PSYC40005 - 2018
ADVANCED DESIGN AND DATA ANALYSIS

Lecture 9:
Multilevel modelling 1

Geoff Saw

Melbourne School of Psychological Sciences
University of Melbourne
Redmond Barry Room 1113
gsaw@unimelb.edu.au

The agenda for this lecture

1. Multiple levels: Macro and micro
2. Changing two levels to one
3. Random effects ANOVA
4. Random intercept multilevel model

GOALS OF THIS LECTURE

- To introduce the idea of different levels in research data
- To show how aggregating or disaggregating data across levels does not permit confident inference
- To extend ANOVA from fixed to random effects
- To introduce a random intercept model for grouped data
- To show how these models may be fitted using SPSS

Section 1: Lecture 9

MULTIPLE LEVELS: MACRO & MICRO

- ❑ Multilevel modeling
- ❑ Macro-micro propositions

The right level?

For example: Constructs in organizational psychology:
Are they properties of individuals or properties of the organization?

Some of the history of ***Organizational climate:***

- James & Jones (1974): distinguished between psychological (at the individual level) and organizational climate (organizational level)
- Subsequent methodological debates:

  How to "aggregate" individual measures of *psychological climate* to measure *organizational climate*? Is such an aggregation a meaningful organizational collective or simply a statistical artefact?

  Glick, 1985; Jackofsky & Slocum, 1990; Payne, 1990
- The search for meaningful organizational collectives as a basis for justifying the collective climate construct.
  - Combination of multi-levels, networks and shared knowledge. Gonzalez-Roma, 1999; Young & Parker, 1999.

Multilevel modeling

Models that permit constructs at more than one level.

Individuals “nested” in groups

Predict individual outcomes from other individual variables as well as group level variables, taking into account the grouping structure

Lower level: micro. Upper level: macro

Macro-level	Micro-level
universities	lecturers
classes	students
neighbourhoods	families
firms	employees
Jawbones	teeth
families	children
litters	animals
doctors	patients
participants	measurements
interviewers	respondents
Judges	suspects

From Snijders & Bosker (2012), p.9

Multilevel modeling

Models that permit constructs at more than one level.

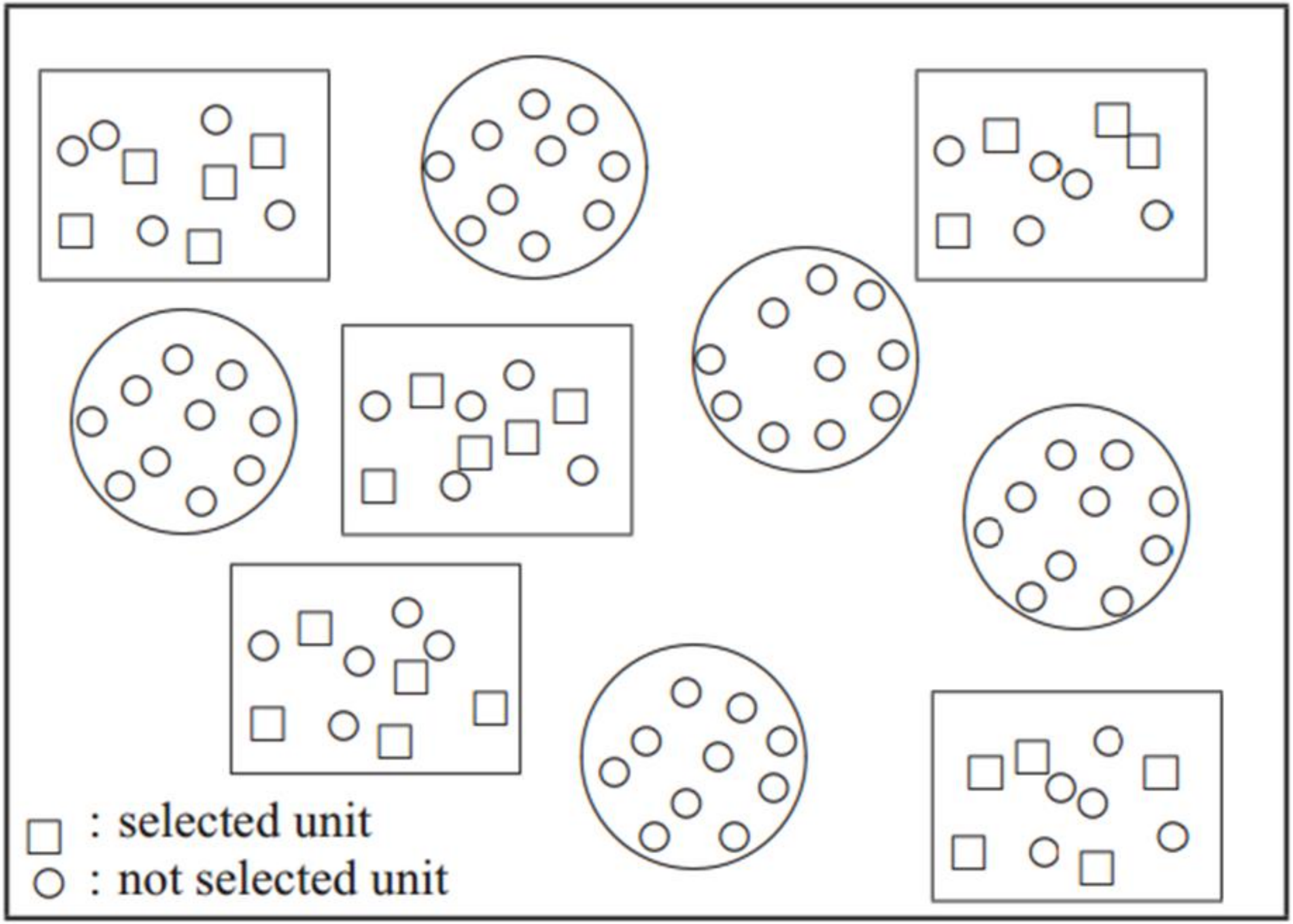
Individuals "nested" in groups

Predict individual outcomes from other individual variables as well as group level variables, taking into account the grouping structure

Lower level: micro. Upper level: macro

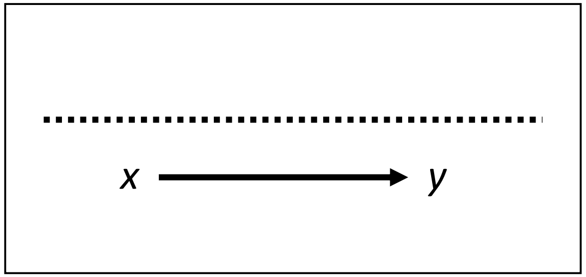
- *The grouping structure sets up dependence among observations.*
- *Sometimes dependence is just a nuisance – you may want independent observations so you can use familiar statistical procedures. But for practical reasons you might need to sample in multiple stages, which can create dependence.*
- *Sometimes dependence in itself is the phenomenon we're interested in. We want to understand group level effects.*

Multistage sampling

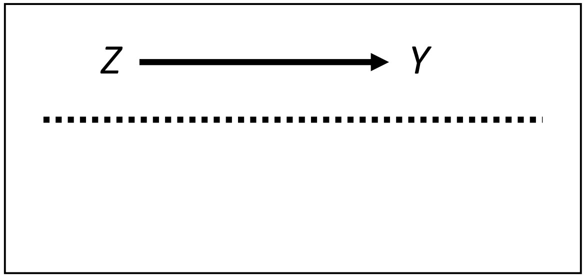


Macro-micro propositions

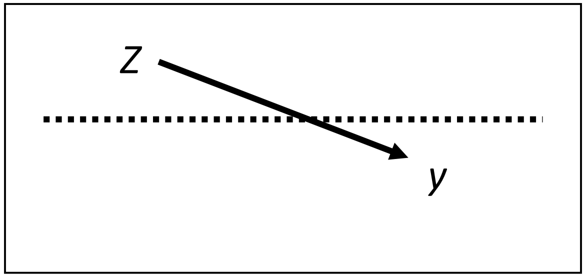
Micro-level propositions:

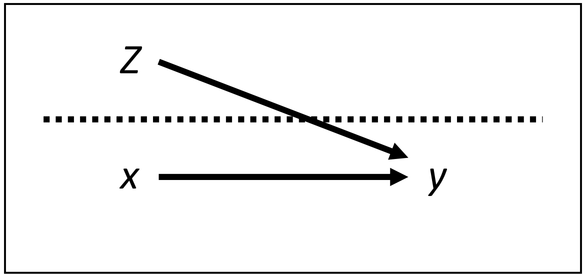


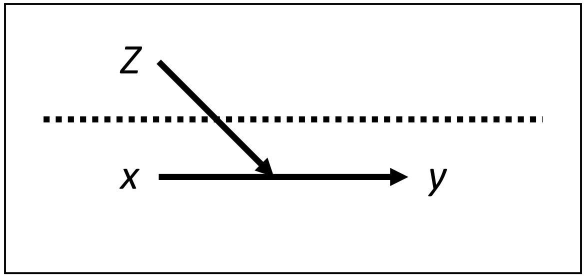
Macro-level propositions:



Macro-micro relations:

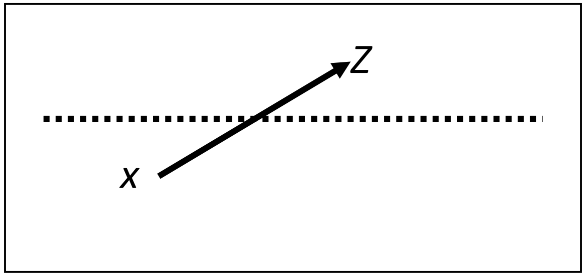




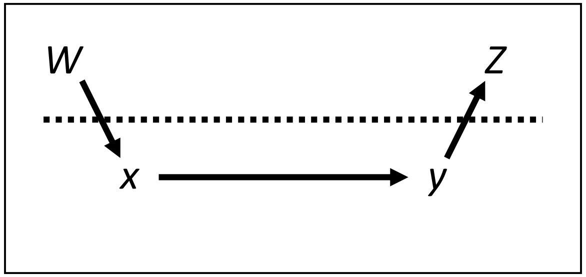


Macro-micro propositions

Micro-macro proposition:



A causal macro-micro-micro-macro chain:



(Snijders & Bosker, 2012, p.10-12)

Section 2: Lecture 9

CHANGING TWO LEVELS TO ONE

- Aggregation
- Disaggregation

Example of a 2-level structure

- Pupils within classrooms:
  - Variables: test result; classroom size (standardized)
  - Research question. Do larger classes affect test results?

