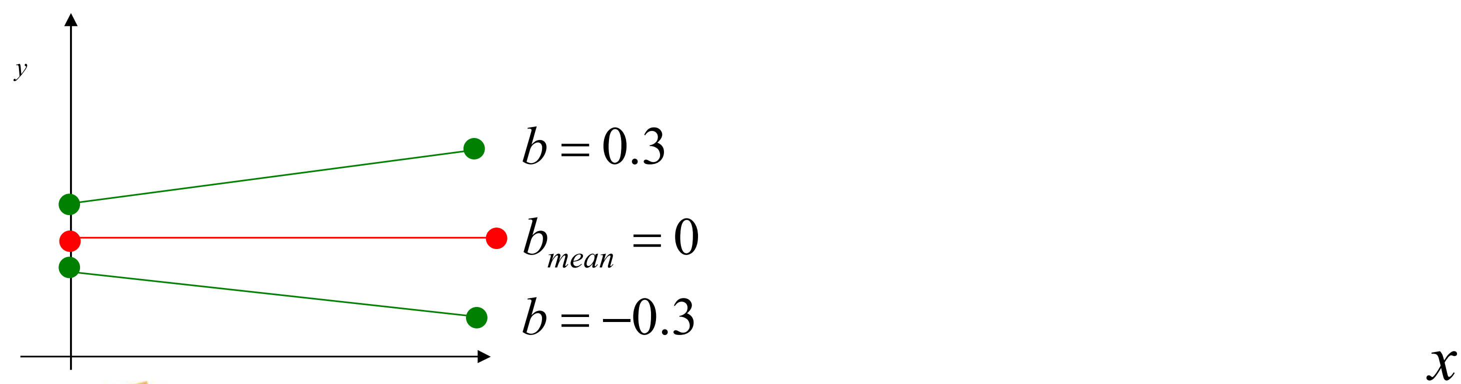
tesc

Remarks on step 4: random slopes

- Testing for random slope variation: variable by variable

- Variables that were omitted in step 2 may be analyzed again:
  - it is possible that an explanatory variable has no significant mean regression slope, but that there is slope variance

- Add all significant slopes simultaneously in a final model.



Step 5: cross-level interactions

Add cross-level interactions between those level one explanatory variables that had significant slope variation and explanatory variables of the second level:

$$Y_{ij} = \beta_{0j} + \beta_{pj} X_{pij} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{0q} Z_{qj} + u_{0j}$$

$$\beta_{pj} = \gamma_{p0} + \gamma_{pq} Z_{qj} + u_{pj}$$

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + \gamma_{0q} Z_{qj} + \gamma_{pq} Z_{qj} X_{pij} + u_{pj} X_{pij} + u_{0j} + e_{ij}$$

Tilburg University tesc

Step 5: cross-level interactions

Add cross-level interactions between those level one explanatory variables that had significant slope variation and explanatory variables of the second level:

$$Y_{ij} = \gamma_{00} + \gamma_{p0} X_{pij} + \gamma_{0q} Z_{qj} + \gamma_{pq} Z_{qj} X_{pij} + u_{pj} X_{pij} + u_{0j} + e_{ij}$$

- $Z_{qj} X_{pij}$: cross-level interactions

- H_0: the explanatory variable Z can't explain the different relations between the explanatory variable X and the dependent variable Y in different level two units

- H_1: the explanatory variable Z explains (a part of) the different relations between the explanatory variable X and the dependent variable Y in different level two units

- Calculate the explained slope variance

Multilevel Regression

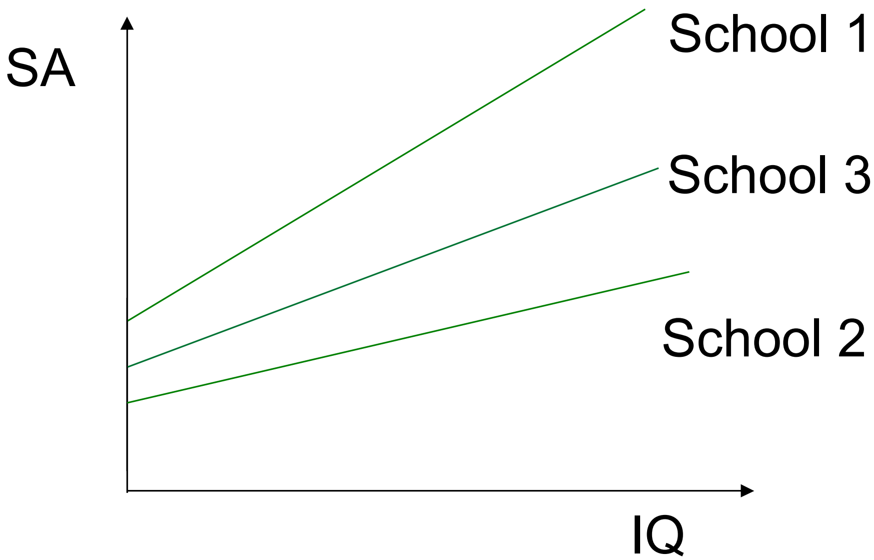
- Level 1: $SA_{ij} = b_{0j} + b_{1j}IQ_{ij} + e_{ij}$
- Level 2: $b_{0j} = \gamma_{00} + \gamma_{01}Schoolsize_j + u_{0j}$

  $b_{1j} = \gamma_{10} + u_{1j}$

Combined:

$$SA_{ij} = \gamma_{00} + \gamma_{10}IQ_{ij} + \gamma_{01}Schoolsize_j + eij + u_{0j} + u_{1j}\ IQ_{ij}$$

Multilevel Regression



Differences in intercept (partially) explained by school size:

$$b_{0j} = \gamma_{00} + \gamma_{01} Schoolsize_j + u_{0j}$$

So far, no predictor for differences in slope:

$$b_{1j} = \gamma_{10} + u_{1j}$$

Could school size explain differences in slope as well?

$$b_{1j} = \gamma_{10} + \gamma_{11} Schoolsize_j + u_{1j}$$

Multilevel Regression

Level 1: $SA_{ij} = b_{0j} + b_{1j}IQ_{ij} + e_{ij}$

Level 2: $b_{0j} = \gamma_{00} + \gamma_{01}Schoolsize_j + u_{0j}$

$$\boxed{b_{1j} = \gamma_{10} + \gamma_{11}Schoolsize_j + u_{1j}}$$

Smaller School → More help → Pupils less reliant on own IQ? → Positive γ_{11}

OR

Smaller School → Less help → Pupils more reliant on own IQ?

Multilevel Regression

- Level 1: $$SA_{ij} = b_{0j} + b_{1j}IQ_{ij} + e_{ij}$$
- Level 2: $$b_{0j} = \gamma_{00} + \gamma_{01}Schoolsize_j + u_{0j}$$ $$b_{1j} = \gamma_{10} + \gamma_{11}Schoolsize_j + u_{1j}$$

Combined:

$$SA_{ij} = \gamma_{00} + \gamma_{10}IQ_{ij} + \gamma_{01}Schoolsize_j \boxed{+ \gamma_{11}Schoolsize_j IQ_{ij}} + eij + u_{0j} + u_{1jIQ}ij$$

Hierarchical Analyses: The Steps

Update!...Don’t do the steps!!

Maximum Approach

- Going through the steps is a good way to see what is going on, but…

- We always know that things will differ between people/units?
  - So why assume fixed effects at first?
  - Why not go for a full model straight away?

Maximum Approach

- Going through the steps is a good way to see what is going on, but…
- We always know that things will differ between people/units?
  - So why assume fixed effects at first?
  - Why not go for a full model straight away? (Maximum Approach)
- What are possible advantages of staring with a full model?
- And what are disadvantages?

