PSYC40005 - 2018
ADVANCED DESIGN AND DATA ANALYSIS

Lecture 10:
Multilevel modelling 2

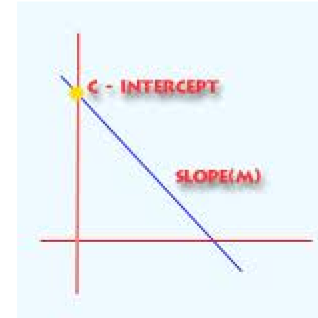
Geoff Saw

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

The agenda for this lecture

1. Random intercept multilevel model
2. Random slope models
3. Some issues for MLMs

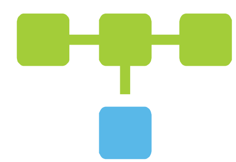
GOALS OF THIS LECTURE
- To review and extend random intercept models
- To introduce random slope models
- To present some additional issues

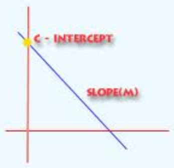


Section 1: Lecture 10

MORE ON RANDOM INTERCEPTS

- ❑ A review of last lecture
- ❑ Level 2 predictors
- ❑ SPSS output

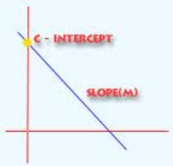




A two stage strategy to investigate variables at two levels of analysis.

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.

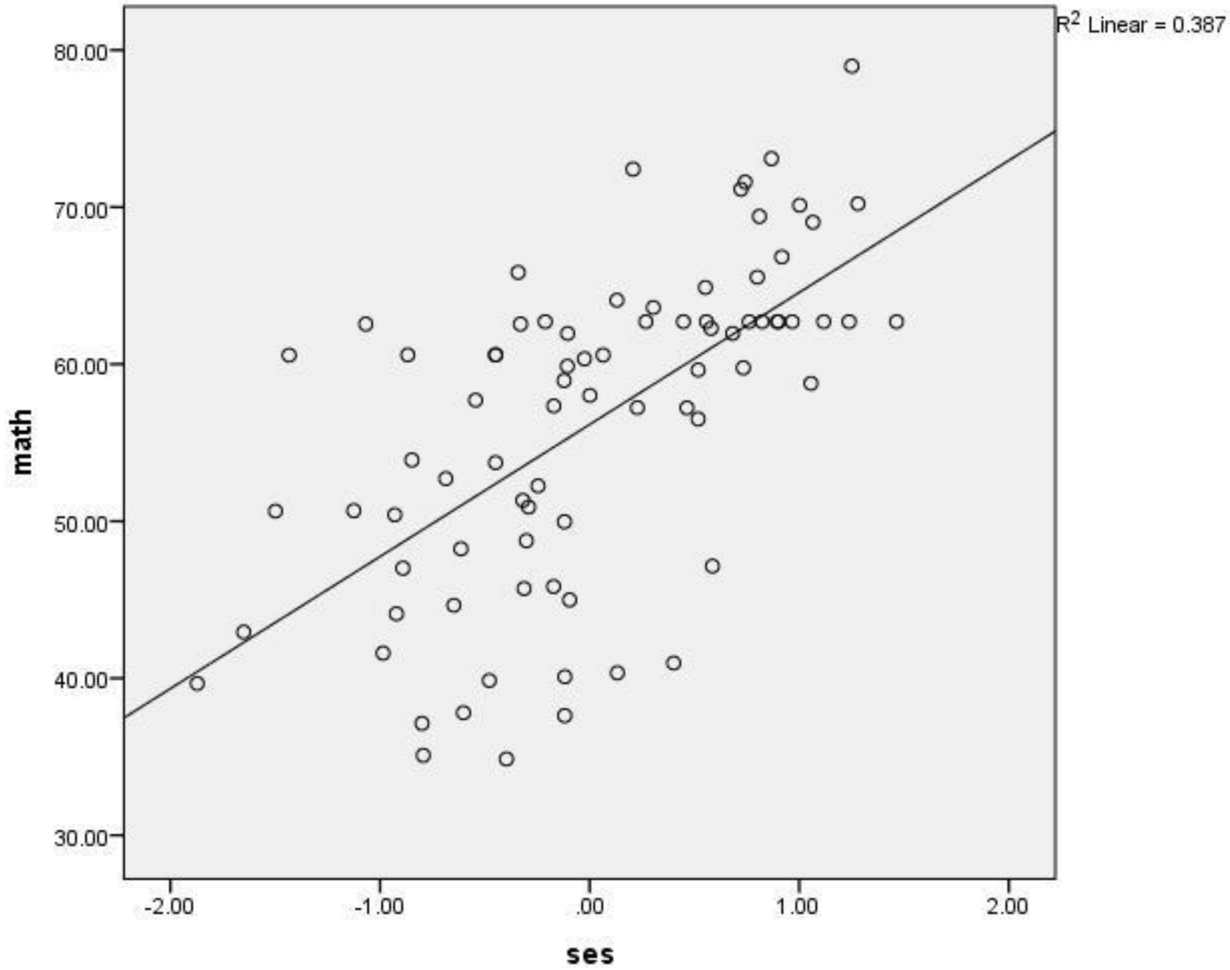


Example

(Heck et al, 2011 – chap 3)

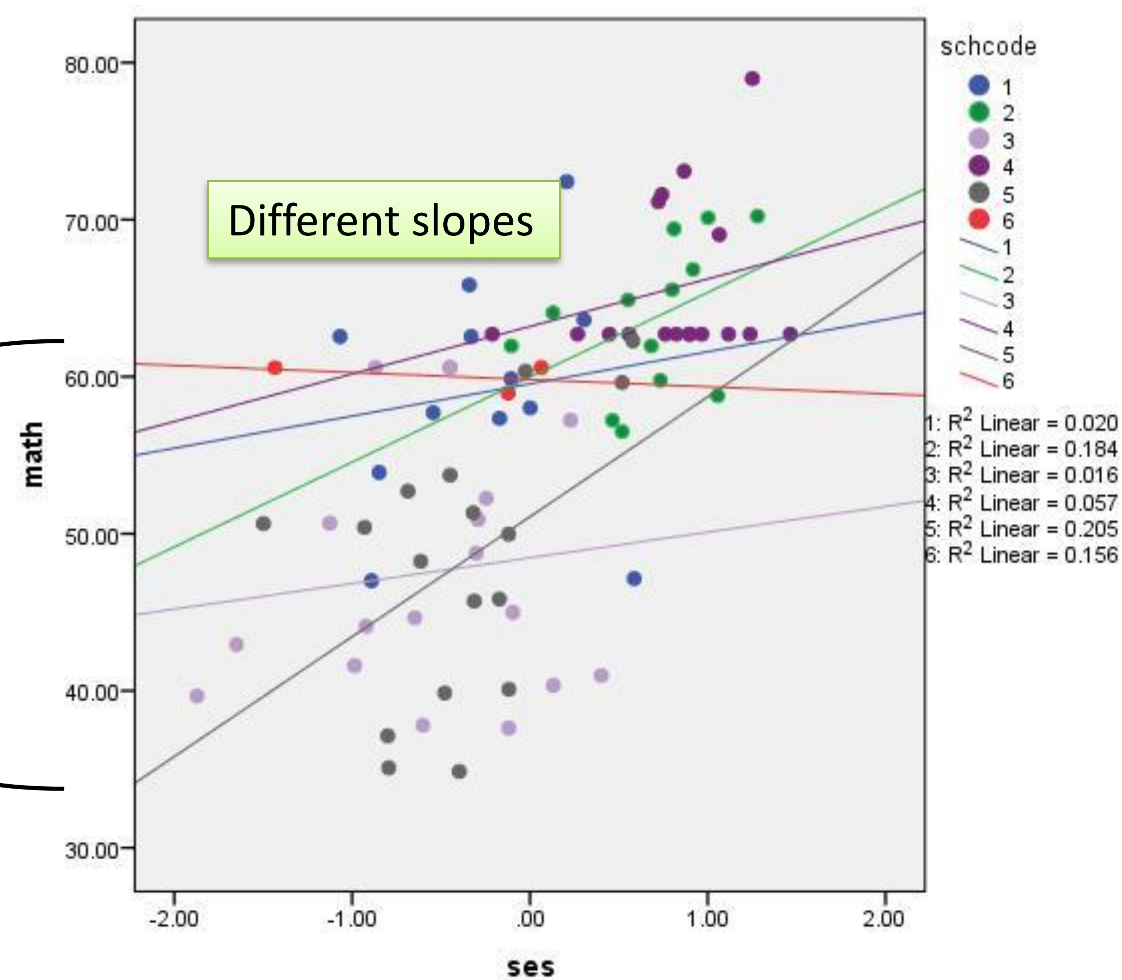
- Students nested within schools
- Dataset contains a number of variables including:
  - Maths test score (*math*)
  - Student socio-economic status (*ses*).
  - Average SES per school (*ses_mean*)
  - Percent of students who intend to study at "4-year universities" (*per4yrc*)
  - Type of school (*public* = 1 if public, 0 otherwise)
- The goal is predict performance on the maths test