Pupil data

	class	test
1	1	31
2	1	24
3	1	35
4	1	32
5	1	25
6	1	33
7	2	22
8	2	33
9	2	22
10	2	25
11	2	27
12	2	23
13	2	30
14	10	26
15	10	25
16	10	31
17	10	20

Class data

	class	size
1	1	2.33
2	2	1.67
3	10	2.00
4	12	1.33
5	15	2.00
6	16	1.67
7	18	1.67
8	21	2.00
9	24	2.67
10	26	3.00
11	27	2.33
12	29	2.00
13	33	2.33
14	35	3.00
15	36	3.33
16	38	3.00
17	40	3.00

Aggregation

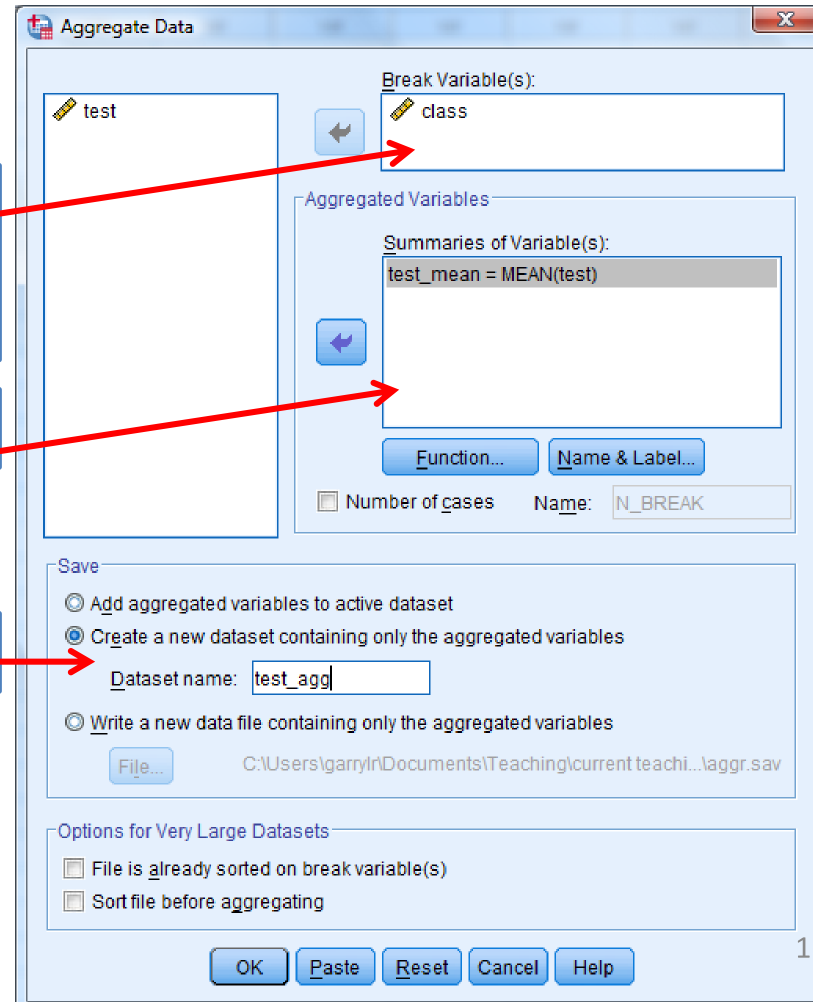
One way to model this is to calculate the means of the test for each class, and then do a regression at the classroom level

SPSS Menus: Data -> Aggregate ...

The Break variable describes how to break the dataset to create means. In this case we want means of *test* within classes

We want the mean of ‘test’ for each class.

We want to save to a new dataset.



Aggregation

One way to model this is to calculate the means of the test for each class, and then do a regression at the classroom level

	class	test_mean
1	1	29.44
2	2	26.00
3	10	26.40
4	12	26.73
5	15	30.63
6	16	37.00
7	18	33.50
8	21	34.76
9	24	36.04
10	26	35.78
11	27	30.14
12	29	37.20
13	33	30.83
14	35	31.79
15	36	33.52
16	38	36.14
17	40	35.60

Class data

Resulting file is at the class level. Need to combine with the other class level file containing the size variable

SPSS Menus: Data -> Merge Files -> Add variables...

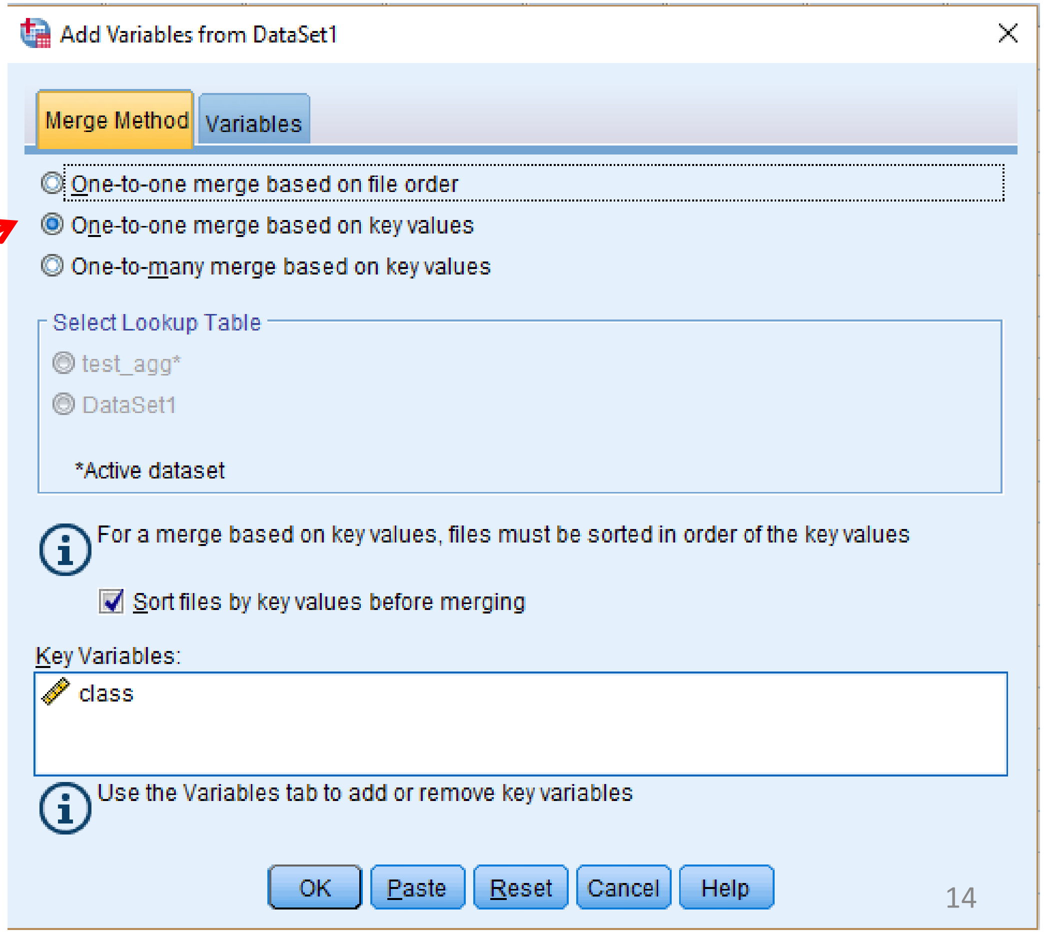
(*then select your other dataset*)

Aggregation

One way to model this is to calculate the means of the test for each class, and then do a regression at the classroom level

SPSS Menus: Data -> Merge Files -> Add variables...