Maximum Approach

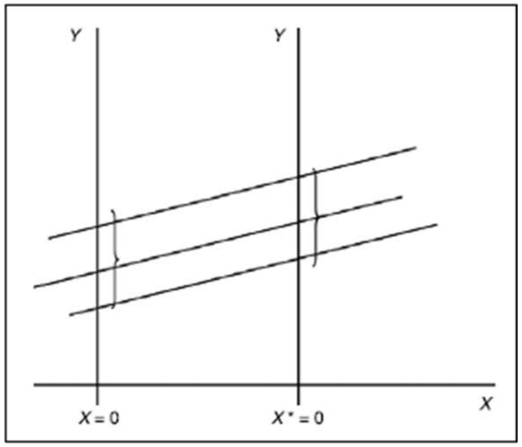


Figure 4.1 Parallel regression lines, with shift on *X*.

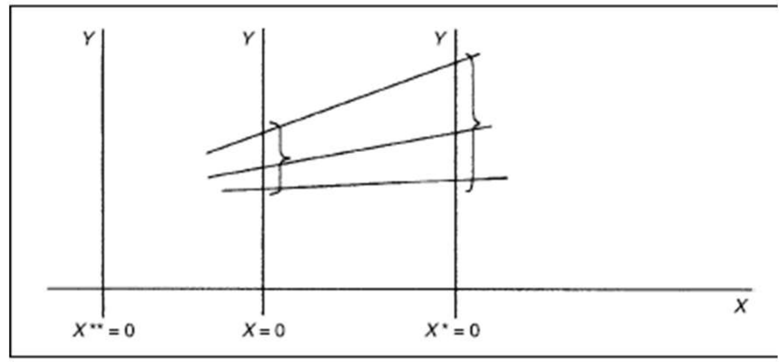
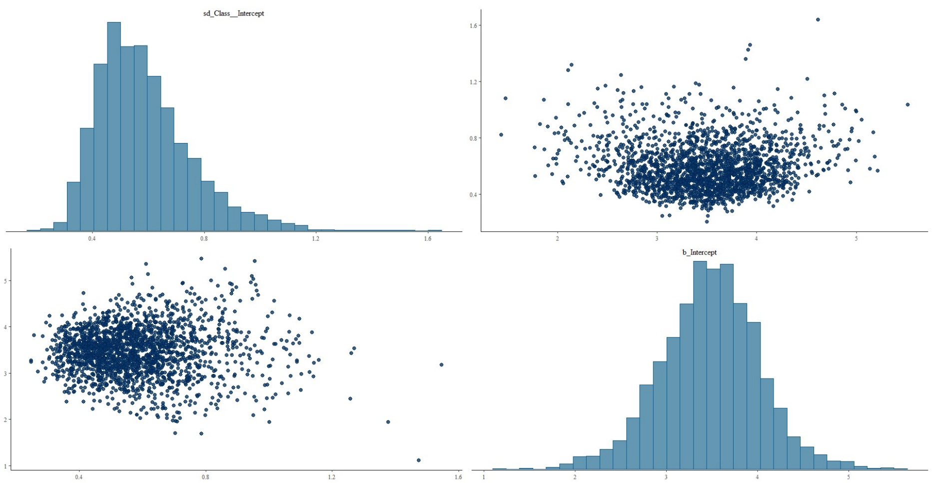
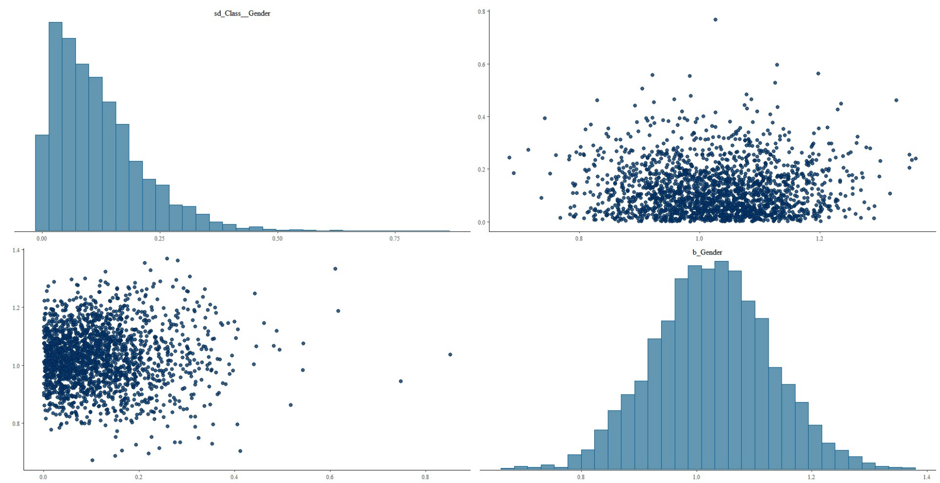
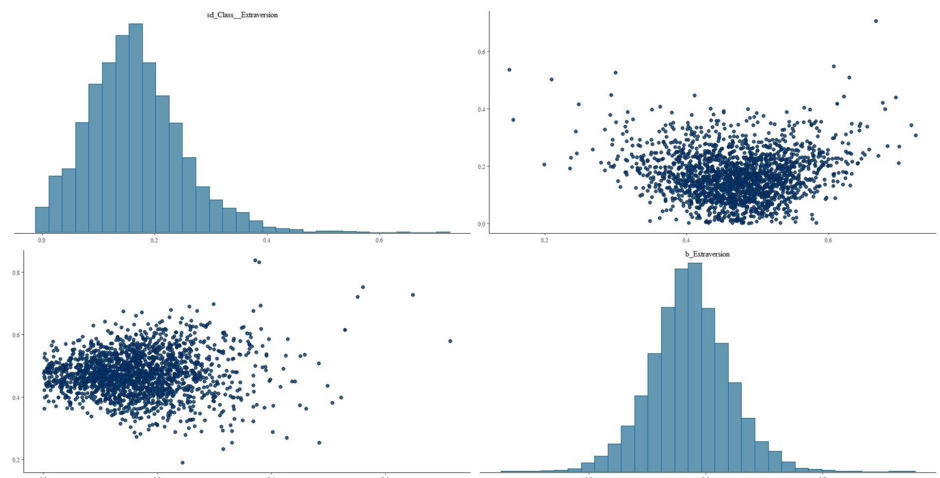


Figure 4.2 Varying regression lines, with shifts on *X*.