One regression model



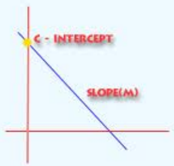
Assumes the one model for all schools

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



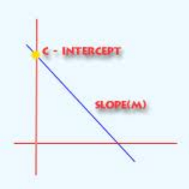
Random intercept model
Notation

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

j is an index for groups
i is an index for individuals within groups

Y_{ij} is the dependent or outcome variable

x_{ij} is a predictor variable at individual level
z_j is a predictor variable at group level



Random intercept model Equations

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

Level 1 equation: $Y_{ij} = \beta_{0j} + \beta_1 x_{ij} + \varepsilon_{ij}$

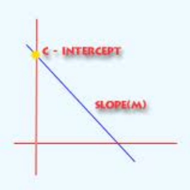
Level 2 equations: $\beta_{0j} = \gamma_{00} + \gamma_{01} z_j + u_{0j}$ — Random intercept

$\beta_1 = \gamma_{10}$ — Fixed slope

So that $Y_{ij} = \gamma_{00} + \gamma_{10} x_{ij} + \gamma_{01} z_j + u_{0j} + \varepsilon_{ij}$

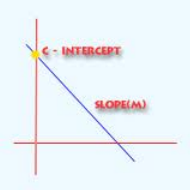
If we have no predictors *x* and *z* (ie set $\gamma_{10} = \gamma_{01} = 0$), then we have the random effects ANOVA – the null model

If we have no predictors *z* (ie set $\gamma_{01} = 0$), then we have a random intercept model with only level 1 predictors – last lecture



Last lecture

- Predict performance on the maths test from student SES
- First we fitted a null model to check the variance components (i.e. calculate an ICC) to see if we needed a multilevel model.
- Then we fitted a random intercept MLM with *ses* as the only predictor



Output 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]

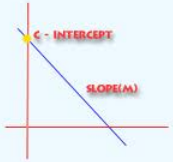
Parameter		Estimate	Std. Error
Residual		66.550655	1.171618
Intercept [subject = schcode]	Variance	10.642209	1.028666

a. Dependent Variable: math.

Random effects ANOVA:

$$Y_{ij} = \gamma_{00} + u_j + \varepsilon_{ij}$$

ICC = [10.64/(10.64+66.55)] = 0.138



Output for random intercept model

Model Dimension[a]

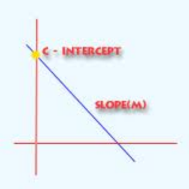
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	ses	1		1	
Random Effects	Intercept[b]	1	Variance Components	1	schcode
Residual				1	
Total		3		4	

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_j)
- One residual effect
  - Four parameters: $Y_{ij} = \gamma_{00} + \gamma_{10}\,(\text{SES})_{ij} + u_j + \varepsilon_{ij}$



Output for random intercept model

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	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:

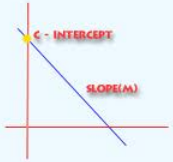
Estimate for fixed intercept γ_{00} = 57.60

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

Both are significant.

Significant positive coefficient for *ses* predicting *math*

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



Output for random intercept model

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

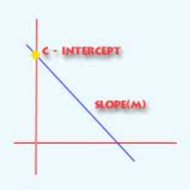
Intercept variance for current model = 3.47

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

= 67%

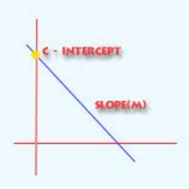
So SES accounts for 67% of between group variance

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



Today

- We want to predict performance on the maths test from student SES as well as from school level variables:
  - Average SES per school (*ses_mean*)
  - Percent of students who intend to study at "4-year universities" (*per4yrc*)
  - Type of school (*public* = 1 if public, 0 otherwise)
- We have already fitted a null model so we know we need a multilevel model.
- Now we fit a random intercept MLM with both individual (student) and group (school) level predictors.



Random intercept model

Level 1 equation: $Y_{ij} = \beta_{0j} + \beta_1 x_{ij} + \varepsilon_{ij}$

Level 2 equations: $\beta_{0j} = \gamma_{00} + \gamma_{01} z_j + u_{0j}$

$\beta_1 = \gamma_{10}$

So that $Y_{ij} = \gamma_{00} + \gamma_{10} x_{ij} + \gamma_{01} z_j + u_{0j} + \varepsilon_{ij}$

Level 1 equation:

$math_{ij} = \beta_{0j} + \beta_1 ses_{ij} + \varepsilon_{ij}$

Level 2 equations:

$\beta_{0j} = \gamma_{00} + \gamma_{01} ses\_mean_j + \gamma_{02} per4yrc_j + \gamma_{03} public_j + u_{0j}$

$\beta_1 = \gamma_{10}$

So that

$math_{ij} = \gamma_{00} + \gamma_{10} ses_{ij} + \gamma_{01} ses\_mean_j + \gamma_{02} per4yrc_j + \gamma_{03} public_j + u_{0j} + \varepsilon_{ij}$

	schcode	ses	math	ses_mean	per4yrc	public
1	1	.59	47.14	-.27	.08	0
2	1	.30	63.61	-.27	.08	0
3	1	-.54	57.71	-.27	.08	0
4	1	-.85	53.90	-.27	.08	0
5	1	.00	58.01	-.27	.08	0
6	1	-.11	59.87	-.27	.08	0
7	1	-.33	62.56	-.27	.08	0
8	1	-.89	47.01	-.27	.08	0
9	1	.21	72.42	-.27	.08	0
10	1	-.34	65.84	-.27	.08	0
11	1	-.17	57.34	-.27	.08	0
12	1	-1.07	62.56	-.27	.08	0
13	2	-.11	61.95	.68	1.00	0
14	2	1.28	70.22	.68	1.00	0
15	2	1.06	58.78	.68	1.00	0
16	2	.80	65.54	.68	1.00	0
17	2	.73	59.77	.68	1.00	0
18	2	.13	64.07	.68	1.00	0
19	2	.68	61.95	.68	1.00	0
20	2	.92	66.83	.68	1.00	0
21	2	.81	69.41	.68	1.00	0
22	2	.52	56.51	.68	1.00	0
23	2	.46	57.22	.68	1.00	0
24	2	.55	64.89	.68	1.00	0
25	2	1.00	70.12	.68	1.00	0
26	3	-.30	48.76	-.55	.33	0
27	3	-.65	44.65	-.55	.33	0
28	3	-.45	60.59	-.55	.33	0
29	3	-.10	44.99	-.55	.33	0
30	3	-.29	50.90	-.55	.33	0

This is a merged datafile from various sources.

A student file with ses and math variables.

A school data file with per4yc and public variables.