We are doing a *One-to-one merge based on key values* (SPSS will default select this anyway)

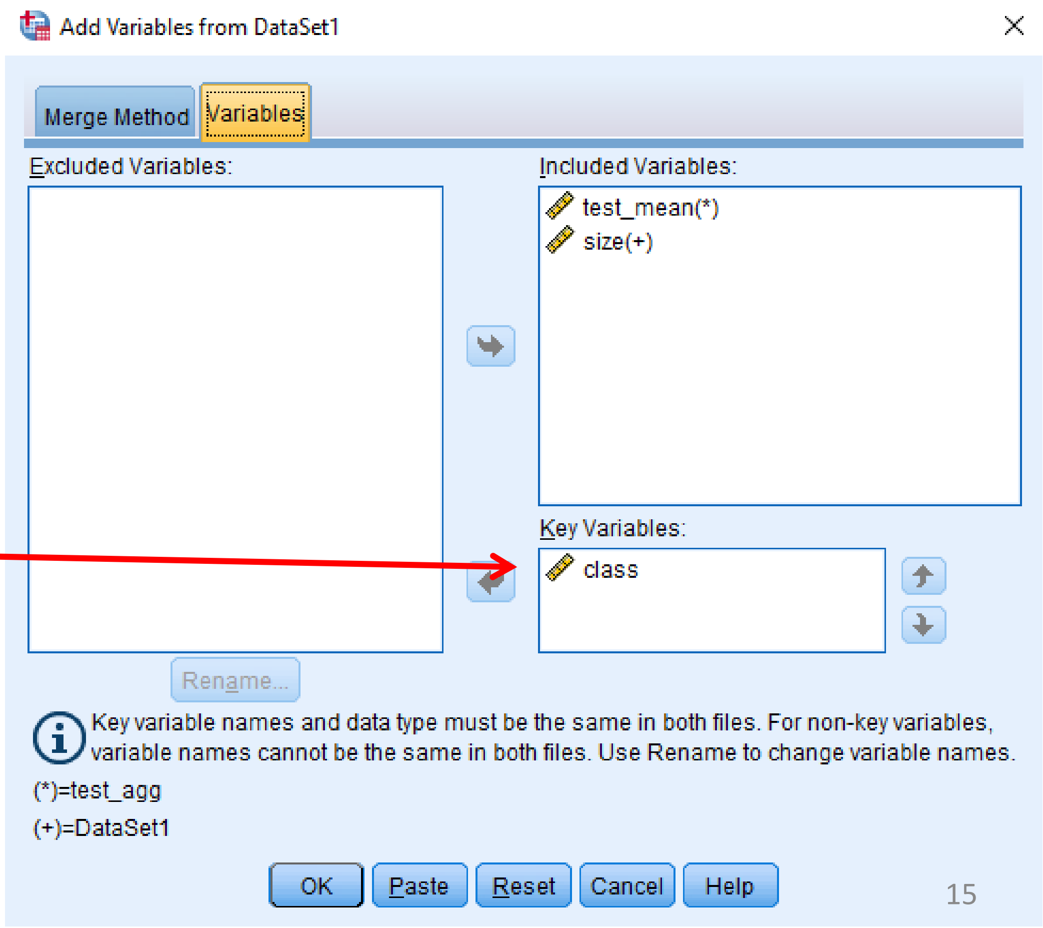


Aggregation

One way to model this is to calculate the means of the test for each class, and then do a regression at the classroom level

SPSS Menus: Data -> Merge Files -> Add variables...

Class is the variable that is identical across the two files. You need to match on this variable.



Aggregation

One way to model this is to calculate the means of the test for each class, and then do a regression at the classroom level

Class data

	class	test_mean	size
1	1	29.44	2.33
2	2	26.00	1.67
3	10	26.40	2.00
4	12	26.73	1.33
5	15	30.63	2.00
6	16	37.00	1.67
7	18	[illegible]	[illegible]
8	21		
9	24		
10	26		
11	27		
12	29		
13	33		
14	35		
15	36		
16	38	36.14	3.00
17	40	35.60	3.00
18	41	[illegible]	[illegible]

a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	32.172	1.478		21.765	.000
	size	.641	.596	.094	1.076	.284

Dependent Variable: test_mean

Aggregation

Aggregation is OK if you are only interested in macro-level propositions but there is potential for gross errors for micro-level or macro-micro propositions.

a. *shift of meaning:* We are no longer predicting student's test scores, but average test scores.

b. neglect of original data structure (between groups relations may even be opposite to within group relations)

c. prevents examination of cross-level interactions

d. *ecological fallacy*: You cannot infer that an association at a macro level translates into an association at micro level

e. danger of *aggregation bias*: (possible to get inflated statistical effects if analyses based on means are interpreted as relating to individuals – James, 1982)

Disaggregation

Another way to proceed is to apply the class size variable to each student and conduct a regression at pupil level

Because class is the variable that links both files, it is good practice to sort the cases first, so that the key variable is in the same order in both files

Data -> Sort cases ...
Sort by *class* in Ascending order
(*Do this for both files*)

Disaggregation

Another way to proceed is to apply the class size variable to each student and conduct a regression at pupil level

After sorting the cases, return to the pupil level datafile (the micro file)

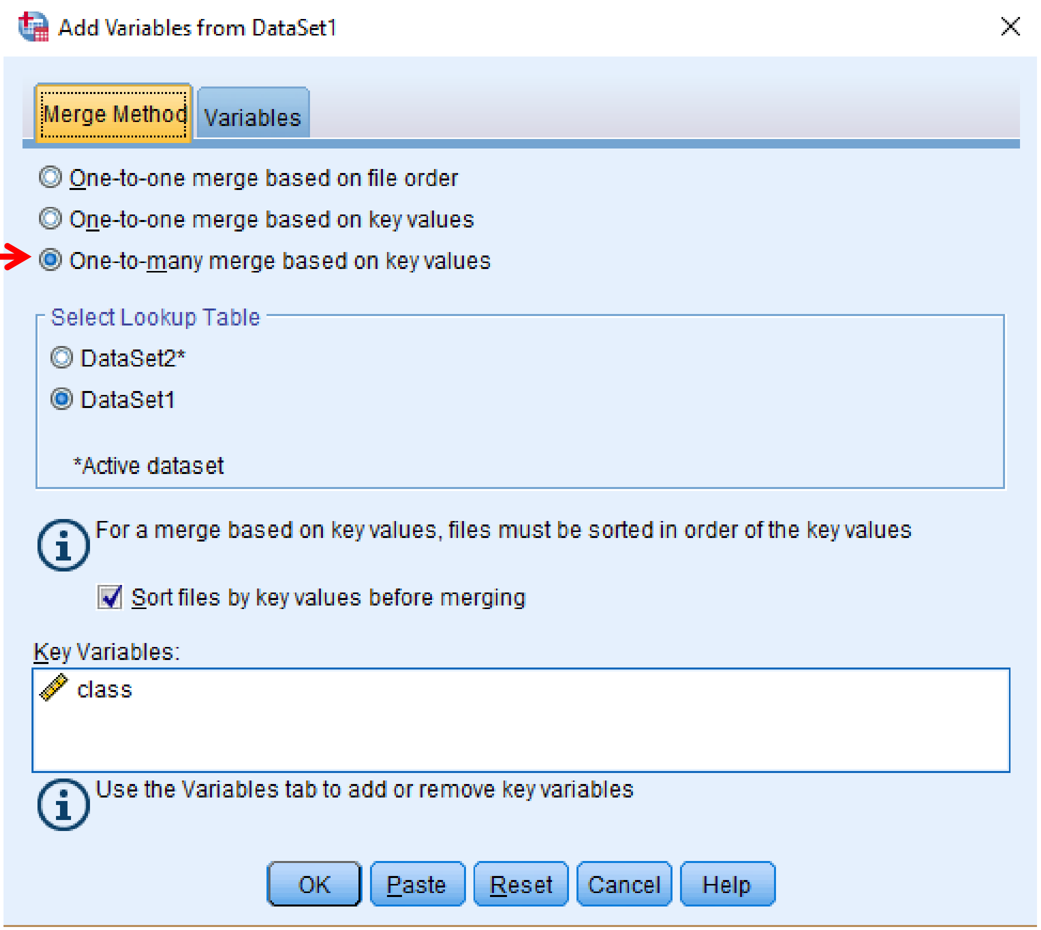
Data -> Merge files -> Add variables

Select the class level (macro-level) dataset to merge

Disaggregation

Another way to proceed is to apply the class size variable to each student and conduct a regression at pupil level

Just as before, ...except now



Disaggregation

Another way to proceed is to apply the class size variable to each student and conduct a regression at pupil level

Pupil data

	class	test	size
1	1	31	2.33
2	1	24	2.33
3	1	35	2.33
4	1	32	2.33
5	1	25	2.33
6	1	33	2.33
7	2		
8	2		
9	2		
10	2		
11	2		
12	2		
13	2		
14	10		
15	10	25	2.00
16	10	31	2.00
17	10	20	2.00

a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	32.405	.687		47.152	.000
	size	.723	.272	.055	2.655	.008

Dependent Variable: test

So, does class size affect test score? Contradictory results

Issues for Disaggregation

a. measure of macro-level variable considered as micro-level
b. *miraculous multiplication of the number of units*
c. risks of type 1 errors
d. Do not take into account that observations within a macro-unit could be correlated.

Snijders & Bosker (p.17) conclude:
"… if the macro-units have any meaningful relation with the phenomenon under study, analysing only aggregated or disaggregated data is apt to lead to misleading and erroneous conclusions. A multi-level approach, in which within-group and between-group relations are combined, is more difficult but much more productive."

Section 3: Lecture 9

RANDOM EFFECTS ANOVA

- General Linear Model: Regression and Random effects ANOVA
- Using SPSS
- Intra class correlation

General linear model: regression & ANOVA

Standard regression model: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$

where Y_i is the outcome variable for the i-th case;

X_i is the predictor variable for the i-th case;

β_0 is the constant or intercept;

β_1 is the regression coefficient for X_i;

and ε_i is the residual for the i-th case.

Assume that ε_i is normally distributed with mean 0 and variance σ^2

Multiple regression: $Y_i = \beta_0 + \sum \beta_k X_{ki} + \varepsilon_i$

where the X_{ki} are P predictor variables

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \varepsilon$$

General linear model: regression & ANOVA

Multiple regression: $Y_i = \beta_0 + \sum\beta_k X_{ki} + \varepsilon_i$

Suppose that the cases are divided into *P* groups and that the X_{ki} are dummy variables (0,1) indicating group membership:

i.e. $X_{ij} = 1$ if *i*-th case is in group *j*

$X_{ij} = 0$ otherwise.

$$Y_{ij} = \beta_0 + \sum\beta_k X_{ki} + \varepsilon_{ij}$$

One-Way ANOVA: $Y_{ij} = \beta_0 + \beta_j + \varepsilon_{ij}$

where Y_{ij} is the outcome variable for *i*-th case in *j*-th group.