Maximum Approach

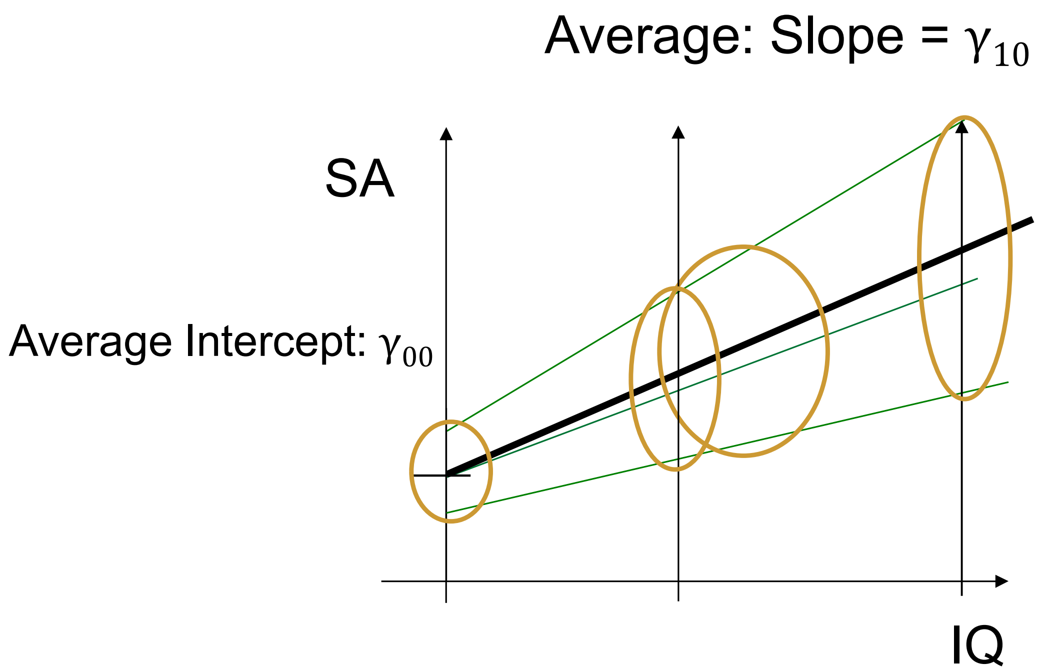






Methodological Considerations: Centering

Multilevel Regression



School 1: Slope = $\gamma_{10} + u_{11}$

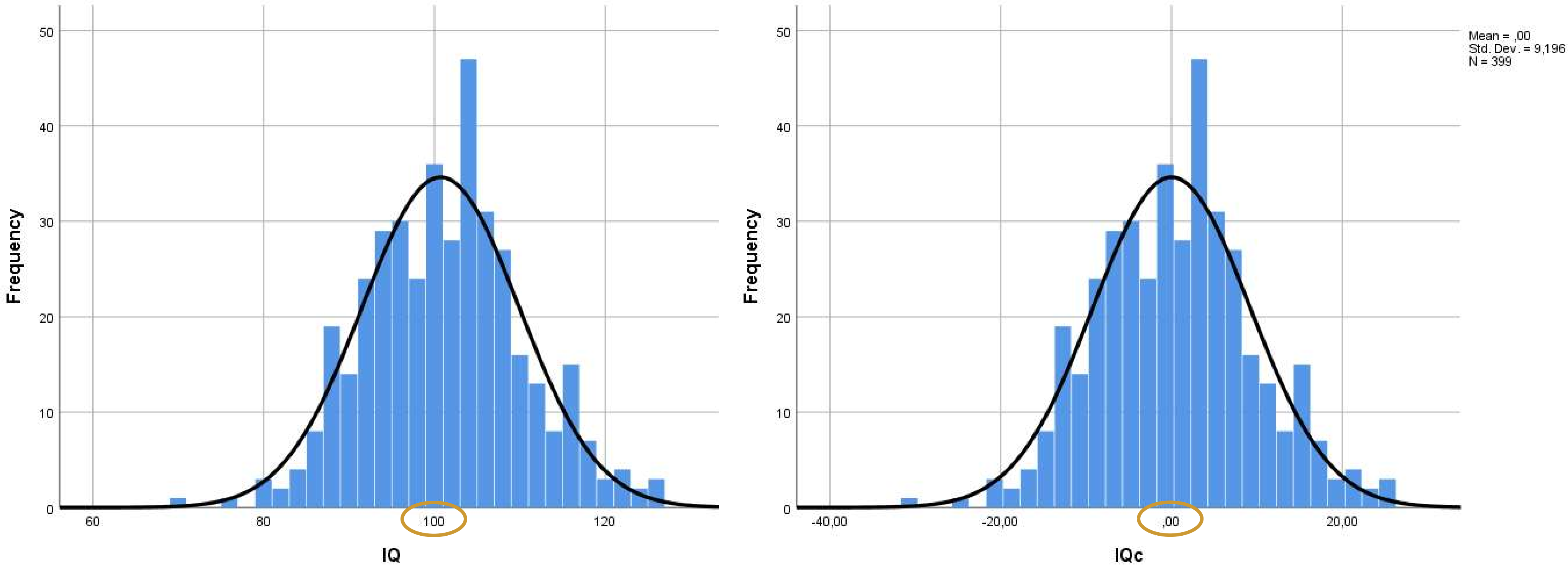
School 2: Slope = $\gamma_{10} + u_{12}$

School 3: Slope = $\gamma_{10} + u_{13}$

$$b_{0j} = \gamma_{00} + \gamma_{01} Schoolsize_j + u_{0j}$$
$$b_{1j} = \gamma_{10} + u_{1j}$$

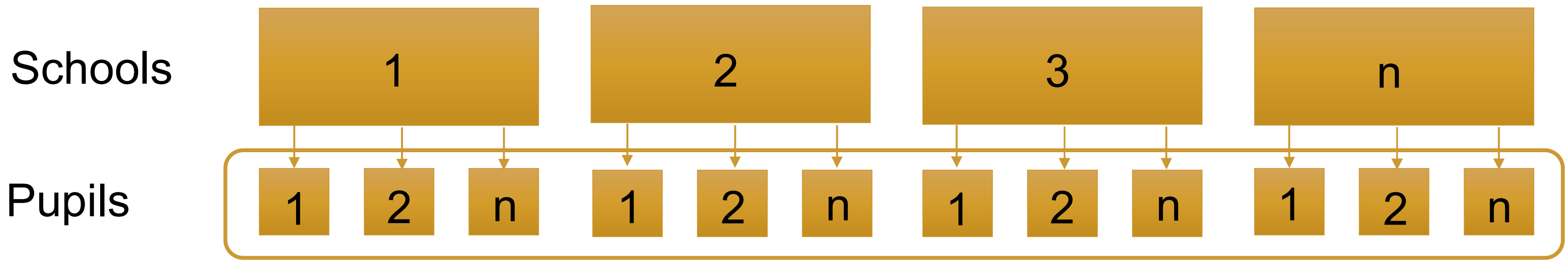
Tilburg University · tesc

Centering



Mean = ,00
Std. Dev. = 9,196
N = 399
Frequency
50
40
30
20
10
0
60
80
100
120
IQ
-40,00
-20,00
,00
20,00
IQc

Grand Mean Centering



DisAgg.sav [DataSet1] - IBM SPSS Statistics Data Editor

Visible: 6 of 6 Variables

	Class	Student	Gender	SA	IQ	Size	var	var
1	1,00	1	1	106	100	30		
2	1,00	2	1	94	88	30		
3	1,00	3	2	107	102	30		
4	1,00	4	1	121	116	30		
5	1,00	5	1	109	104	30		
6	1,00	6	2	100	95	30		
7	1,00	7	1	101	96	30		
8	1,00	8	1	95	89	30		
9	1,00	9	1	98	92	30		
10	1,00	10	1	114	110	30		
11	1,00	11	1	121	115	30		
12	1,00	12	1	107	101	30		
13	1,00	13	2	97	93	30		
14	1,00	14	1	111	107	30		
15	1,00	15	2	100	96	30		
16	1,00	16	2	91	87	30		
17	1,00	17	2	110	104	30		
18	1,00	18	1	97	90	30		
19	1,00	19	2	111	106	30		
20	1,00	20	1	107	102	30		
21	1,00	21	1	117	112	30		
22	1,00	22	2	104	99	30		
23	1,00	23	1	117	110	30		
24	1,00	24	1	109	105	30		
25	1,00	25	2	99	93	30		
26	1,00	26	2	112	107	30		
27	1,00	27	1	113	108	30		

Data View | Variable View

DisAggLvl2.sav [DisAggLvl2] - IBM SPSS Statistics Data Editor

Visible: 4 of 4 Variables

	Class	SA	IQ	Size	var	var	var	var
1	1	106,03	100,80	30				
2	2	110,46	105,42	24				
3	3	106,06	101,03	31				
4	4	106,76	101,79	29				
5	5	104,21	99,31	29				
6	6	104,58	99,75	24				
7	7	107,64	102,50	22				
8	8	106,43	101,29	21				
9	9	103,45	98,73	33				
10	10	108,32	102,89	19				
11	11	104,74	99,70	23				
12	12	104,45	99,48	31				
13	13	102,81	97,84	32				
14	14	108,87	103,52	23				
15	15	104,89	99,75	28				
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								