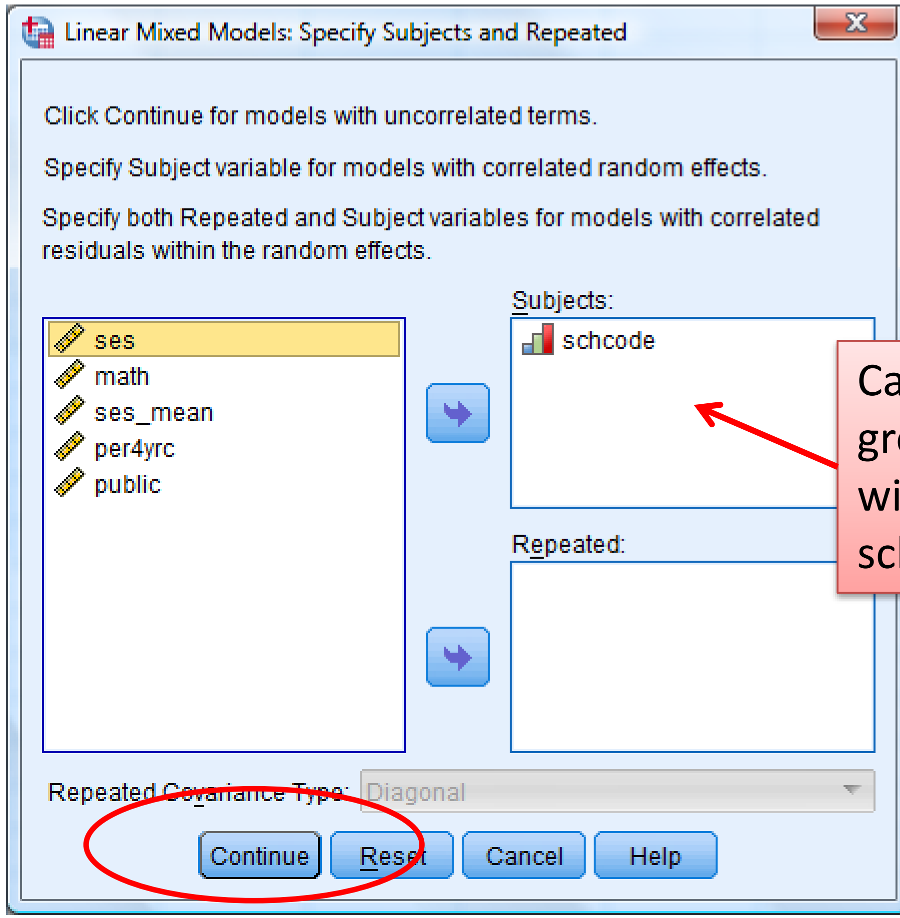
ses_mean was produced by aggregation and added to the school datafile

The school level variables were disaggregated and merged into the student level file.

We saw how to do aggregation and disaggregation last lecture.

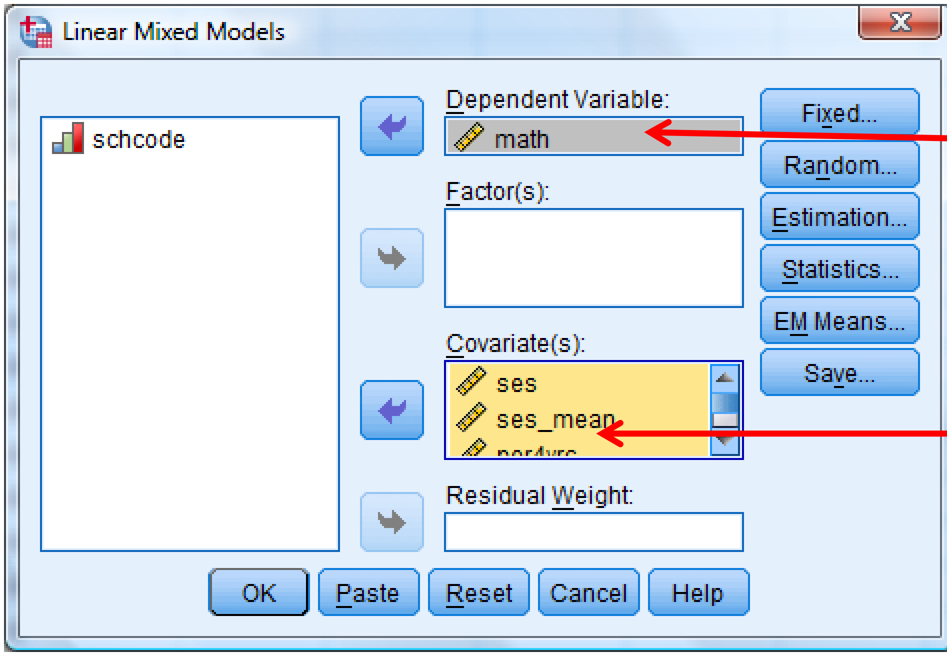
Random intercept model

Analyze -> Mixed Models -> Linear...