Define $\beta_0 + \beta_j = \beta_{0j}$

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

Random effects ANOVA

- For fixed effects ANOVA, we assume that the groups refer to categories, each with its own distinct interpretation.
  - e.g. gender, religious denomination, treatment vs control groups.
- But sometimes the groups are *samples* from a population (actual or hypothetical) of possible macro-units.
  - e.g. three treatment groups based on different levels of drug intake
  - e.g. a study comparing five workteams within an organisation.
- In this case, β_{0j} is not a fixed, but a *random factor*.
  - Set $\beta_{0j} = \gamma_{00} + u_{0j}$ to denote this situation,
  - with γ_{00} a fixed effect (the intercept)
  - u_{0j} is assumed normally distributed with mean of 0 and variance τ^2.

Random effects ANOVA:

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

Random effects ANOVA

Random effects ANOVA:

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

- Notice that for a random effects ANOVA, there are two sources of variance.
  - u_{0j} assumed normally distributed with mean of 0 and variance τ^2.
  - ε_{ij} assumed normally distributed with mean 0 and variance σ^2
  - Total variance of $Y_{ij} = \tau^2 + \sigma^2$
  - τ^2 is the variance due to the group structure
  - σ^2 is the residual variance

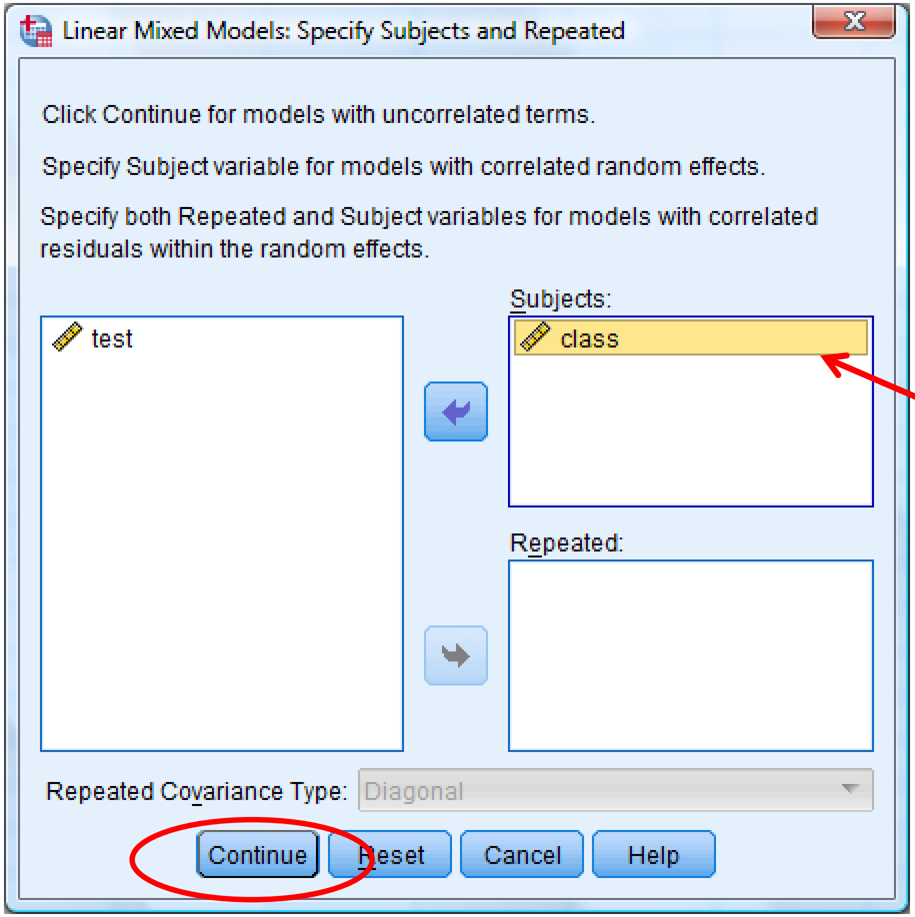
In effect, the intercept varies across all groups by an amount u_{0j} for group j. So the j-th group has intercept $\gamma_{00} + u_{0j}$. The variance of the intercept term across all groups is τ^2.

SPSS: Random effects ANOVA

Pupil data

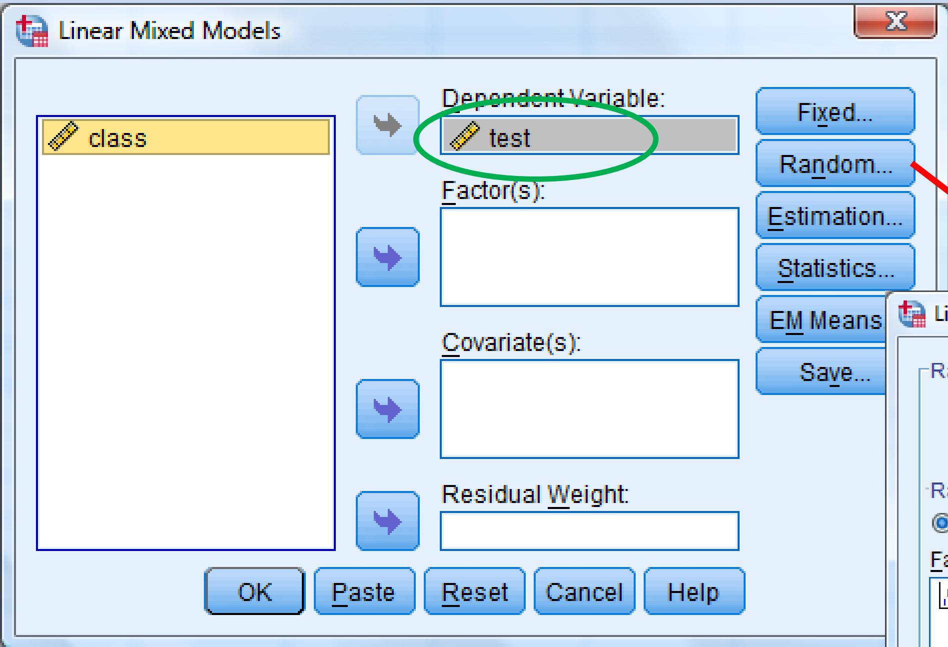
	class	test
1	1	31
2	1	24
3	1	35
4	1	32
5	1	25
6	1	33
7	2	22
8	2	33
9	2	22
10	2	25
11	2	27
12	2	23
13	2	30
14	10	26
15	10	25
16	10	31
17	10	20

Analyze -> Mixed Models -> Linear...