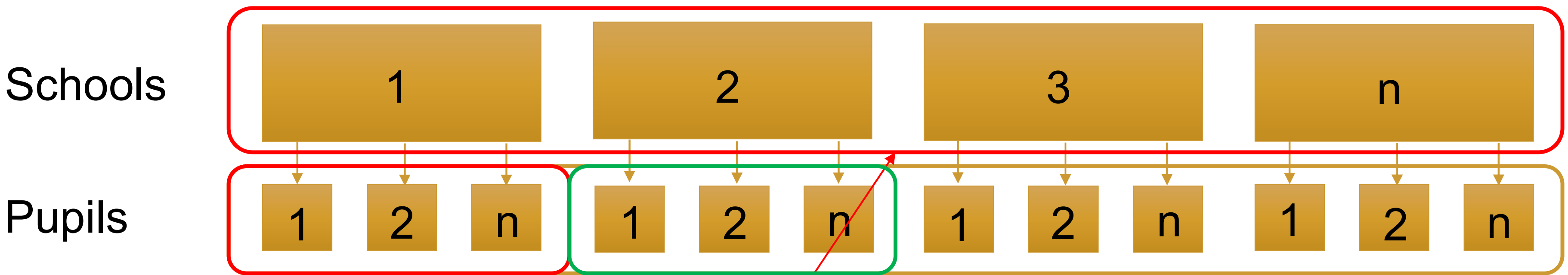
Data View | Variable View

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
IQ	399	70	126	100,72	9,196
Valid N (listwise)	399				

$$IQc = \boxed{IQ} - \boxed{100.72}$$

Group Mean Centering



DisAgg.sav [DataSet1] - IBM SPSS Statistics Data Editor

Visible: 6 of 6 Variables

	Class	Student	Gender	SA	IQ	Size	var	var
1	1,00	1	1	106	100	30		
2	1,00	2	1	94	88	30		
3	1,00	3	2	107	102	30		
4	1,00	4	1	121	116	30		
5	1,00	5	1	109	104	30		
6	1,00	6	2	100	95	30		
7	1,00	7	1	101	96	30		
8	1,00	8	1	95	89	30		
9	1,00	9	1	98	92	30		
10	1,00	10	1	114	110	30		
11	1,00	11	1	121	115	30		
12	1,00	12	1	107	101	30		
13	1,00	13	2	97	93	30		
14	1,00	14	1	111	107	30		
15	1,00	15	2	100	96	30		
16	1,00	16	2	91	87	30		
17	1,00	17	2	110	104	30		
18	1,00	18	1	97	90	30		
19	1,00	19	2	111	106	30		
20	1,00	20	1	107	102	30		
21	1,00	21	1	117	112	30		
22	1,00	22	2	104	99	30		
23	1,00	23	1	117	110	30		
24	1,00	24	1	109	105	30		
25	1,00	25	2	99	93	30		
26	1,00	26	2	112	107	30		
27	1,00	27	1	113	108	30		

Data View | Variable View

DisAggLvl2.sav [DisAggLvl2] - IBM SPSS Statistics Data Editor

Visible: 4 of 4 Variables

	Class	SA	IQ	Size	var	var	var	var
1	1	106,03	100,80	30				
2	2	110,46	105,42	24				
3	3	106,06	101,03	31				
4	4	106,76	101,79	29				
5	5	104,21	99,31	29				
6	6	104,58	99,75	24				
7	7	107,64	102,50	22				
8	8	106,43	101,29	21				
9	9	103,45	98,73	33				
10	10	108,32	102,89	19				
11	11	104,74	99,70	23				
12	12	104,45	99,48	31				
13	13	102,81	97,84	32				
14	14	108,87	103,52	23				
15	15	104,89	99,75	28				

Data View | Variable View

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
IQ	399	70	126	100,72	9,196
Valid N (listwise)	399				

$$IQc=\boxed{IQ}-\boxed{100.72}$$

$$IQc\text{-}group=IQ_{ij}-\overline{IQ}_{j}$$

Centering

- Group mean centering changes interpretation of variable:
  - No longer “raw” score.
  - Relative position within higher level unit.

- Used for so-called frog-pond effects:
  - Someone’s actual level of an attribute doesn’t matter, what matters is how one compares to the rest of their group (or pond).

Conclusion

- With random slopes important to think about the intercept, and therefore the meaning of a 0-score on predictors.

- Centering advised, but…
  - 2 methods: Grand mean centering and group mean centering.
  - Group mean centering complicates interpretation of model.

Practical 3

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?

$$SA_{ij} = b_{01}Class1 + b_{02}Class2 + \ \dots + b_{0x}ClassX + b_{1j}IQ_{ij} + e_{ij}$$

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

Fixed Effects Models

$$SA_{ij} = b_{01}Class1 + b_{02}Class2 + \ \dots + b_{0x}ClassX + b_{1j}IQ_{ij} + e_{ij}$$

- 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

$$SA_{ij} = b_{01}Class1 + b_{02}Class2 + \ \ldots + b_{0x}ClassX + b_{1j}IQ_{ij} + e_{ij}$$

- 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.

Three-level Models

Multilevel Regression

Level 1: $Pop_{ij} = b_{0j} + b_{1j}Ext_{ij} + e_{ij}$

Level 2: $b_{0j} = \gamma_{00} + \gamma_{01}Texp_j + u_{0j}$

$b_{1j} = \gamma_{10} + u_{1j}$

Combined:

$$
\begin{aligned}
Pop_{ij} &= b_{0j} + b_{1j}Ext_{ij} + eij \\
&= \gamma_{00} + \gamma_{10}Ext_{ij} + \gamma_{01}Texp_j + e_{ij} + u_{0j} + u_{1j}Ext_{ij}
\end{aligned}
$$

Tilburg University

tesc

Multilevel Regression

Level 1: $$Pop_{ijk} = b_{0jk} + b_{1jk}Ext_{ijk} + e_{ijk}$$

Level 2: $$b_{0jk} = \boxed{\gamma_{00k}} + \gamma_{01k}Texp_{jk} + u_{0jk}$$
$$b_{1jk} = \boxed{\gamma_{10k}} + u_{1jk}$$

Level 3: $$\boxed{\gamma_{00k}} = \gamma_{000} + v_{ok}$$
$$\boxed{\gamma_{10k}} = \gamma_{100} + v_{1k}$$

Combined:

$$\begin{aligned} Pop_{ijk} &= b_{0jk} + b_{1jk}Ext_{ijk} + e_{ij} \\ &= \gamma_{000} + \gamma_{100}Ext_{ijk} + \gamma_{01k}Texp_{jk} + e_{ijk} + u_{0jk} + u_{1jk}Ext_{ijk} + v_{ok} \ + v_{1k}Ext_{ijk} \end{aligned}$$

TILBURG UNIVERSITY

tesc

Multilevel Regression

Level 1: $$Pop_{ijk} = b_{0jk} + b_{1jk}Ext_{ijk} + e_{ijk}$$

Level 2: $$\begin{aligned} b_{0jk} &= \gamma_{00k} + \gamma_{01k}Texp_{jk} + u_{0jk} \\ b_{1jk} &= \gamma_{10k} + u_{1jk} \end{aligned}$$

Level 3: $$\begin{aligned} \gamma_{00k} &= \gamma_{000} + v_{ok} \\ \gamma_{10k} &= \gamma_{100} + v_{1k} \end{aligned}$$

Combined:

$$\begin{aligned} Pop_{ijk} &= b_{0jk} + b_{1jk}Ext_{ijk} + e_{ij} \\ &= \gamma_{000} + \gamma_{100}Ext_{ijk} + \gamma_{01k}Texp_{jk} + e_{ij} \ + u_{0jk} + u_{1jk}Ext_{ijk} + v_{ok} \ + v_{1k}Ext_{ijk} \end{aligned}$$

Multilevel Regression

Level 1: $Pop_{ijk} = b_{0jk} + b_{1jk}Ext_{ijk} + e_{ijk}$

Level 2:
$$b_{0jk} = \boxed{\gamma_{00k}} + \gamma_{01k}Texp_{jk} + u_{0jk}$$
$$b_{1jk} = \boxed{\gamma_{10k}} + u_{1jk}$$