Cases grouped within school

Random intercept model

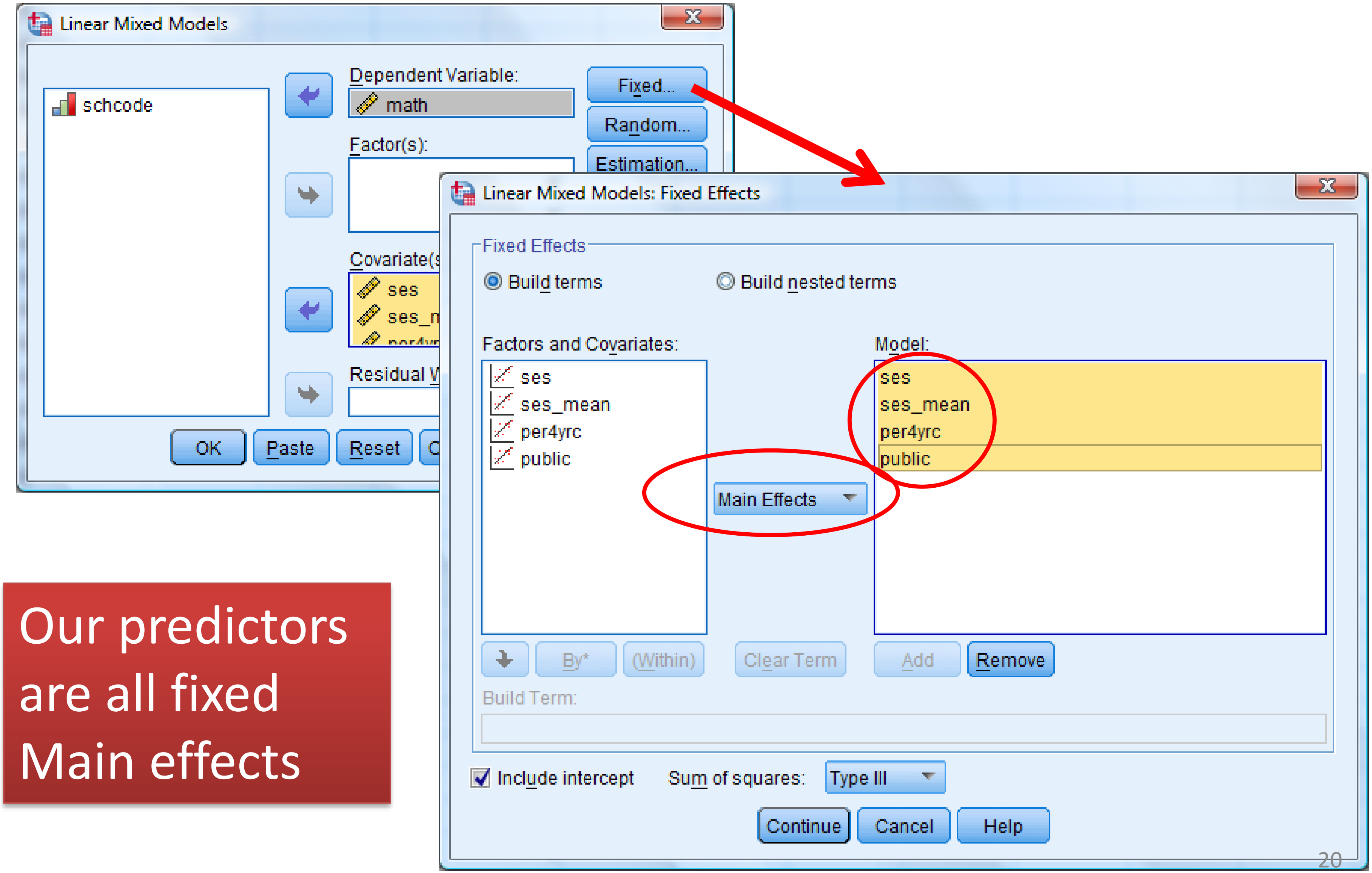


Predicting *math*

from the predictor variables

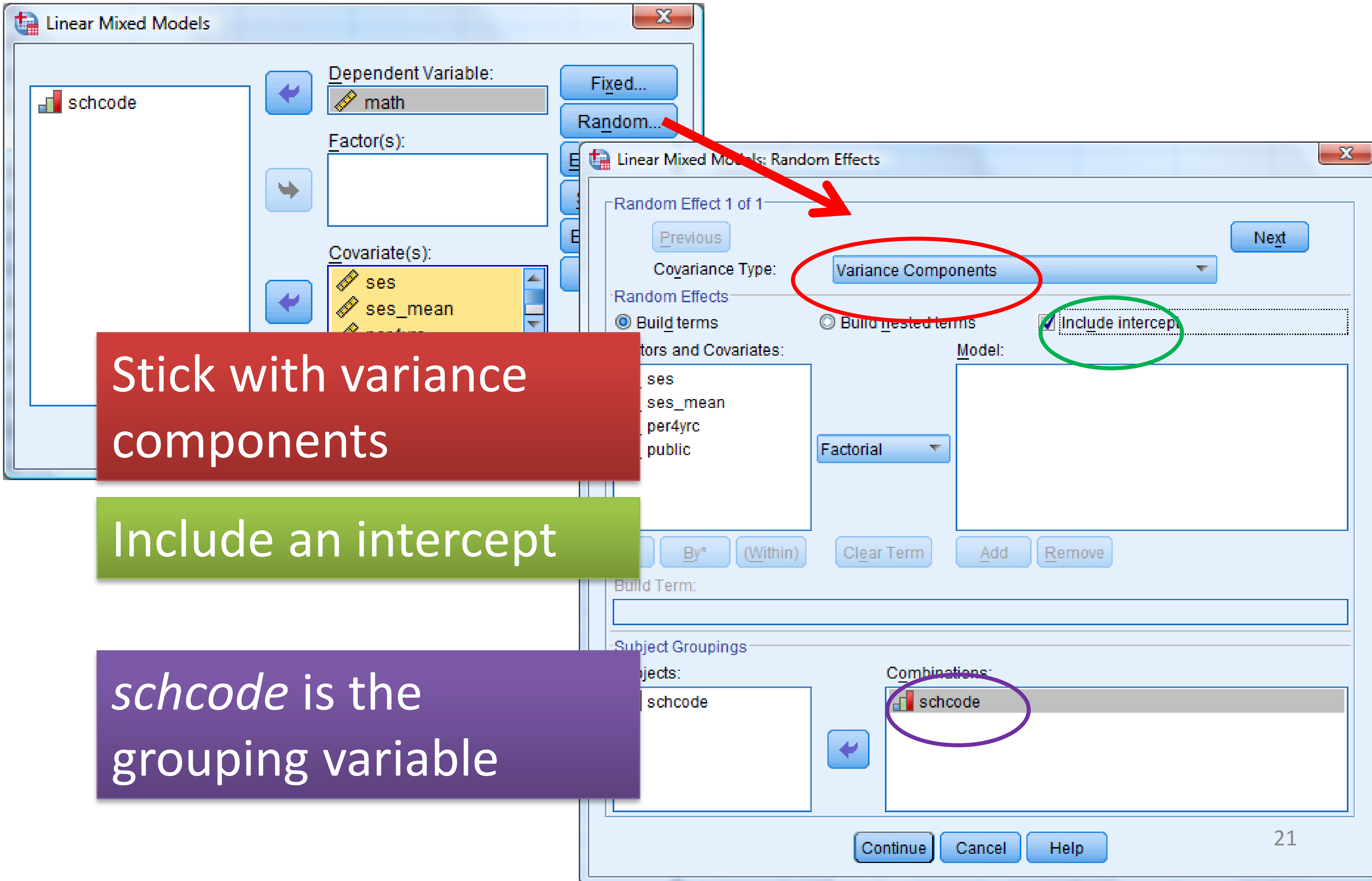
Note here that *public* is treated as a covariate because it is a binary variable (i.e coded 0,1). Binary variables are OK in regressions but variables with three or more categories, as well as dichomotous variables coded (say) 1,2, would have to be treated as factors (a little bit more complicated – I prefer to use binary variables if I can.

Random intercept model



Our predictors are all fixed Main effects

Random intercept model

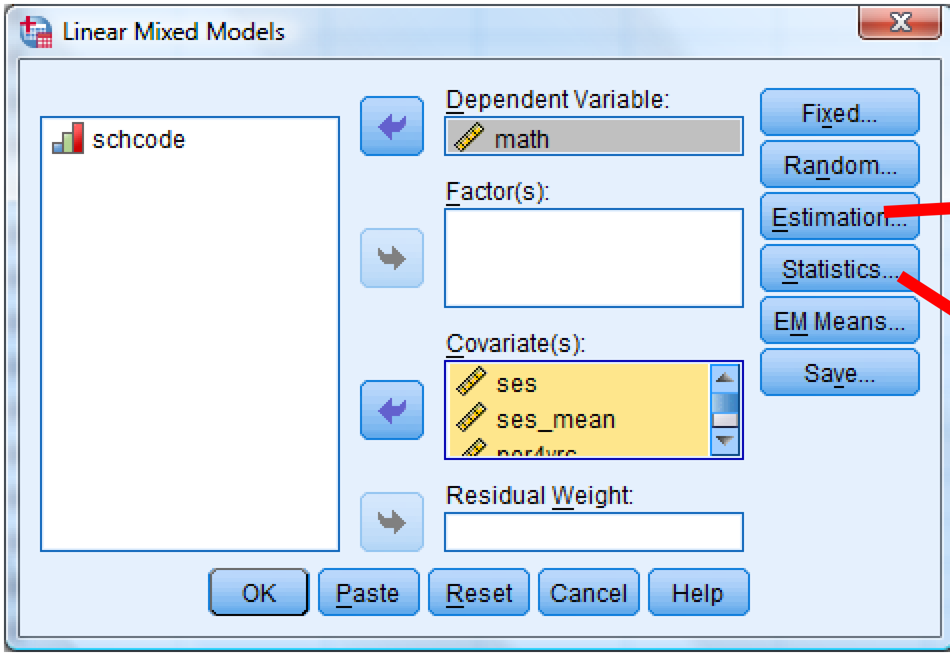


Stick with variance components

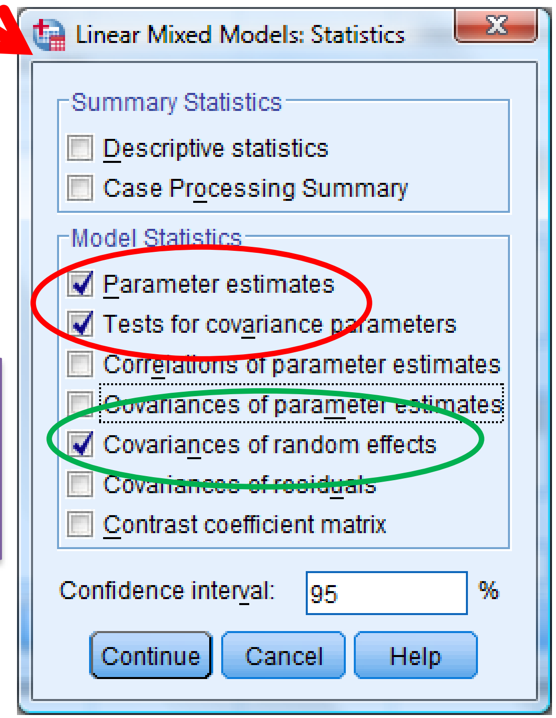
Include an intercept

schcode is the grouping variable

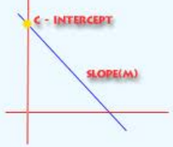
Random intercept model



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	
	ses_mean	1		1	
	per4yrc	1		1	
	public	1		1	
Random Effects	Intercept[b]	1	Variance Components	1	schcode
Residual				1	
Total		6		7	

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.