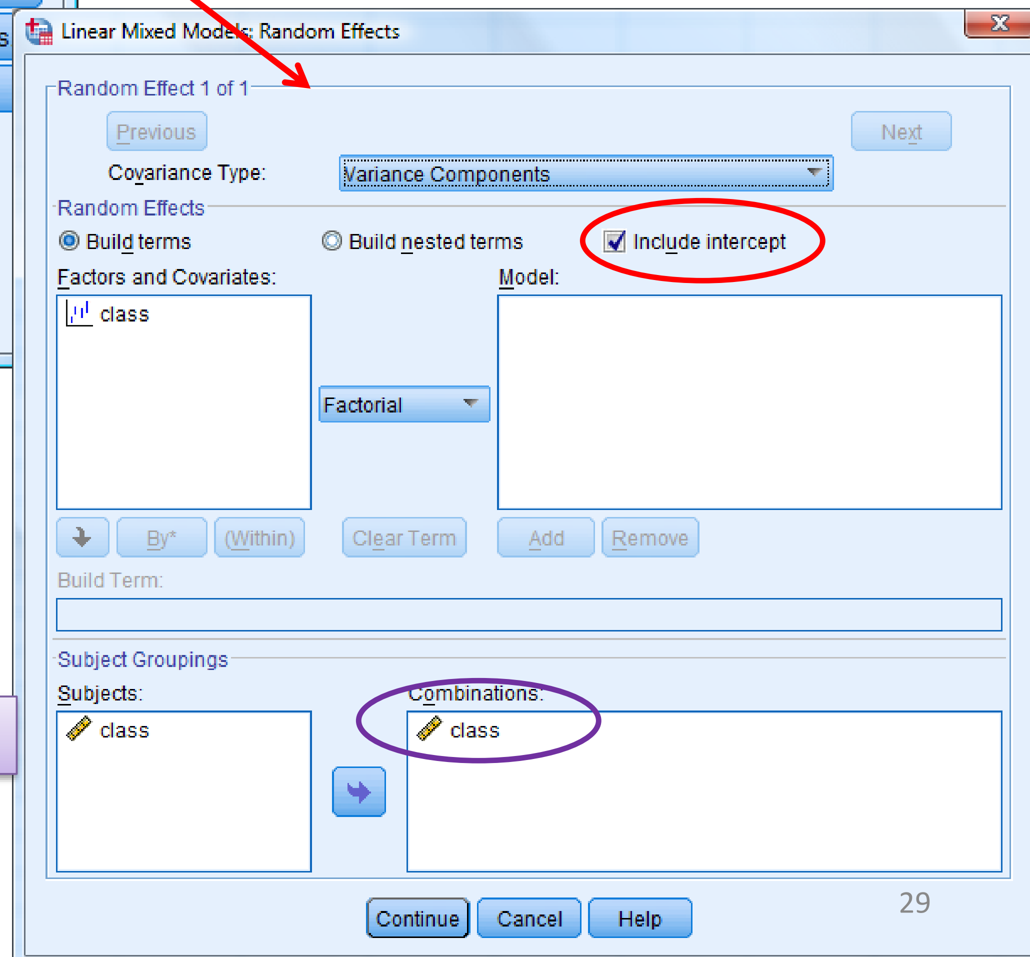


Cases grouped within class

SPSS: Random effects ANOVA



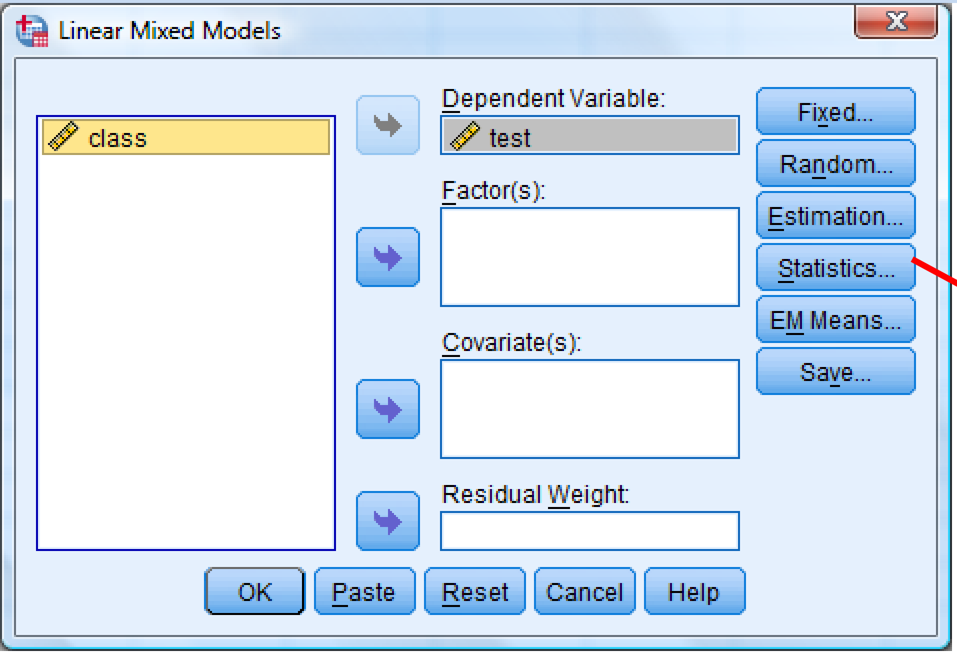
DV is the test score

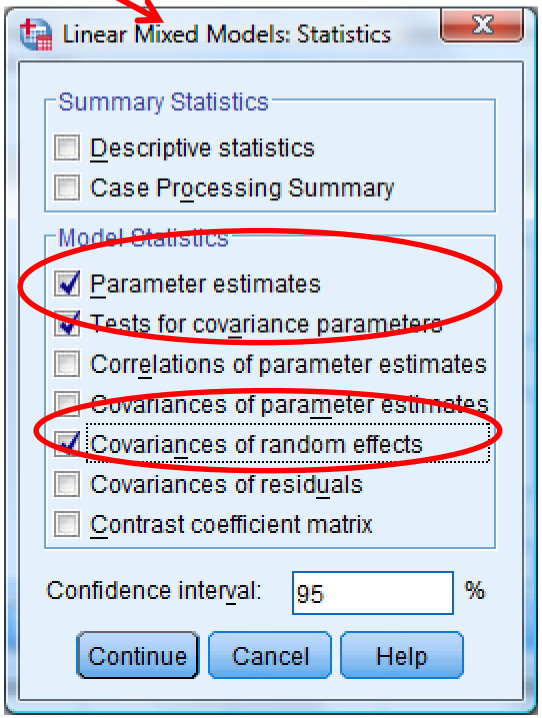


There is an intercept in our model

Class is our random subject grouping

SPSS: Random effects ANOVA





Under statistics, choose at least:
Parameter estimates
Tests for covariance parameters
Covariances of randfom effects

Output: Random effects ANOVA

Model Dimension[a]

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
Random Effects	Intercept[b]	1	Variance Components	1	class
Residual				1	
Total		2		3	

a. Dependent Variable: test.

b. As of version 11.5, the syntax rules for the RANDOM subcommand have changed. Your command syntax may yield results that differ from those produced by prior versions. If you are using version 11 syntax, please consult the current syntax reference guide for more information.

Model dimension sets out the shape of the data.

One fixed effect (the fixed intercept γ_{00})

One random effect (the variance of the intercept u_{0j})

One residual effect

Variance components simply divides the variance into these two components

Output : Random effects ANOVA

Information Criteria[a]

-2 Restricted Log Likelihood	15043.169
Akaike's Information Criterion (AIC)	15047.169
Hurvich and Tsai's Criterion (AICC)	15047.174
Bozdogan's Criterion (CAIC)	15060.638
Schwarz's Bayesian Criterion (BIC)	15058.638

The information criteria are displayed in smaller-is-better forms.

a. Dependent Variable: test.

These are various measures of fit, which are not much use unless comparing two models which we are not doing here.

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	33.920817	.266744	107.560	127.166	.000	33.392059	34.449575

a. Dependent Variable: test.

The fixed effect γ_{00}. The estimate is 33.9. It is significantly different from 0.

Intra class correlation

- The ICC is the proportion of variance explained by the group structure.
  - It is also the correlation between two randomly drawn individuals in one randomly drawn group.
  - There are several ICCs – the one we are dealing with today is also often called ICC(1) – McGraw & Wong, 1996.

Random effects ANOVA:

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

Variance of intercept = var ($\gamma_{00} + u_{0j}$) = var (u_{0j}) = variance due to group structure

Variance due to residual = var (ε_{ij})

Total variance = var (u_{0j}) + var (ε_{ij})

ICC = var (u_{0j}) / total variance = var (u_{0j}) / [var (u_{0j}) + var (ε_{ij})]

The proportion of variance explained by the group level

Output : Random effects ANOVA

Estimates of Covariance Parameters[a]

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		38.986611	1.193325	32.671	.000	36.716512	41.397064
Intercept [subject = class]	Variance	6.735678	1.253613	5.373	.000	4.676916	9.700698

a. Dependent Variable: test.

Random effects ANOVA:

Variance of intercept = var (u_{0j}) = 6.74

Variance due to residual = 38.99

Total variance = 6.74 + 38.99 = 45.73

ICC = var (u_{0j}) / total variance = 6.74 / 45.73 = 0.147

So 14.7% of variance in test score is due to the group structure

The intercept variance is significant.

The ICC at 14.7% is greater than 5% (used as a rough cut-off value).

So, the group structure is important to explaining test scores – this argues in favour of doing multilevel modelling.

Section 4: Lecture 9

RANDOM INTERCEPT MULTILEVEL MODEL

- ❑ Hierarchical Linear Models
- ❑ Random intercept: The null model
- ❑ Random intercept with one predictor

Hierarchical Linear Models
Multilevel models

A two stage strategy to investigate variables at two levels of analysis.

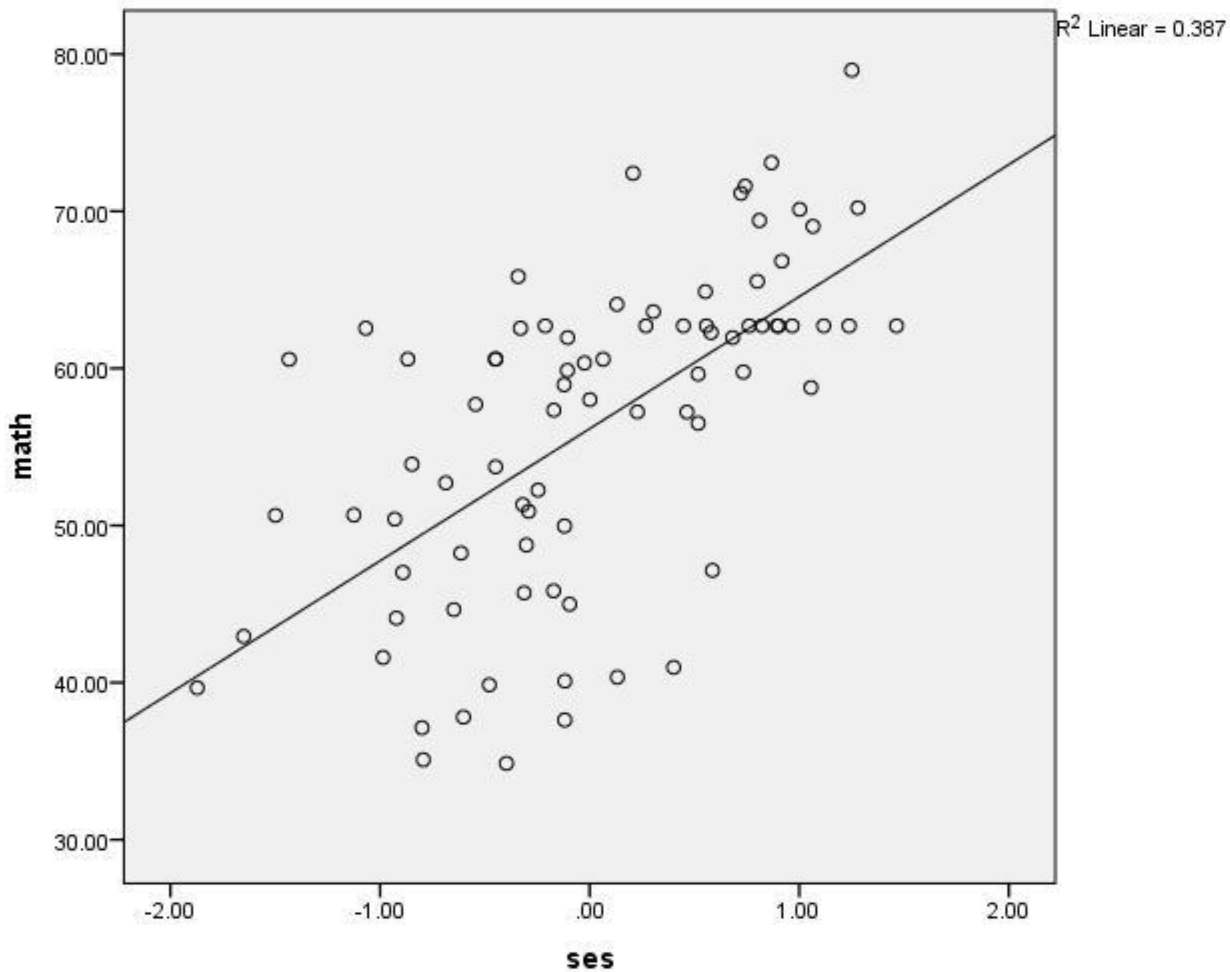
i. Level 1: relationships among level 1 variables estimated separately for each higher level (level 2) unit.