Level 3:
$$\boxed{\gamma_{00k}} = \boxed{\gamma_{000}} + \boxed{v_{ok}}$$
$$\boxed{\gamma_{10k}} = \boxed{\gamma_{100}} + \boxed{v_{1k}}$$

Combined:

$$\begin{aligned} Pop_{ijk} &= b_{0jk} + b_{1jk}Ext_{ijk} + e_{ij} \\ &= \gamma_{000} + \gamma_{100}Ext_{ijk} + \gamma_{01k}Texp_{jk} + e_{ijk} + u_{0jk} + \boxed{u_{1jk}Ext_{ijk}} + v_{ok} + \boxed{v_{1k}Ext_{ijk}} \end{aligned}$$

Multilevel Regression

Level 1: $Pop_{ijk} = b_{0jk} + b_{1jk}Ext_{ijk} + e_{ijk}$

Level 2: $b_{0jk} = \gamma_{00k} + \gamma_{01k}Texp_{jk} + u_{0jk}$

$b_{1jk} = \gamma_{10k} + u_{1jk}$

Level 3: $\gamma_{00k} = \gamma_{000} + v_{ok}$

$\gamma_{10k} = \gamma_{100} + v_{1k}$

Combined:

$$
\begin{aligned}
Pop_{ijk} &= b_{0jk} + b_{1jk}Ext_{ijk} + e_{ijk} \\
&= \gamma_{000} + \gamma_{100}Ext_{ijk} + \gamma_{01k}Texp_{jk} + e_{ijk} + u_{0jk} + u_{1jk}Ext_{ijk} + v_{ok} + v_{1k}Ext_{ijk}
\end{aligned}
$$

Multilevel Regression

Level 1: $Pop_{ijk} = b_{0jk} + b_{1jk}Ext_{ijk} + e_{ijk}$

Level 2: $b_{0jk} = \gamma_{00k} + \boxed{\gamma_{01k}}Texp_{jk} + u_{0jk}$

$b_{1jk} = \gamma_{10k} + u_{1jk}$

Level 3: $\gamma_{00k} = \gamma_{000} + v_{ok}$

$\gamma_{10k} = \gamma_{100} + v_{1k}$

Combined:

$$
\begin{aligned}
Pop_{ijk} &= b_{0jk} + b_{1jk}Ext_{ijk} + e_{ijk} \\
&= \gamma_{000} + \gamma_{100}Ext_{ijk} + \gamma_{01k}Texp_{jk} + e_{ijk} + u_{0jk} + u_{1jk}Ext_{ijk} + v_{ok} + v_{1k}Ext_{ijk}
\end{aligned}
$$

Multilevel Regression

Level 1: $Pop_{ijk} = b_{0jk} + b_{1jk}Ext_{ijk} + e_{ijk}$

Level 2: $b_{0jk} = \gamma_{00k} + \boxed{\gamma_{01k}}Texp_{jk} + u_{0jk}$

$b_{1jk} = \gamma_{10k} + u_{1jk}$

Level 3: $\gamma_{00k} = \gamma_{000} + v_{ok}$

$\gamma_{10k} = \gamma_{100} + v_{1k}$

$\boxed{\gamma_{01k}} = \gamma_{010} + v_{2k}$

Combined:

$$Pop_{ijk} = b_{0jk} + b_{1jk}Ext_{ijk} + e_{ij}$$
$$= \gamma_{000} + \gamma_{100}Ext_{ijk} + \boxed{\gamma_{010}}Texp_{jk} + e_{ijk} + u_{0jk} + u_{1jk}Ext_{ijk} + v_{ok} + v_{1k}Ext_{ijk} + \boxed{v_{2k}Texp_{jk}}$$

Intraclass correlation

- $\sigma^2_{v_0}$ variance of the third level errors v_{0j}
- $\sigma^2_{u_0}$ variance of the second level errors u_{0j}
- σ^2_{e} variance of the first level errors e_{ij}

$$\rho_{class} = \frac{\sigma^2_{u_0}}{\sigma^2_{v_0} + \sigma^2_{u_0} + \sigma^2_{e}} \qquad \rho_{school} = \frac{\sigma^2_{v_0}}{\sigma^2_{v_0} + \sigma^2_{u_0} + \sigma^2_{e}}$$

- The ‘percentage’ variance at the second and third level

$$\rho_{class} = \frac{\sigma^2_{v_0} + \sigma^2_{u_0}}{\sigma^2_{v_0} + \sigma^2_{u_0} + \sigma^2_{e}}$$

- The expected correlation between the observations on the dependent variable of two randomly chosen units that are in the same class (and thus in the same school).

Extra difficulty with 3 level models

- What do you do when there is a significant amount of variance on level 3, but not level 2?

Extra difficulty with 3 level models

- What do you do when there is a significant amount of variance on level 3, but not level 2?

- The “higher you go” the smaller your sample size.
  - What do you do if the number of level 3 units is very small?

Fixed Effect Models

Level 1: $Pop_{ijk} = b_{0j} + b_{1jk}Ext_{ij} + e_{ij}$

Level 2: $b_{0j} = \gamma_{01}School1 + \gamma_{02}School2 + \ \ldots + \gamma_{0x}SchoolX + \gamma_{0,x+1}Texp_{jk} + u_{0s1} + \ldots + u_{0sx}$

$b_{1j} = \gamma_{11}School1 + \gamma_{12}School2 + \ \ldots + \gamma_{1x}SchoolX + u_{1s1} + \ \ldots \ + \ u_{1s}$

Combined:

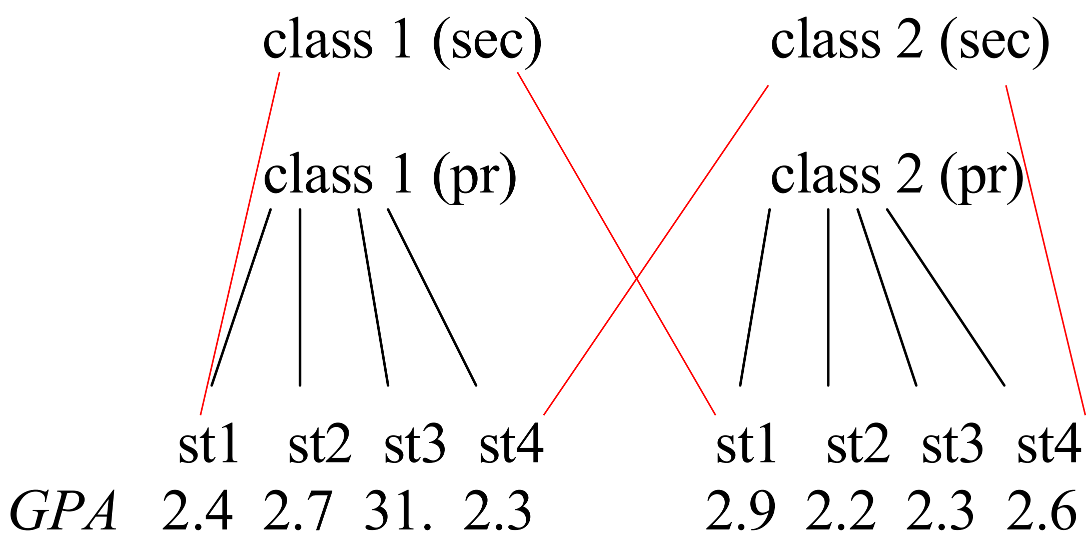
$$
\begin{aligned}
Pop_{ijk} &= b_{0jk} + b_{1jk}Ext_{ijk} + e_{ijk} \\
&= \gamma_{01}School1 + \gamma_{02}School2 + \ \ldots + \gamma_{0x}SchoolX + \gamma_{0,x+1}Texp_{jk} \\
&\quad \gamma_{11}School1 * Ext_{ij} + \gamma_{12}School2 * Ext_{ij} + \ \ldots + \gamma_{1x}SchoolX * Ext_{ij} + \\
&\quad e_{ij} + u_{0s1} + \ \ldots \ + u_{0sx} + u_{1s1} + \ \ldots \ + \ u_{1sx} + u_{1j}Ext_{ij}
\end{aligned}
$$

Cross-Nested Data

Cross-Classified Models

- Sometimes the nesting structure is not constant.
  - You follow students as the transition from primary school to secondary school.
  - Some students that attend the same primary school will attend different secondary schools.
  - So students nested in a cross-classification of primary and secondary schools.