- Five fixed effects (the fixed intercept γ_{00} and fixed slope for each predictor)
- One random effect (the variance of the intercept u_{0j})
- One residual effect (the variance of the error term ε_{ij})

Seven parameters:

$$math_{ij} = \gamma_{00} + \gamma_{10}\, ses_{ij} + \gamma_{01}\, ses\_mean_{j} + \gamma_{02}\, per4yrc_j + \gamma_{03}\, public_j + u_{0j} + \varepsilon_{ij}$$

Output

SPSS

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	56.441552	.474433	421.055	118.966	.000	55.509001	57.374104
ses	3.190801	.157803	6448.937	20.220	.000	2.881455	3.500147
ses_mean	2.473244	.306897	709.247	8.059	.000	1.870709	3.075779
per4yrc	1.419812	.471391	413.879	3.012	.003	.493192	2.346432
public	-.164264	.275903	409.345	-.595	.552	-.706627	.378098

a. Dependent Variable: math.

Estimates for fixed effects:

Estimate for fixed intercept $\gamma_{00} = 56.44^*$

Estimate for slope of ses, $\gamma_{10} = 3.19^*$

of ses_mean, $\gamma_{01} = 2.47^*$

of per4yrc, $\gamma_{02} = 1.42^*$

of public, $\gamma_{03} = -0.16$

$$math_{ij} = \gamma_{00} + \gamma_{10}\, ses_{ij} + \gamma_{01}\, ses\_mean_{j} + \gamma_{02}\, per4yrc_j + \gamma_{03}\, public_j + u_{0j} + \varepsilon_{ij}$$

Output

Estimates of Covariance Parameters[a]

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		62.630370	1.102966	56.784	.000	60.505479	64.829885
Intercept [subject = schcode]	Variance	2.395178	.443654	5.399	.000	1.665987	3.443531

a. Dependent Variable: math.

Variance:

Previous intercept variance for null model = 10.64

Intercept variance for current model = 2.40

Reduction from null model = (10.64-2.40)/10.64 = 77%

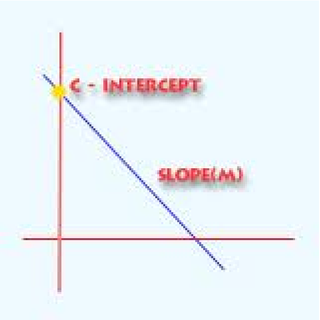
So the predictors account for 77% of level 2 variance (between groups)

Residual (level 1) variance in null model = 66.55

Residual variance here = 62.63

Reduction from null model = (66.55-62.63)/66.55 = 6%

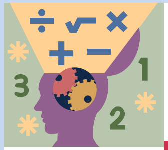
So, student SES accounts for 6% of level 1 variance in individual math performance.

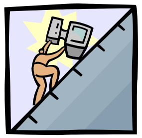


Section 2: Lecture 10

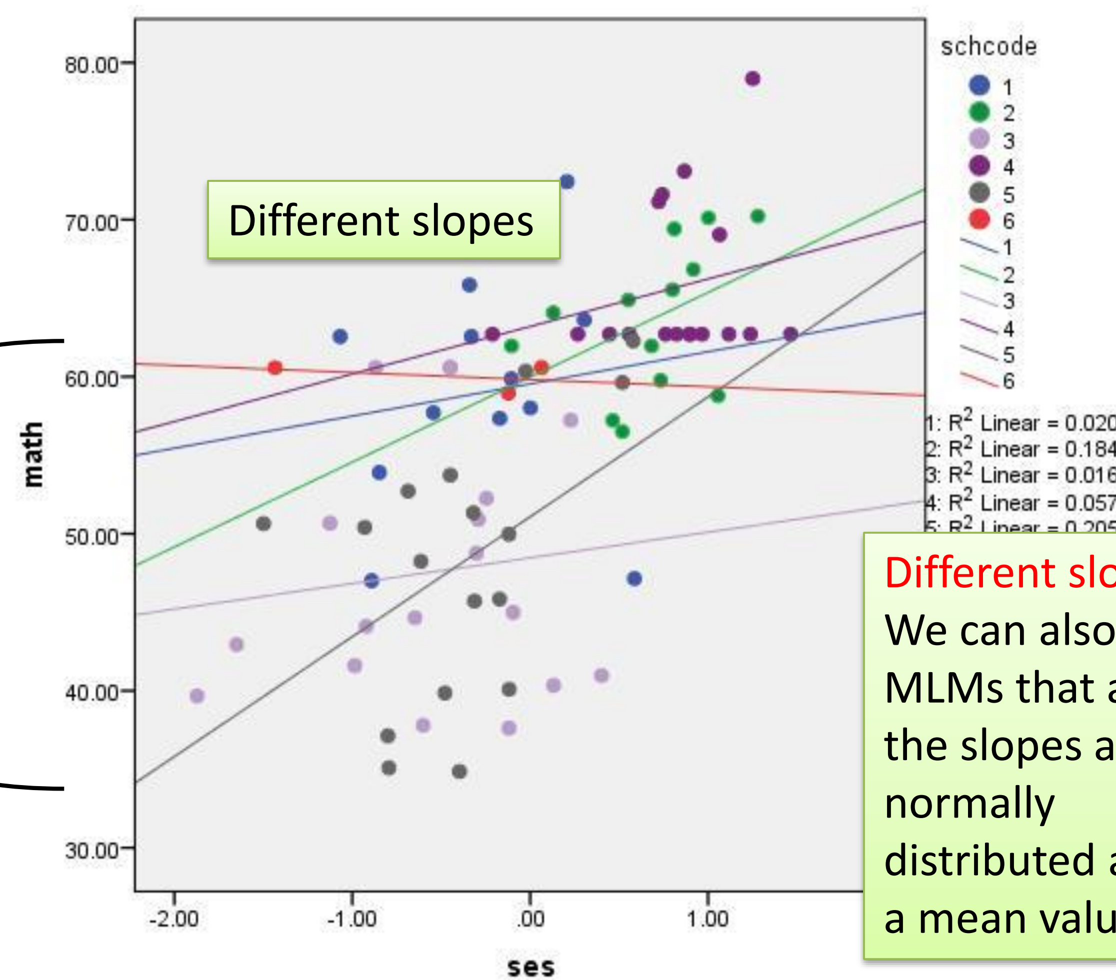
RANDOM SLOPE MODELS

- ❑ Basic ideas
- ❑ SPSS output
- ❑ Predicting slopes
- ❑ More SPSS output





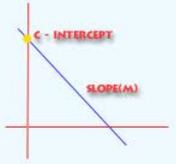
Regression models for each class



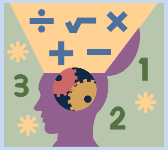
Different intercepts: Multilevel models assume that these intercepts are normally distributed around a mean value

Different slopes: We can also fit MLMs that assume the slopes are normally distributed around a mean value

Different linear models for 6 different schools



Random slope model



Level 1 equation: $Y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + \varepsilon_{ij}$