ii. These relationships are then used as outcome variables for the variables at level 2.

Example
(Heck et al, 2014)

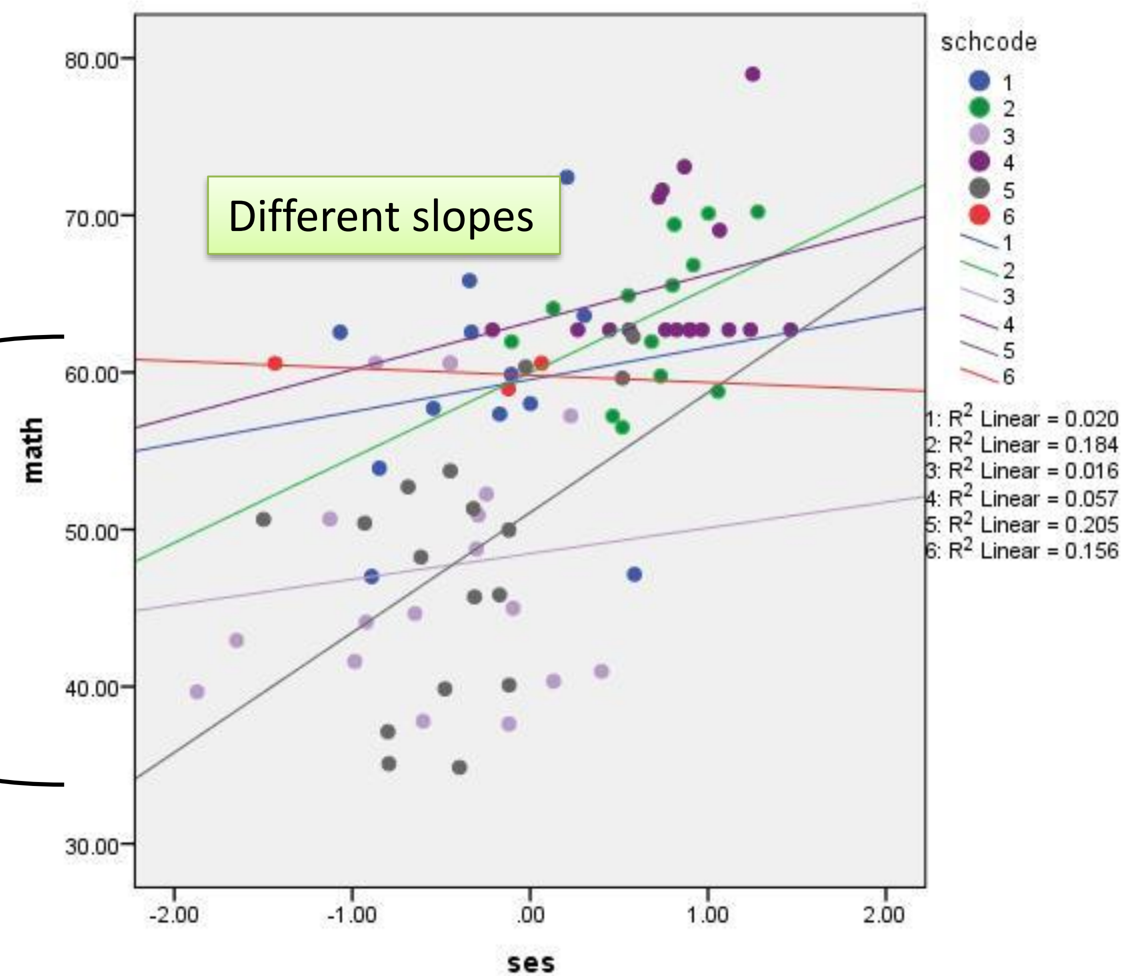
- Students nested within schools
- Dataset contains a number of variables including:
  - Maths test score
  - Student socio-economic status (a standardised continuous measure).
- Is SES associated with performance on the maths test?

One regression model



Assumes the one model for all schools

Regression models for each class



Different intercepts: Multilevel models assume that these intercepts are normally distributed around a mean value

Different linear models for 6 different schools

Random intercept model with one Level 1 predictor

Random intercept model treats the intercepts as random but has a fixed effect for slope.

Level 1 equation: $Y_{ij} = \beta_{0j} + \beta_1 (\mathrm{SES})_{ij} + \varepsilon_{ij}$

Level 2 equations: $\beta_{0j} = \gamma_{00} + u_{0j}$ ← Random intercept

$\beta_1 = \gamma_{10}$ ← Fixed slope

So that $Y_{ij} = \gamma_{00} + \gamma_{10} (\mathrm{SES})_{ij} + u_{0j} + \varepsilon_{ij}$

Notice that if we ignore the predictor SES (ie set $\gamma_{10} = 0$), then we have the random effects ANOVA – the null model

Variance components for null model

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	57.674234	.188266	416.066	306.344	.000	57.304162	58.044306

a. Dependent Variable: math.

Estimates of Covariance Parameters[a]

Parameter		Estimate	Std. Error
Residual		66.550655	1.171618
Intercept [subject = schcode]	Variance	10.642209	1.028666

a. Dependent Variable: math.

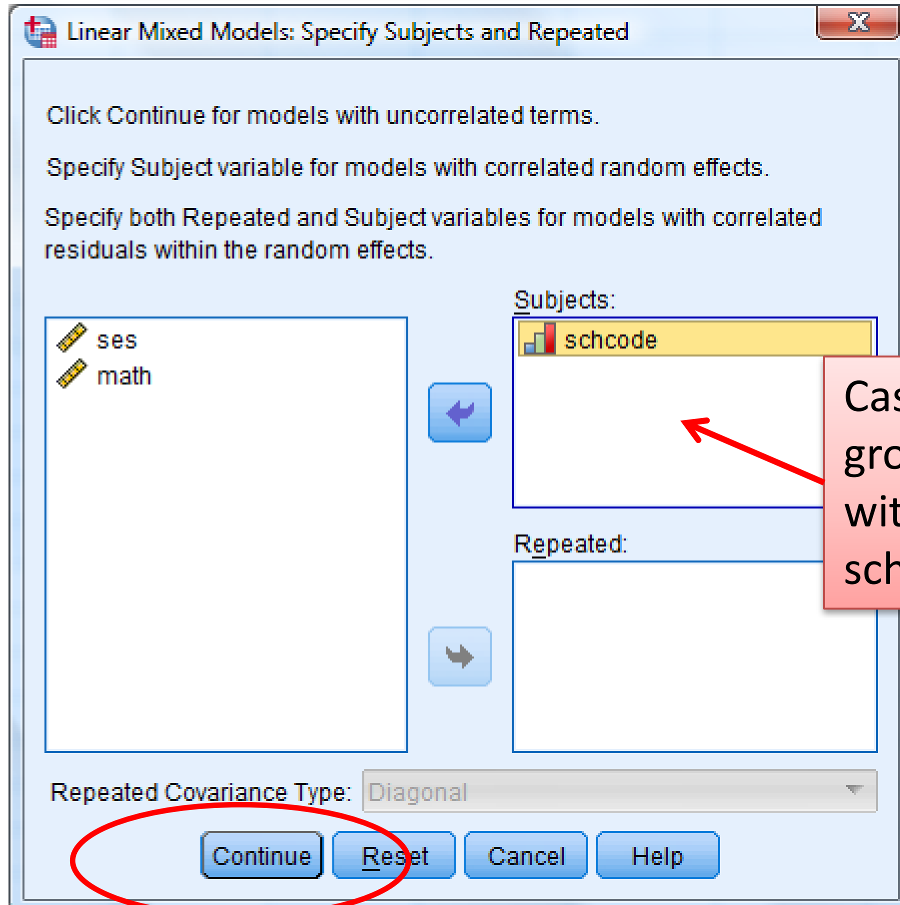
ICC = 0.138

Random intercept model
Predicting *math* from *ses*

Analyze -> Mixed Models -> Linear...