Cross-classified data



Cross-Classified Models

- $Y_{i(jk)}$ Score of student i within cross-classification of primary school j and secondary school k

$$Y_{i(jk)} = \beta_{0(jk)} + e_{i(jk)}$$

- $\beta_{0(jk)}$ is the intercept/overall mean

- $e_{i(jk)}$ is the residual error term

- the subscript *(jk)* is written in parentheses to indicate that they are conceptually at the same level

TILBURG UNIVERSITY tesc

Cross-Classified Models

- The intercept $\beta_{0(jk)}$ varies independently across both primary and secondary schools, so

$$Y_{i(jk)} = \beta_{0(jk)} + e_{i(jk)}$$

$$\beta_{0(jk)} = \gamma_{00} + u_{0j} + v_{ok}$$

- u_{0j} error for primary schools

- v_{ok} error for secondary schools

Cross-Classified Models

- Individual-level explanatory variables can be added to the equation.
  - Their regression slopes may be allowed to vary across primary and/or secondary schools.

$$Y_{i(jk)} = \beta_{0(jk)} + \beta_{1(jk)} X_{i(jk)} + e_{i(jk)}$$

- School-level variables can also be added,
  - Used to explain variation in the slopes of individual-level variables across schools
  - Similar to ordinary multilevel regression models.

Cross-Classified Models

- How do you specify these things?

- Do they sort of remind you of a model we saw earlier when looking at the equations?

$$Y_{i(jk)} = \beta_{0(jk)} + e_{i(jk)}$$

$$\beta_{0(jk)} = \gamma_{00} + u_{0j} + v_{ok}$$

Cross-Classified Models

- How do you specify these things?
- We can start by ignoring the secondary school level, specify the individual and primary school level as usual.
  - individuals at the first level and primary schools at the second level.
- But now what?
  - We need to add another level, without actually adding another level.
  - Have we seen this before?

Fixed Effect Models

Level 1: $Pop_{ijk} = b_{0j} + b_{1jk}Ext_{ij} + e_{ij}$

Level 2: $b_{0j} = \gamma_{01}School1 + \gamma_{02}School2 + \ \dots + \gamma_{0x}SchoolX + \gamma_{0,x+1}Texp_{jk} + u_{0s1} + \dots + u_{0sx}$

$b_{1j} = \gamma_{11}School1 + \gamma_{12}School2 + \ \dots + \gamma_{1x}SchoolX + u_{1s1} + \ \dots \ + \ u_{1sx}$

Combined:

$$
\begin{aligned}
Pop_{ijk} &= b_{0jk} + b_{1jk}Ext_{ijk} + e_{ijk} \\
&= \gamma_{01}School1 + \gamma_{02}School2 + \ \dots + \gamma_{0x}SchoolX + \gamma_{0,x+1}Texp_{jk} \\
&\quad \gamma_{11}School1 * Ext_{ij} + \gamma_{12}School2 * Ext_{ij} + \ \dots + \gamma_{1x}SchoolX * Ext_{ij} + \\
&\quad e_{ij} + u_{0s1} + \ \dots \ + u_{0sx} + u_{1s1} + \ \dots \ + \ u_{1sx} + u_{1j}Ext_{ij}
\end{aligned}
$$

Cross-Classified Models

- We can start by ignoring the secondary school level, specify the individual and primary school level as usual.
  - individuals at the first level and primary schools at the second level.

- To create a place to specify the crossed effects of the secondary school level, we introduce a third ‘dummy’ level

- At the pupil level: specify a full set of dummy variables to indicate all of the secondary schools.

Cross-Classified Models

- At the pupil level: specify a full set of dummy variables to indicate all of the secondary schools.
- But this has one problem left! Which one?

Fixed Effect Models

Level 1: $Pop_{ijk} = b_{0j} + b_{1jk}Ext_{ij} + e_{ij}$

Level 2:

$$b_{0j} = \boxed{\gamma_{01}School1 + \gamma_{02}School2 + \ \dots + \gamma_{0x}SchoolX} + \gamma_{0,x+1}Texp_{jk} + \boxed{u_{0s1} + \dots + u_{0sx}}$$

$$b_{1j} = \gamma_{11}School1 + \gamma_{12}School2 + \ \dots + \gamma_{1x}SchoolX + u_{1s1} + \ \dots \ + \ u_{1sx}$$

Combined:

$$\begin{aligned} Pop_{ijk} &= b_{0jk} + b_{1jk}Ext_{ijk} + e_{ijk} \\ &= \gamma_{01}School1 + \gamma_{02}School2 + \ \dots + \gamma_{0x}SchoolX + \gamma_{0,x+1}Texp_{jk} \\ &\quad \gamma_{11}School1 * Ext_{ij} + \gamma_{12}School2 * Ext_{ij} + \ \dots + \gamma_{1x}SchoolX * Ext_{ij} + \\ &\quad e_{ij} + u_{0s1} + \ \dots \ + u_{0sx} + u_{1s1} + \ \dots \ + \ u_{1sx} + u_{1j}Ext_{ij} \end{aligned}$$

Multilevel Regression

Level 1: $$Pop_{ij} = b_{0j} + b_{1j}Ext_{ij} + e_{ij}$$

Level 2: $$b_{0j} = \gamma_{00} + \gamma_{01}Texp_j + \boxed{u_{0j}}$$

$$b_{1j} = \gamma_{01} + u_{1j}$$

Combined:

$$\begin{aligned} Pop_{ij} &= b_{0j} + b_{1j}Ext_{ij} + eij \\ &= \gamma_{00} + \gamma_{01}Ext_{ij} + \gamma_{01}Texp_j + e_{ij} + u_{0j} + u_{1j}Ext_{ij} \end{aligned}$$

Cross-Classified Models

- The fixed regression coefficients of these dummies are excluded from the model,
  - their slopes are allowed to vary at the third, ‘dummy’, level.
  - covariances between dummies all constrained to be zero,
  - variances all constrained to be equal.
- Thus, in the end we estimate one variance component for the secondary schools
- By putting the secondary schools in a separate “level” we assure that there are no covariances between residuals for the primary and the secondary schools.