Level 2 equations: $\beta_{0j} = \gamma_{00} + \gamma_{01} z_j + u_{0j}$

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

so that

$$Y_{ij} = (\gamma_{00} + \gamma_{01} z_j + u_{0j}) + (\gamma_{10} + u_{1j}) x_{ij} + \varepsilon_{ij}$$

$$Y_{ij} = \gamma_{00} + \gamma_{10} x_{ij} + \gamma_{01} z_j + u_{0j} + u_{1j} x_{ij} + \varepsilon_{ij}$$

Level 1 equation:

$math_{ij} = \beta_{0j} + \beta_{1j}\, ses_{ij} + \varepsilon_{ij}$

Level 2 equations:

$\beta_{0j} = \gamma_{00} + \gamma_{01}\, ses\_mean_j + \gamma_{02}\, per4yrc_j + \gamma_{03}\, public_j + u_{0j}$

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

So that

$$math_{ij} = \gamma_{00} + \gamma_{10}\, ses_{ij} + \gamma_{01}\, ses\_mean_j + \gamma_{02}\, per4yrc_j + \gamma_{03}\, public_j + u_{0j} + \mathbf{u_{1j}\, ses_{ij}} + \varepsilon_{ij}$$

Random slope model

Level 1 equation:

$math_{ij} = \beta_{0j} + \beta_{1j}\, ses_{ij} + \varepsilon_{ij}$

Level 2 equations:

$\beta_{0j} = \gamma_{00} + \gamma_{01}\, ses\_mean_{j} + \gamma_{02}\, per4yrc_j + \gamma_{03}\, public_j + u_{0j}$

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

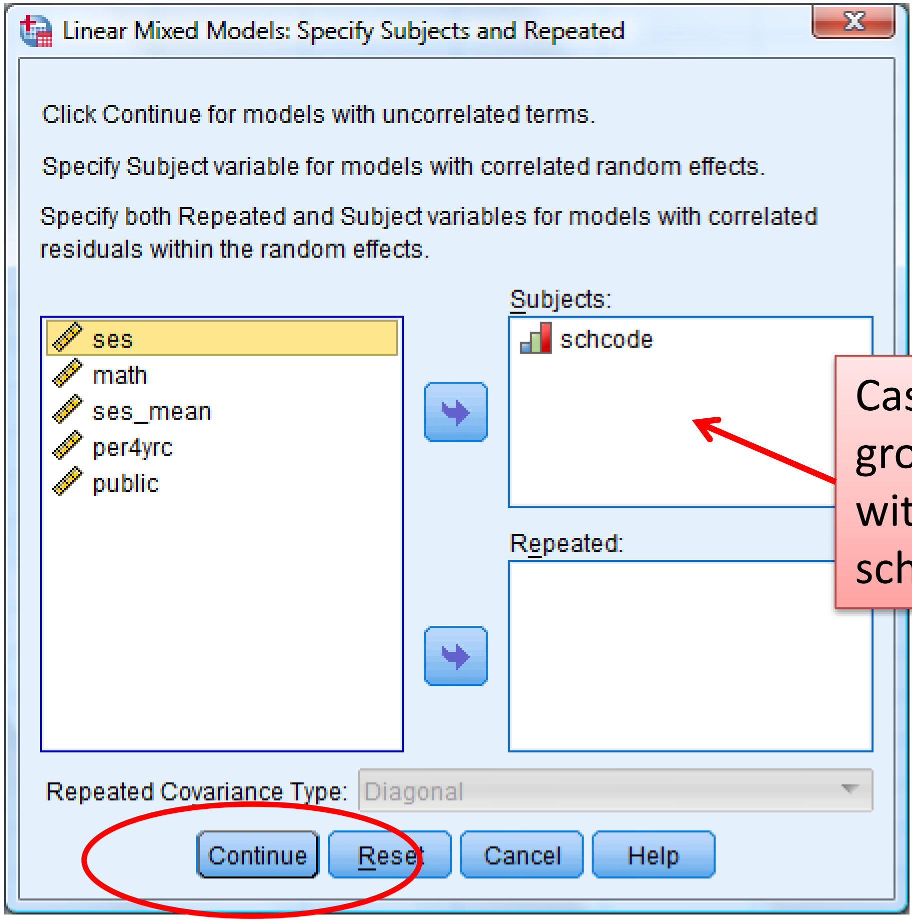
So that

$math_{ij} = \gamma_{00} + \boldsymbol{\gamma_{10}\, ses_{ij}} + \gamma_{01}\, ses\_mean_{j} + \gamma_{02}\, per4yrc_j + \gamma_{03}\, public_j + u_{0j} + \boldsymbol{u_{1j}\, ses_{ij}} + \varepsilon_{ij}$

ses is now both a fixed effect (as before) AND a random effect

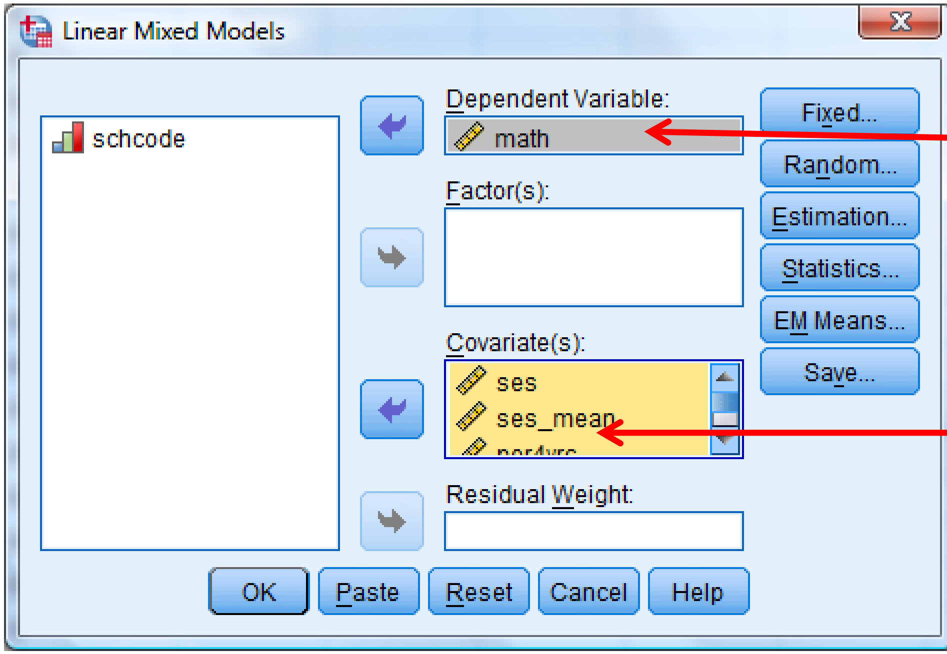
Random slope model

Analyze -> Mixed Models -> Linear...