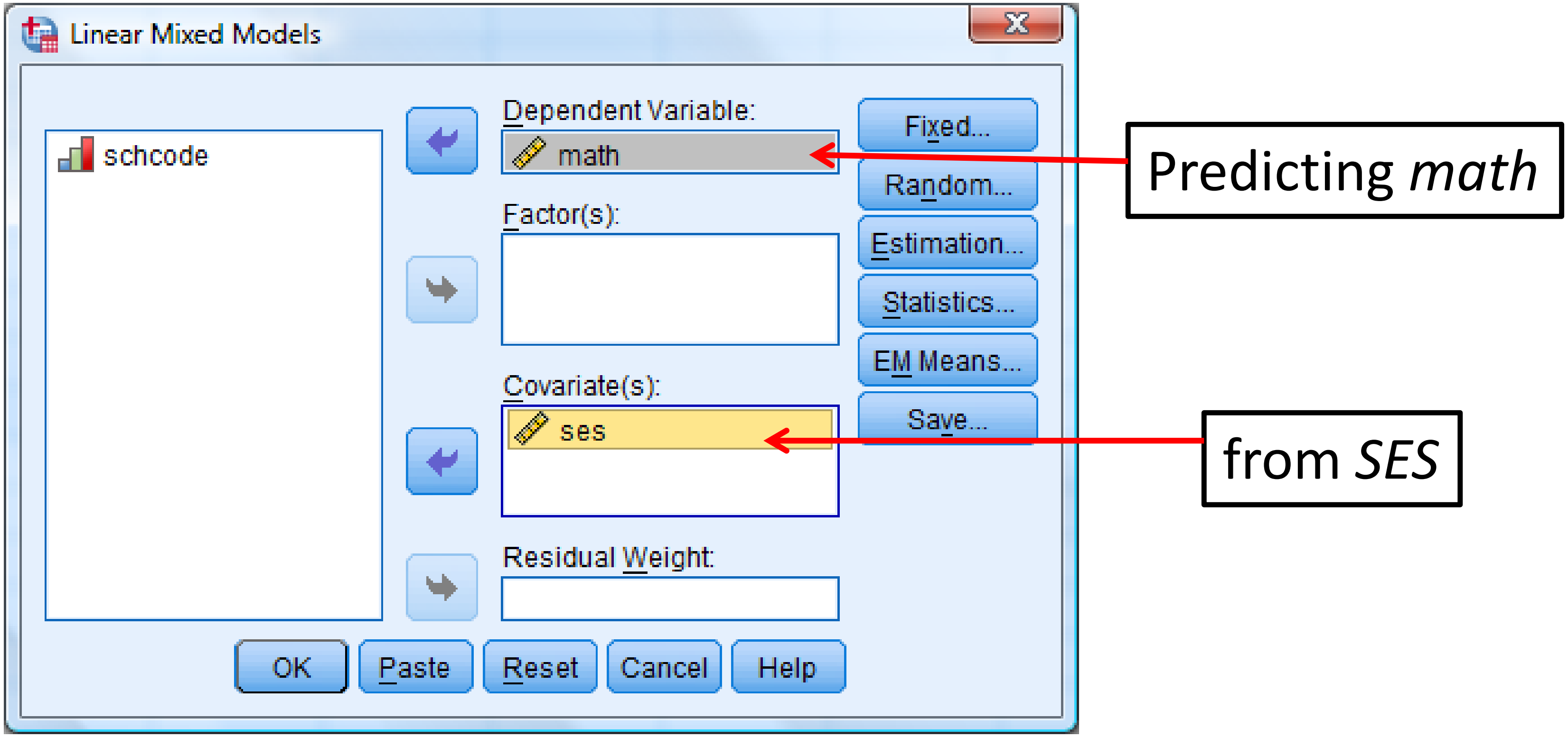
	schcode	ses	math
1	1	.59	47.14
2	1	.30	63.61
3	1	-.54	57.71
4	1	-.85	53.90
5	1	.00	58.01
6	1	-.11	59.87
7	1	-.33	62.56
8	1	-.89	47.01
9	1	.21	72.42
10	1	-.34	65.84
11	1	-.17	57.34
12	1	-1.07	62.56
13	2	-.11	61.95
14	2	1.28	70.22
15	2	1.06	58.78
16	2	.80	65.54
17	2	.73	59.77