Fixed Effect Models

$$b_{0j} = \boxed{\gamma_{01}School1 + \gamma_{02}School2 + \ \dots + \gamma_{0x}SchoolX} + \gamma_{0,x+1}Texp_{jk} + \boxed{u_{0s1} + \dots + u_{0sx}}$$

$$\gamma_{01} \text{-} \gamma_{0x} = 0$$

TILBURG UNIVERSITY

tesc

Fixed Effect Models

$$b_{0j} = \boxed{\gamma_{01}School1 + \gamma_{02}School2 + \ \ldots + \gamma_{0x}SchoolX} + \gamma_{0,x+1}Texp_{jk} + \boxed{u_{0s1} + \ldots + u_{0sx}}$$

$$\gamma_{01} - \gamma_{0x} = 0$$

$$b_{0j} = \gamma_{0,x+1}Texp_{jk} + u_{0s1} + \ldots + u_{0sx}$$

Fixed Effect Models

$$b_{0j} = \boxed{\gamma_{01}School1 + \gamma_{02}School2 + \ \dots + \gamma_{0x}SchoolX} + \gamma_{0,x+1}Texp_{jk} + \boxed{u_{0s1} + \ \dots + u_{0sx}}$$

$$\gamma_{01} \text{-} \gamma_{0x} = 0$$

$$b_{0j} = \gamma_{0,x+1}Texp_{jk} + u_{0s1} + \ \dots + u_{0sx}$$

$$sd(u_{0s1}) = sd(u_{0s2}) = \cdots = sd(u_{0sx})$$

TILBURG UNIVERSITY

tesc

Cross-Classified Models

$$\text{achiev}_{ijk} \sim \text{N}(XB, \Omega)$$

$$\begin{aligned}\text{achiev}_{ijk} = {} & \beta_{30ij}\text{const} + \text{v}_{0k}\text{c101}_{ijk} + \text{v}_{1k}\text{c102}_{ijk} + \text{v}_{2k}\text{c103}_{ijk} + \text{v}_{3k}\text{c104}_{ijk} + \text{v}_{4k}\text{c105}_{ijk} \\ & + \text{v}_{5k}\text{c106}_{ijk} + \text{v}_{6k}\text{c107}_{ijk} + \text{v}_{7k}\text{c108}_{ijk} + \text{v}_{8k}\text{c109}_{ijk} + \text{v}_{9k}\text{c110}_{ijk} \\ & + \text{v}_{10k}\text{c111}_{ijk} + \text{v}_{11k}\text{c112}_{ijk} + \text{v}_{12k}\text{c113}_{ijk} + \text{v}_{13k}\text{c114}_{ijk} + \text{v}_{14k}\text{c115}_{ijk} \\ & + \text{v}_{15k}\text{c116}_{ijk} + \text{v}_{16k}\text{c117}_{ijk} + \text{v}_{17k}\text{c118}_{ijk} + \text{v}_{18k}\text{c119}_{ijk} + \text{v}_{19k}\text{c120}_{ijk} \\ & + \text{v}_{20k}\text{c121}_{ijk} + \text{v}_{21k}\text{c122}_{ijk} + \text{v}_{22k}\text{c123}_{ijk} + \text{v}_{23k}\text{c124}_{ijk} + \text{v}_{24k}\text{c125}_{ijk} \\ & + \text{v}_{25k}\text{c126}_{ijk} + \text{v}_{26k}\text{c127}_{ijk} + \text{v}_{27k}\text{c128}_{ijk} + \text{v}_{28k}\text{c129}_{ijk} + \text{v}_{29k}\text{c130}_{ijk}\end{aligned}$$

$$\beta_{30ij} = 6.349(0.078) + u_{30jk} + e_{30ijk}$$

$$\begin{bmatrix} v_{0k} \\ v_{1k} \\ v_{2k} \\ v_{3k} \\ v_{4k} \end{bmatrix} \qquad \begin{bmatrix} 0.065(.022) & & & & \\ 0 & 0.065(.022) & & & \\ 0 & 0 & 0.065(.022) & & \\ 0 & 0 & 0 & 0.065(.022) & \\ 0 & 0 & 0 & 0 & \end{bmatrix}$$

Cross-Classified Models

Model	Intercept-only Coeff. (s.e.)	+ pupil vars Coeff. (s.e.)	+ school vars Coeff. (s.e.)	+ *ses* random Coeff. (s.e.)
Fixed part				
Intercept	6.35 (.08)	5.76 (.11)	5.52 (.19)	5.53 (.14)
Pupil gender		0.26 (.05)	0.26 (.05)	0.25 (.05)
Pupil ses		0.11 (.02)	0.11 (.02)	0.11 (.02)
Primary denom			0.20 (.12)	0.20 (.12)
Secondary denom			0.18 (.10)	0.17 (.09)
Random part				
$\sigma^2_{\text{int/pupil}}$	0.51 (.02)	0.47 (.02)	0.47 (.02)	0.46 (.02)
$\sigma^2_{\text{int/primary}}$	0.17 (.04)	0.17 (.04)	0.16 (.04)	0.14 (.08)
$\sigma^2_{\text{int/secondary}}$	0.07 (.02)	0.06 (.02)	0.06 (.02)	0.05 (.02)
$\sigma^2_{\text{ses/primary}}$				0.008 (.004)
Deviance	2317.8	2243.5	2237.5	2224.5
AIC	2325.8	2255.5	2253.5	2244.5

Cross-Classified Models

- Alternatively you might have network data.
- E.g. you ask all members of a group to rate all other members
  - how much they would like to share some activity with the rated person
  - Sociometric rating

Sociometric rating

	group	child	age	sex	grsize	rating1	rating2	rating3	rating4	rating5	rating6	rating7	rating8	rating9
1	1	1	8	1	7	.	3	6	4	4	7	6	.	.
2	1	2	10	1	7	5	.	6	4	5	7	5	.	.
3	1	3	11	1	7	4	6	.	4	5	7	6	.	.
4	1	4	9	0	7	4	4	6	.	5	7	5	.	.
5	1	5	11	0	7	5	5	6	5	.	7	6	.	.
6	1	6	10	1	7	4	5	6	3	4	.	6	.	.
7	1	7	10	1	7	3	5	6	5	3	6	.	.	.
8	2	1	9	0	9	.	3	5	3	4	6	6	4	5
9	2	2	9	0	9	2	.	4	5	6	5	4	4	5
10	2	3	9	0	9	5	3	.	4	3	6	5	4	6
11	2	4	8	1	9	3	2	5	.	6	6	5	3	4
12	2	5	9	1	9	4	4	5	5	.	5	7	4	5
13	2	6	9	0	9	3	4	4	4	4	.	5	4	5
14	2	7	9	1	9	4	4	6	5	6	5	.	4	5
15	2	8	11	0	9	3	4	5	4	5	6	6	.	5
16	2	9	8	1	9	3	4	5	5	4	6	7	5	.
17	3	1	11	0	5	.	5	7	5	6	.	.	.	.
18	3	2	11	0	5	5	.	7	6	6	.	.	.	.
19	3	3	13	1	5	5	5	.	6	8	.	.	.	.
20	3	4	12	1	5	4	4	6	.	6	.	.	.	.

Cross-Classified Models

- DV: Ratings
- Nested:

TILBURG UNIVERSITY

tesc

Cross-Classified Models

- DV: Ratings
- Nested: Cross-clasification of senders and receivers
- Can be further nested within groups

Sociometric rating

	group	sender	receiver	rating	agesend	sexsend	agerec	sexrec	grsize
1	1	1	2	3	8	1	10	1	7
2	1	1	3	6	8	1	11	1	7
3	1	1	4	4	8	1	9	0	7
4	1	1	5	4	8	1	11	0	7
5	1	1	6	7	8	1	10	1	7
6	1	1	7	6	8	1	10	1	7
7	1	2	1	5	10	1	8	1	7
8	1	2	3	6	10	1	11	1	7
9	1	2	4	4	10	1	9	0	7
10	1	2	5	5	10	1	11	0	7
11	1	2	6	7	10	1	10	1	7
12	1	2	7	5	10	1	10	1	7
13	1	3	1	4	11	1	8	1	7
14	1	3	2	6	11	1	10	1	7
15	1	3	4	4	11	1	9	0	7
16	1	3	5	5	11	1	11	0	7
17	1	3	6	7	11	1	10	1	7
18	1	3	7	6	11	1	10	1	7
19	1	4	1	4	9	0	8	1	7
20	1	4	2	4	9	0	10	1	7

Cross-Classified Models

- $Y_{i(jk)l}$ Score of student i within cross-classification of sender j and receiver k, and within group l

$$Y_{i(jk)l} = \beta_{0(jk)l} + e_{i(jk)l}$$

- $\beta_{0(jk)l}$ is the intercept/overall mean
- $e_{i(jk)l}$ is the residual error term
- the subscript *(jk)* is written in parantheses to indicate that they are conceptually at the same level

Cross-Classified Models

- The intercept $\beta_{0(jk)l}$ varies independently across both senders, receivers, and group, so

$$Y_{i(jk)l} = \beta_{0(jk)l} + e_{i(jk)l}$$

$$\beta_{0(jk)l} = \beta_{0l} + u_{0jl} + v_{okl}$$

$$\beta_{0l} = \gamma_{00} + f_{0l}$$

Cross-Classified Models

$$Y_{i(jk)l} = \beta_{0(jk)l} + e_{i(jk)l}$$

$$\beta_{0(jk)} = \beta_{0l} + u_{0jl} + v_{okl}$$

$$\beta_{0l} = \gamma_{00} + f_{0l}$$

- Score consists of:
  - Overall mean γ_{00},
  - residual error term f_{0l} for group l,
  - individual-level residual error terms u_{jl} for sender j in group l
  - individual-level residual error v_{kl} for receiver k in group l, and
  - the measurement-level error term $e_{i(jk)l}$.