Cases grouped within school

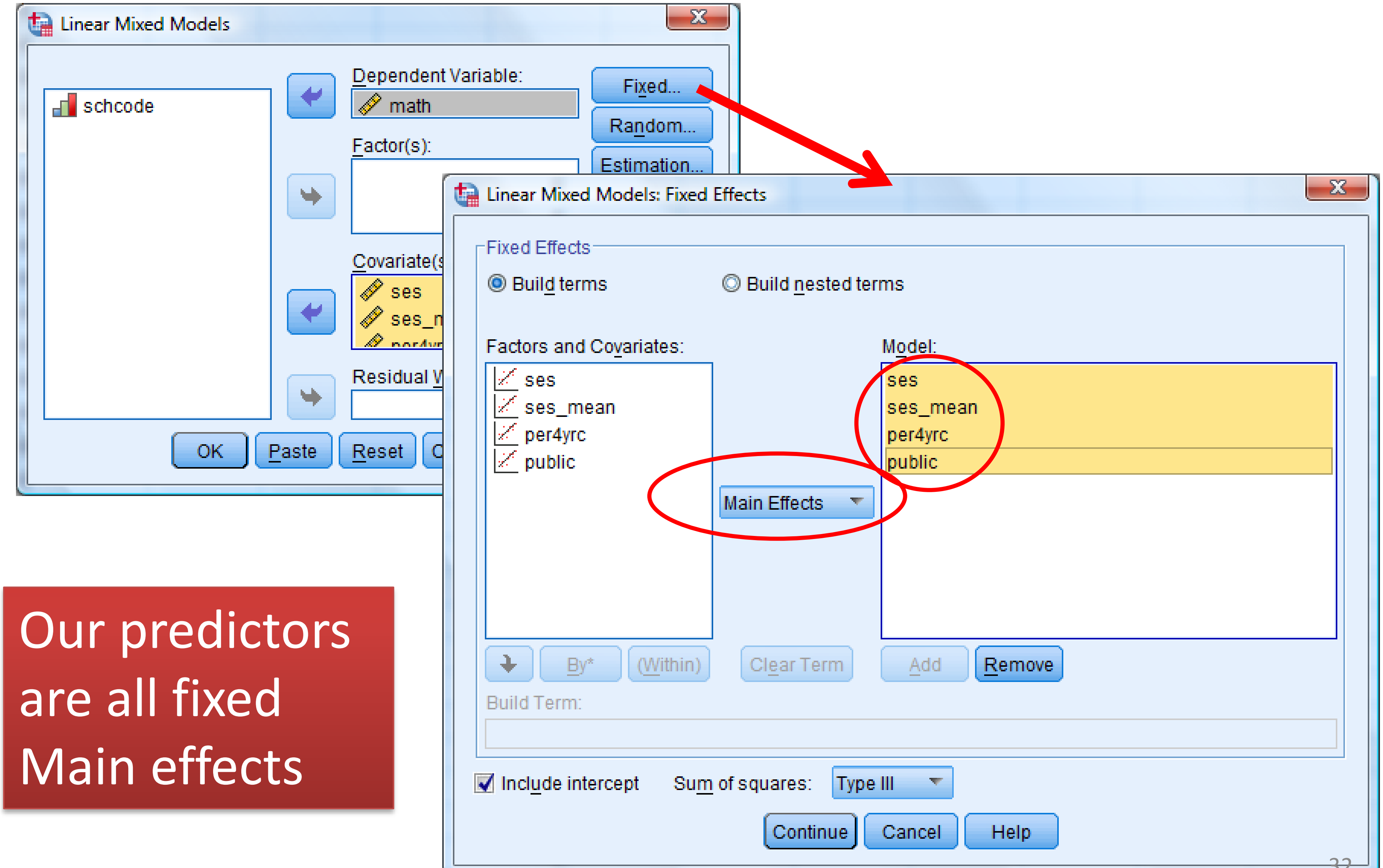
Random slope model



Predicting *math*

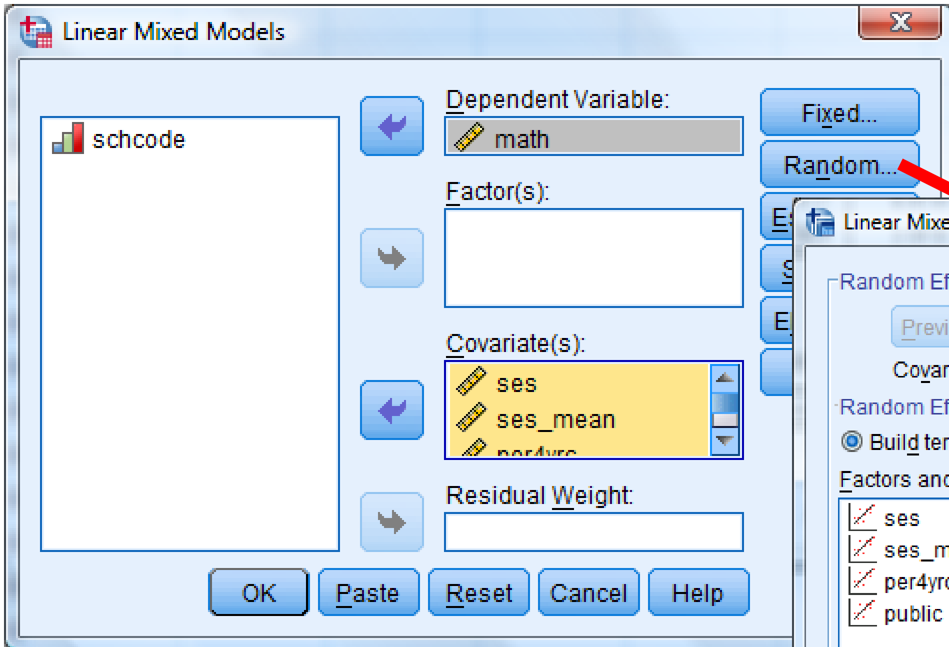
from the predictor variables

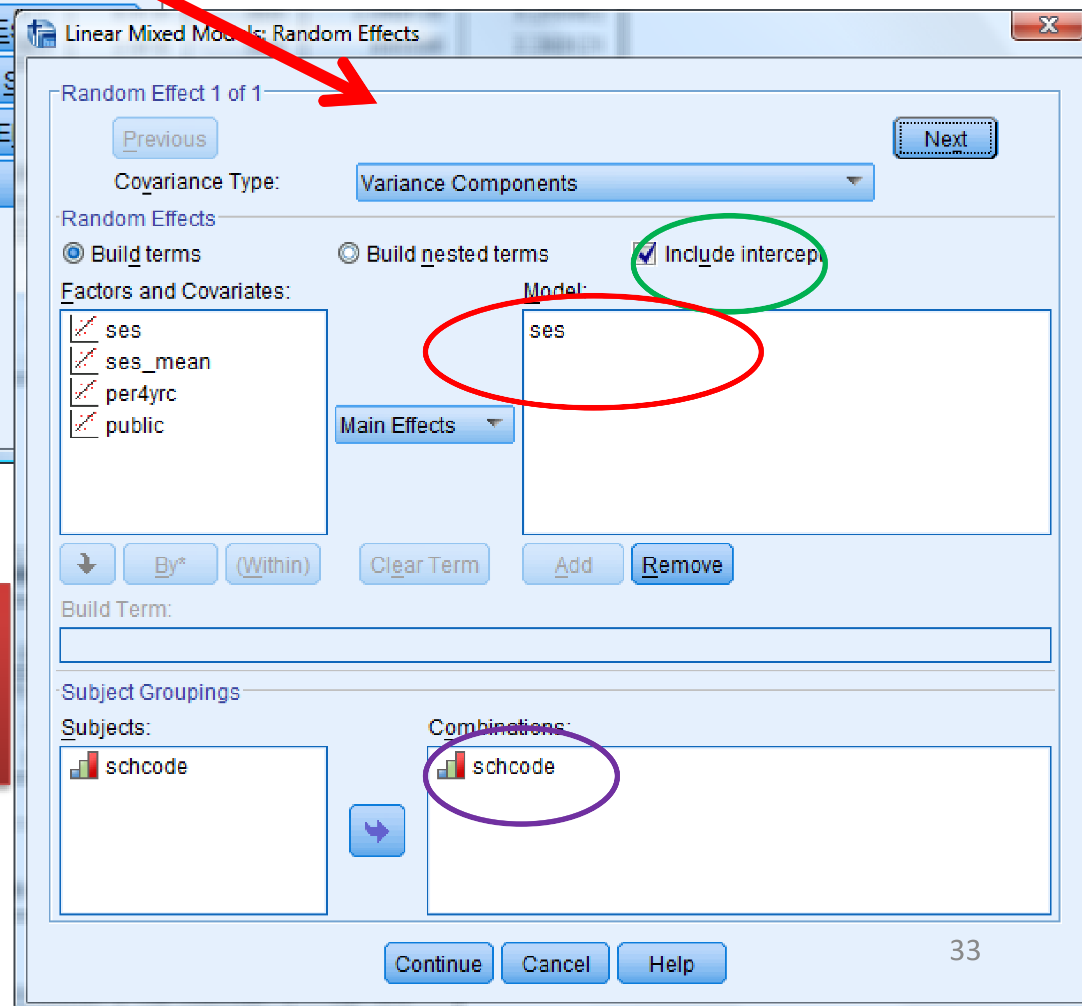
Random slope model



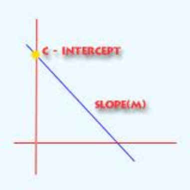
Our predictors are all fixed Main effects