schcode is the code number for each school

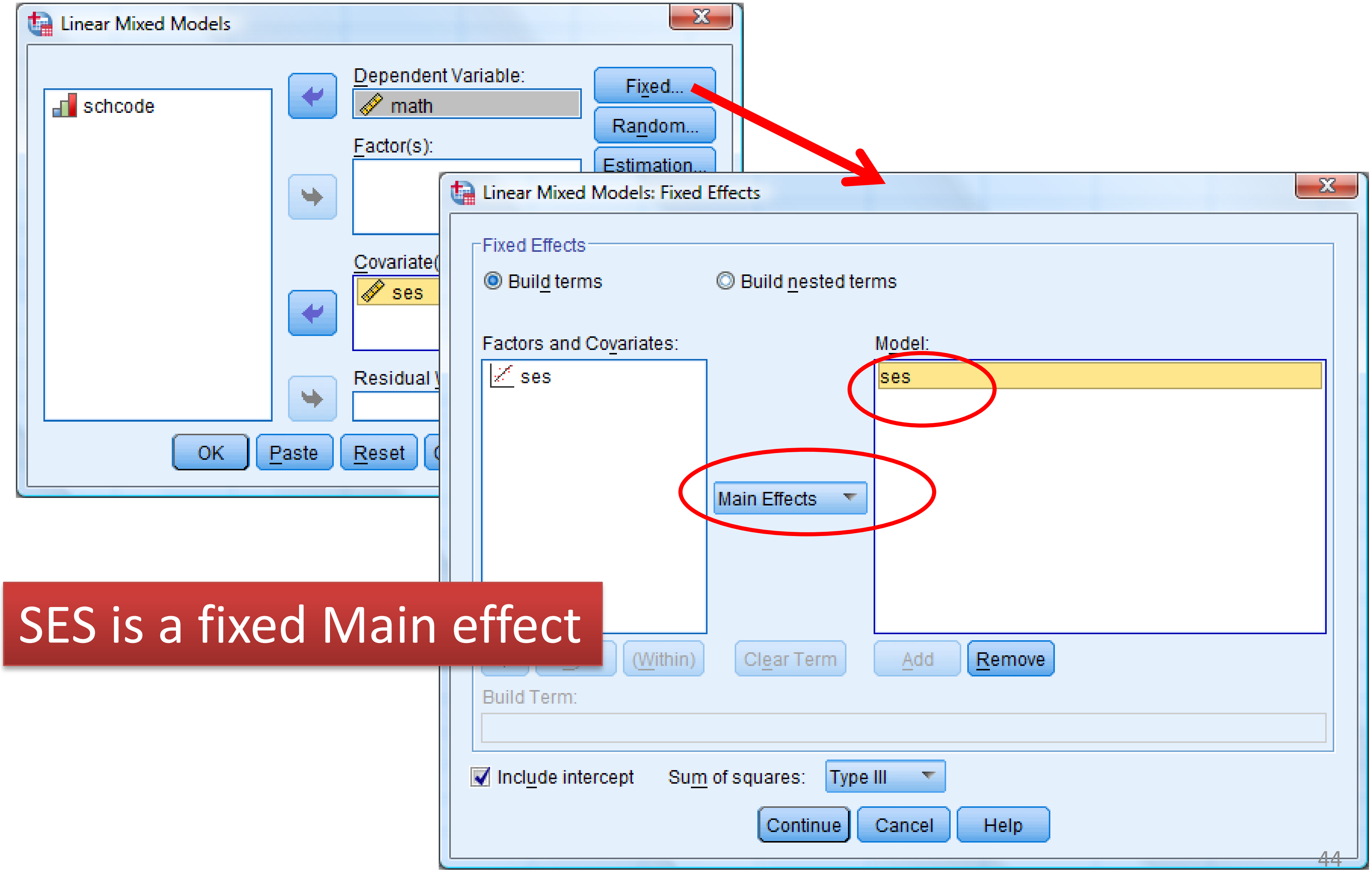


Cases grouped within school

Random intercept model
Predicting *math* from *ses*

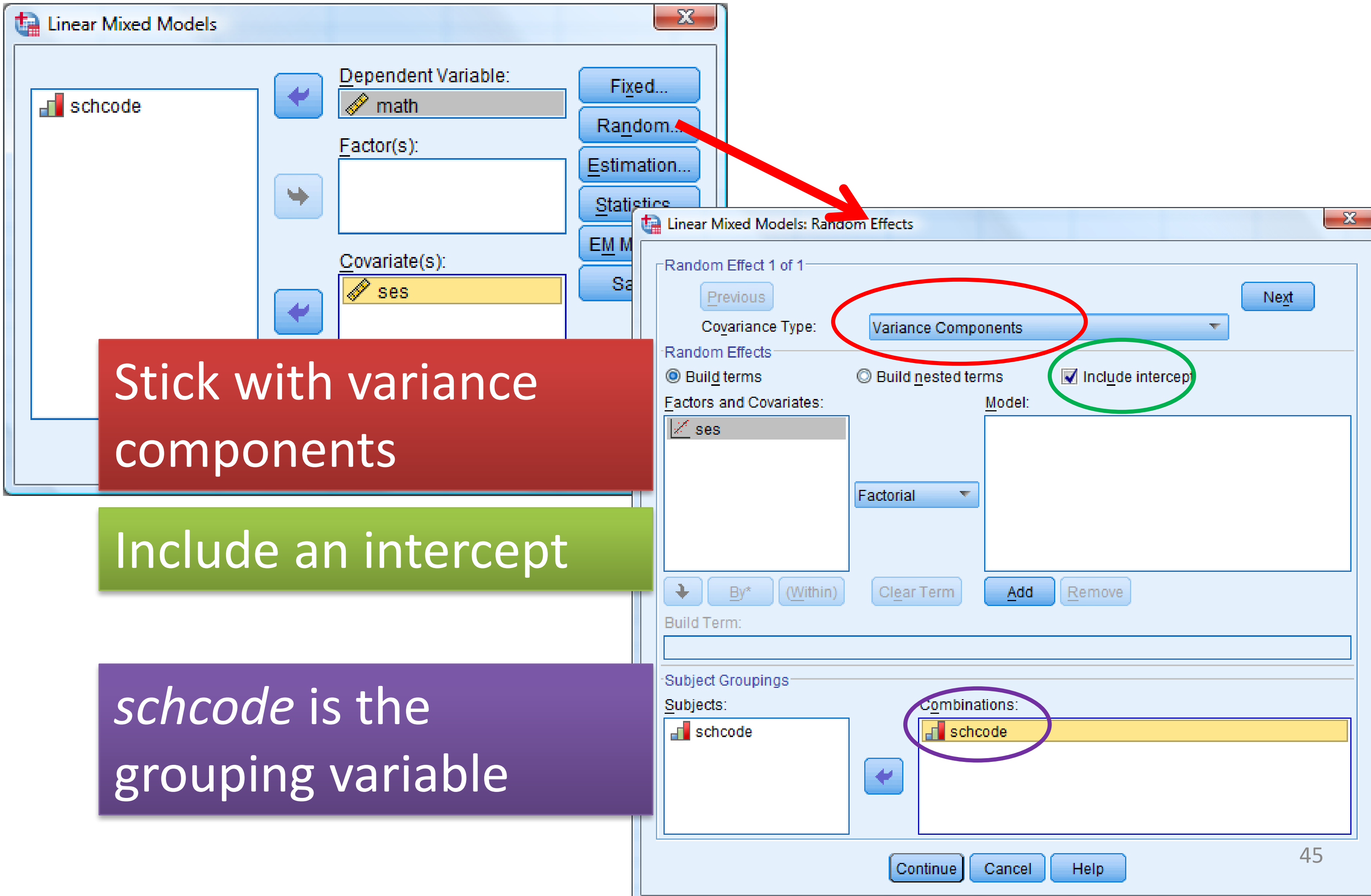


Random intercept model
Predicting *math* from *ses*



SES is a fixed Main effect

Random intercept model
Predicting *math* from *ses*

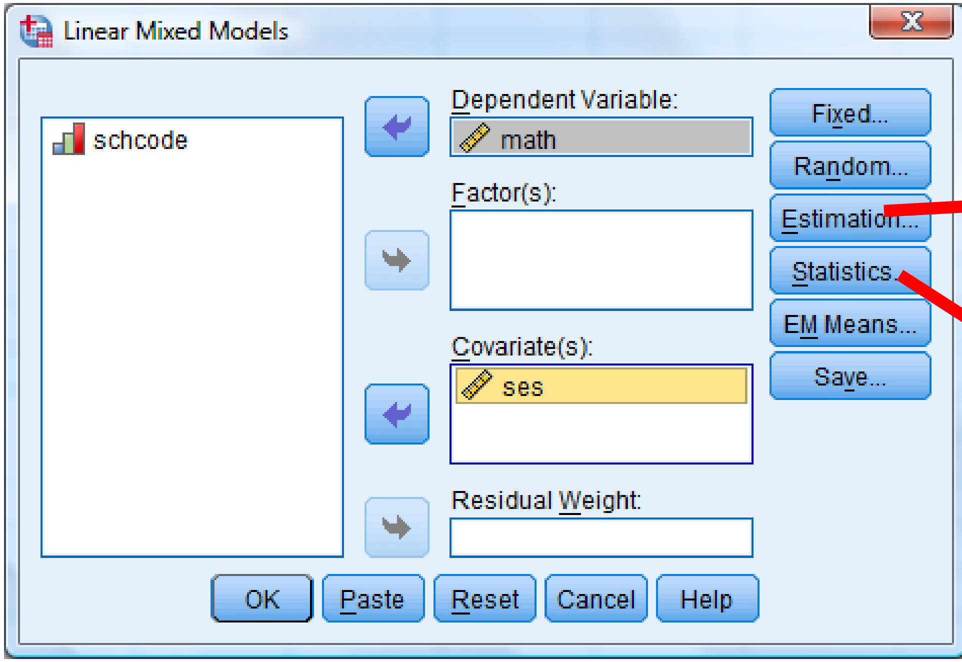


Stick with variance components

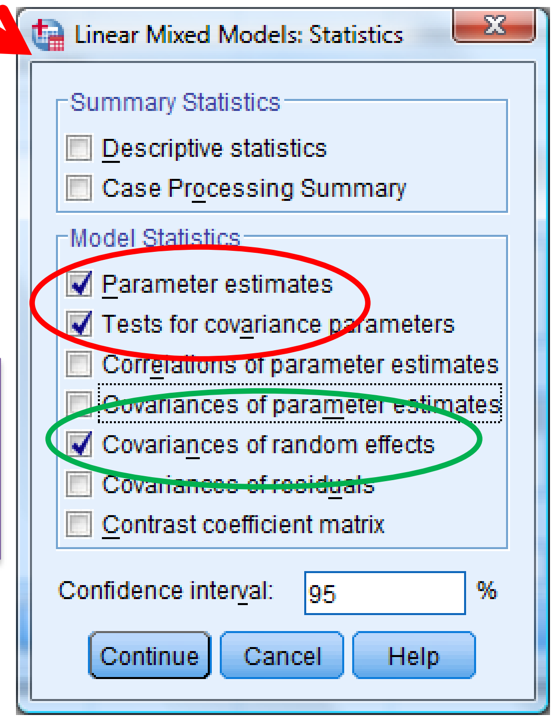
Include an intercept

schcode is the grouping variable

Random intercept model
Predicting *math* from *ses*



Check that restricted maximum likelihood is used



Select at least these statistics

Output

Model Dimension[a]

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	ses	1		1	
Random Effects	Intercept[b]	1	Variance Components	1	schcode
Residual				1	
Total		3		4	

a. Dependent Variable: math.

b. As of version 11.5, the syntax rules for the RANDOM subcommand have changed. Your command syntax may yield results that differ from those produced by prior versions. If you are using version 11 syntax, please consult the current syntax reference guide for more information.

Model dimension sets out the shape of the data.

- Two fixed effects (the fixed intercept γ_{00} and fixed slope for *ses*)
- One random effect (the variance of the intercept u_{0j})
- One residual effect
  - Four parameters: $Y_{ij} = \gamma_{00} + \gamma_{10}\,(SES)_{ij} + u_{0j} + \varepsilon_{ij}$

Output

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	57.595965	.132905	375.699	433.362	.000	57.334634	57.857296
ses	3.873861	.136624	3914.638	28.354	.000	3.605999	4.141722

a. Dependent Variable: math.

Estimates for fixed effects:

Estimate for fixed intercept γ_{00} = 57.60

Estimate for slope of ses, γ_{10} = 3.87

Both are significant.

Significant positive slope for *ses* predicting *math*

Four parameters: $Y_{ij} = \gamma_{00} + \gamma_{10}\,(\mathrm{SES})_{ij} + u_{0j} + \varepsilon_{ij}$

Output

Estimates of Covariance Parameters[a]

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		62.807187	1.108877	56.640	.000	60.671000	65.018587
Intercept [subject = schcode]	Variance	3.469256	.538821	6.439	.000	2.558783	4.703696

a. Dependent Variable: math.

Variance:

Previous intercept variance for null model = 10.64

Intercept variance for current model = 3.47

Reduction of variance from null model = (10.64-3.47)/10.64

= 67%

So SES accounts for 67% of between group variance

ICC = 3.47/(3.47+62.81) = 0.05 – much reduced compared to null model.

Summary

IN THIS LECTURE, you

- were introduced to the idea of multilevel research
- learnt how aggregating or disaggregating data across levels is not the best way to deal with multilevel data
- learnt about random effects ANOVA
- Were introduced to a random intercept model for grouped data
- learnt how to fit these models in SPSS

References

Books:

Heck, Thomas & Tabata (2014). *Multilevel and longitudinal modeling with IBM SPSS.* Routledge. (chaps 2,3).

Snijders, T., & Bosker, R. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* London: Sage.

Articles

James, L.R. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology, 67,* 219-229

McGraw, K.O., & Wong, S.P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1,* 30-46.

Organizational climate

Glick, W.H. (1985). Conceptualizing and measuring organizational and psychological climate: pitfalls in multilevel research. *Academy of Management Review, 10*, 601-616.

Gonzalez-Roma, V., et al. (1999). The validity of collective climates. *Journal of Occupational & Organizational Psychology, 72*, 25-40.

Jackofsky, E.F. & Slocum, J.W. (1990). Rejoinder to Payne's comment on ‘A longitudinal study of climates’. *Journal of Organizational Behavior, 11*, 81-83.

James, L.R., & Jones, A.P. (1974). Organizational climate: A review of theory and research. *Psychological Bulletin, 81,* 1096-1112.

Payne, R. (1990). Madness in our method. A comment on Jackofsky and Slocum’s paper, ‘A longitudinal study of climates’. *Journal of Organizational Behavior, 11*, 77-80.

Young. S.A. & Parker, C.P. (1999). Predicting collective climates: assessing the role of shared work values, needs, employee interaction and work group membership. *Journal of Organizational Behavior, 20,* 1199-1218.