TILBURG UNIVERSITY

tesc

Cross-Classified Models

- The crossed effects of the receiver level are incorporated using dummies.
- At the lowest level, the ratings, we specify:
  - Dummy variables that indicate the receivers.
  - The fixed coefficients of these dummies are excluded from the model, but their slopes are allowed to vary.
  - Since the cross-classification is nested within the sociometric groups, the slopes of the dummy variables are set to vary at a third group level.
  - In addition, the covariances between the receiver dummies are constrained to be zero, and
  - the variances are constrained to be equal.

Cross-Classified Models

- The covariances between the receiver dummies are constrained to be zero, and the variances are constrained to be equal.
  - Thus, we estimate one variance component for the receivers.

- By putting the variance term(s) for the receivers on a “separate” level we assure that there are no covariances between the residuals for the sender and the receiver level.

Cross-Classified Models

- Both sender and receiver characteristics like *age* and *gender* and group characteristics like *group size* can be added to the model as predictors.

- Child characteristics may be allowed to have random slopes at the group level.

- The analysis proceeds along exactly the same lines as outlined for the cross-classification of primary and secondary schools.

Cross-Classified Models: Hard in some software!

- Since the third ‘group’ level is already used to specify the random variation for the receiver dummies:
  - we must make sure that the intercept and possible slope variation at the ‘real’ group level are not correlated with the dummies.
  - This can be accomplished by adding the appropriate constraints to the model.

- When the software supports more than three levels (e.g., R/MLWin/HLM), the same result can be accomplished more conveniently:
  - Add a fourth level to the model; also for the groups
  - Used for the random part at the real group level.
  - Conceptually we still have three levels, with senders and receivers crossed at the second level

Cross-Classified Models

- Individual-level explanatory variables can be added to the equation.
  - Their regression slopes may be allowed to vary across primary and/or secondary schools.

$$Y_{i(jk)} = \beta_{0(jk)} + \beta_{1(jk)} X_{i(jk)} + e_{i(jk)}$$

- School-level variables can also be added,
  - Used to explain variation in the slopes of individual-level variables across schools
  - Similar to ordinary multilevel regression models.

Non-Normal Data

Dichotomous data and proportions

- Standard regression: continuous outcome variable with normal distribution of errors
  - What do we do again if one or more violations are violated?

- Some outcome variables *must* violate this assumption

  - Dichotomous outcome variable

  - Outcome variable is a proportion

Nonlinear Regression models

- Classic solution: transform the dependent variable
  - Proportions: $f(p) = \text{logit}(p) = \ln((p)/(1-p))$
  - Breaks down if $p = 0,1$
    - Does not work with dichotomous variable

- Modern solution: use generalized linear model
  - i.e., Logistic regression

Nonlinear Regression models

- Generalized linear model:
- Don’t transform DV and then run a “normal” regression.
- Run a “normal” regression and then transform the prediction of the analysis!

Generalized Linear Model

Three common generalized linear models

(a) Linear regression	(b) Logistic regression	(c) Poisson regression
$\alpha + \beta x_i$	$logistic(\alpha + \beta x_i)$	$exp(\alpha + \beta x_i)$
$\downarrow$	$\downarrow$	$\downarrow$
$y_i \sim Normal(\mu_i, \sigma)$	$y_i \sim Bernoulli(p_i)$	$y_i \sim Poisson(\lambda_i)$

$$y = b_0 + b_1 X_1 + b_2 X_2 + e$$

$$\text{pred} = b_0 + b_1 X_1 + b_2 X_2 + e$$

$$\text{logistic(pred)} = \frac{exp^{(pred)}}{1 + exp^{(pred)}}$$

$$y \sim Bernoulli(logistic(pred))$$

Generalized Linear Model

Three common generalized linear models

$$\alpha + \beta x_i \downarrow y_i \sim Normal(\mu_i, \sigma)$$

(a) Linear regression

$$logistic(\alpha + \beta x_i) \downarrow y_i \sim Bernoulli(p_i)$$

(b) Logistic regression

$$exp(\alpha + \beta x_i) \downarrow y_i \sim Poisson(\lambda_i)$$

(c) Poisson regression

- The output of $b_0 + b_1 X_1 + b_2 X_2 + e$ can be any value ranging from negative to positive infinity.
- A 'link' function is introduced to ensure that only sensible parameter values are allowed.
- Logistic regression: the logistic function ensures that values from negative to positive infinity are compressed into the range [0, 1].

$$\text{pred} = b_0 + b_1 X_1 + b_2 X_2 + e$$

$$\text{logistic(pred)} = \frac{exp^{(pred)}}{1 + exp^{(pred)}}$$

$$y \sim Bernoulli(logistic(pred))$$

Generalized Linear Model

- Easy to do in R!!

- Just have to change the distribution we use in an analysis!
  - Just like changing from normal to t-distribution
  - But now use Bernoulli or Poisson or…

- Just need to remember that the parameter-estimates are in terms of the non-transformed “pred” part, and so we need to transform them to the “original” scale for interpretation.

Example: Multilevel Logistic model

- Thailand Education Data
  - 8582 pupils in 411 primary schools
  - Outcome variable dichotomous
  - Repeat a class: yes (1) or no (0)

  - Pupil level predictors
  - *sex* (1 = male, 0 = female)
  - *pre-primary education* (PPED: 1 = yes, 0 = no)
  - School level predictor
  - Mean *SES*

Example: Multilevel Logistic model

Final estimation of fixed effects: (Unit-specific model)

Fixed Effect	Coefficient	Standard Error	T-ratio	Approx. d.f.	P-value
For INTRCPT1, B0					
INTRCPT2, G00	-2.038493	0.093685	-21.759	355	0.000
For SEX slope, B1					
INTRCPT2, G10	0.509270	0.073925	6.889	7513	0.000
For PPED slope, B2					
INTRCPT2, G20	-0.609289	0.095220	-6.399	7513	0.000

$$p_{ij} = logistic(-2.04 + 0.51 sex - 0.61 pped)$$

Example: Multilevel Logistic model

- Interpretation: $p_{ij} = logistic(-2.04 + 0.51 sex - 0.61 pped)$

  - girls, no pped: predict: –2.04 on the underlying continuous scale.
  Transform: $g(x) = \dfrac{e^x}{1+e^x}$ = 0.115

  - boys, no pped: predict: -2.04 + 0.51 = -1.53
  Transformed: 0.178

  - The estimated repeat rate for the girls is 11.5% and for
  the boys 17.8% (no pped).

Example: Multilevel Logistic model

- Interpretation: $p_{ij} = logistic(-2.04 + 0.51 sex - 0.61 pped)$
  - girls, with pped: predict: –2.65 on the underlying continuous scale
    Transformed: 0.066

  - boys, with pped: predict: -2.04 + 0.51 – 0.61 = -2.14
    Transformed: 0.105

- The estimated repeat rate for the girls with pped is 6.6% and for the boys 10.5%.