Random slope model

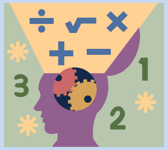




Now *ses* is a random effect

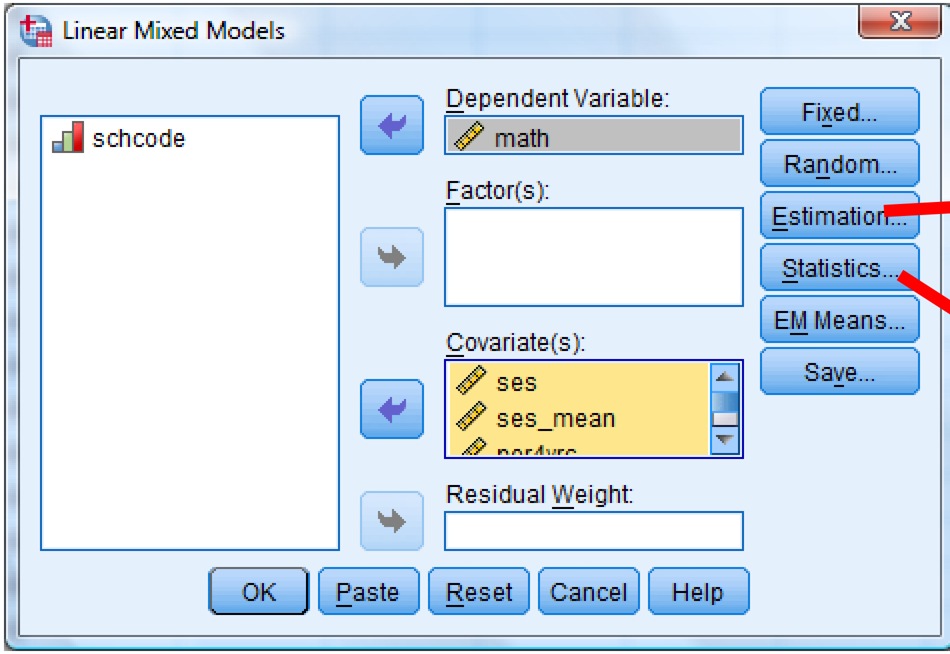


Why variance components?

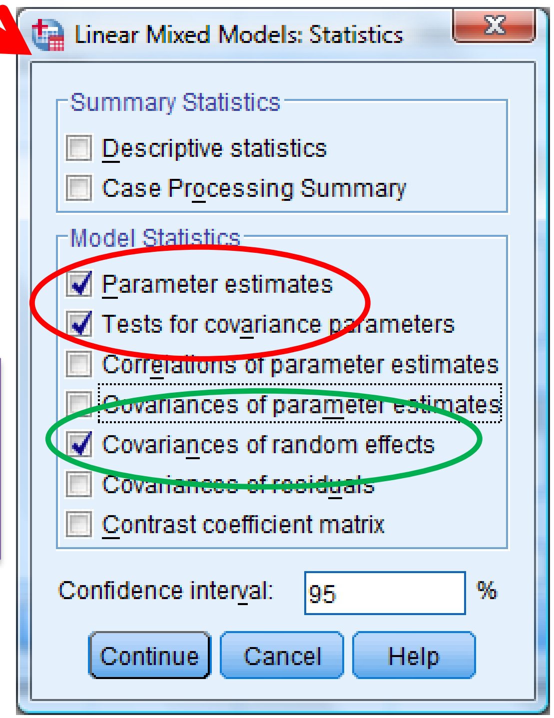


- The slopes could be correlated with the intercepts!
- Variance components does not include a parameter for the correlation.
- If you want to investigate whether the slopes and intercepts might be correlated, you can try “Unstructured”.
  - This does not impose any constraints on the random effects covariance matrix, whereas variance components assumes it is diagonal.
- However, “unstructured” can result in warning messages about lack of convergence.
  - If you get one of these, drop some variables or go back to Variance components

Random slope model



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	
	ses_mean	1		1	
	per4yrc	1		1	
	public	1		1	
Random Effects	Intercept + ses[b]	2	Variance Components	2	schcode
Residual				1	
Total		7		8	

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.

Five fixed effects (the fixed intercept γ_{00} and fixed slope for each predictor)

Two random effects

One residual effect

Eight parameters:

$$math_{ij} = \gamma_{00} + \gamma_{10}\, ses_{ij} + \gamma_{01}\, ses\_mean_{j} + \gamma_{02}\, per4yrc_j + \gamma_{03}\, public_j + u_{0j} + u_{1j}\, ses_{ij} + \varepsilon_{ij}$$

Output

SPSS

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	56.469785	.471568	419.501	119.749	.000	55.542854	57.396716
ses	3.163898	.168888	635.541	18.734	.000	2.832252	3.495544
ses_mean	2.659588	.313631	698.064	8.480	.000	2.043815	3.275361
per4yrc	1.360179	.467933	410.212	2.907	.004	.440334	2.280025
public	-.119986	.274402	407.915	-.437	.662	-.659405	.419433

a. Dependent Variable: math.

Estimates for fixed effects:

Estimate for fixed intercept $\gamma_{00} = 56.47^*$

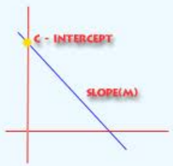
Estimate for slope of ses, $\gamma_{10} = 3.16^*$

of ses_mean, $\gamma_{01} = 2.66^*$

of per4yrc, $\gamma_{02} = 1.36^*$

of public $\gamma_{03} = -0.12$

$$math_{ij} = \gamma_{00} + \gamma_{10}\, ses_{ij} + \gamma_{01}\, ses\_mean_j + \gamma_{02}\, per4yrc_j + \gamma_{03}\, public_j + u_{0j} + u_{1j}\, ses_{ij} + \varepsilon_{ij}$$



Output

Estimates of Covariance Parameters[a]

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		62.114614	1.111312	55.893	.000	59.974229	64.331386
Intercept [subject = schcode]	Variance	2.112261	.445499	4.741	.000	1.397075	3.193563
ses [subject = schcode]	Variance	1.314246	.566455	2.320	.020	.564675	3.058824

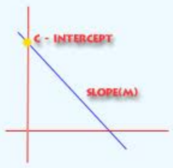
a. Dependent Variable: math.

Variance:

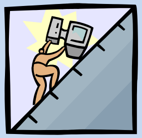
Much of the results here are as in the random intercept model, but now there is a significant slope variance of 1.31.

So we have evidence that the slopes vary across the schools in the sample.

Can we explain this variation?



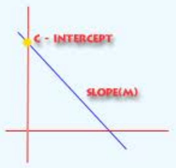
More complex random slope models



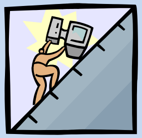
Level 1 equation: $Y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + \varepsilon_{ij}$

Level 2 equations: $\beta_{0j} = \gamma_{00} + \gamma_{01} z_j + u_{0j}$

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



More complex random slope models



Level 1 equation: $Y_{ij} = \beta_{0j} + \beta_{1j}\, x_{ij} + \varepsilon_{ij}$

Level 2 equations: $\beta_{0j} = \gamma_{00} + \gamma_{01}\, z_j + u_{0j}$

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

so that

$$Y_{ij} = (\gamma_{00} + \gamma_{01}\, z_j + u_{0j}) + (\gamma_{10} \boldsymbol{+ \gamma_{11}\, z_j} + u_{1j})x_{ij} + \varepsilon_{ij}$$

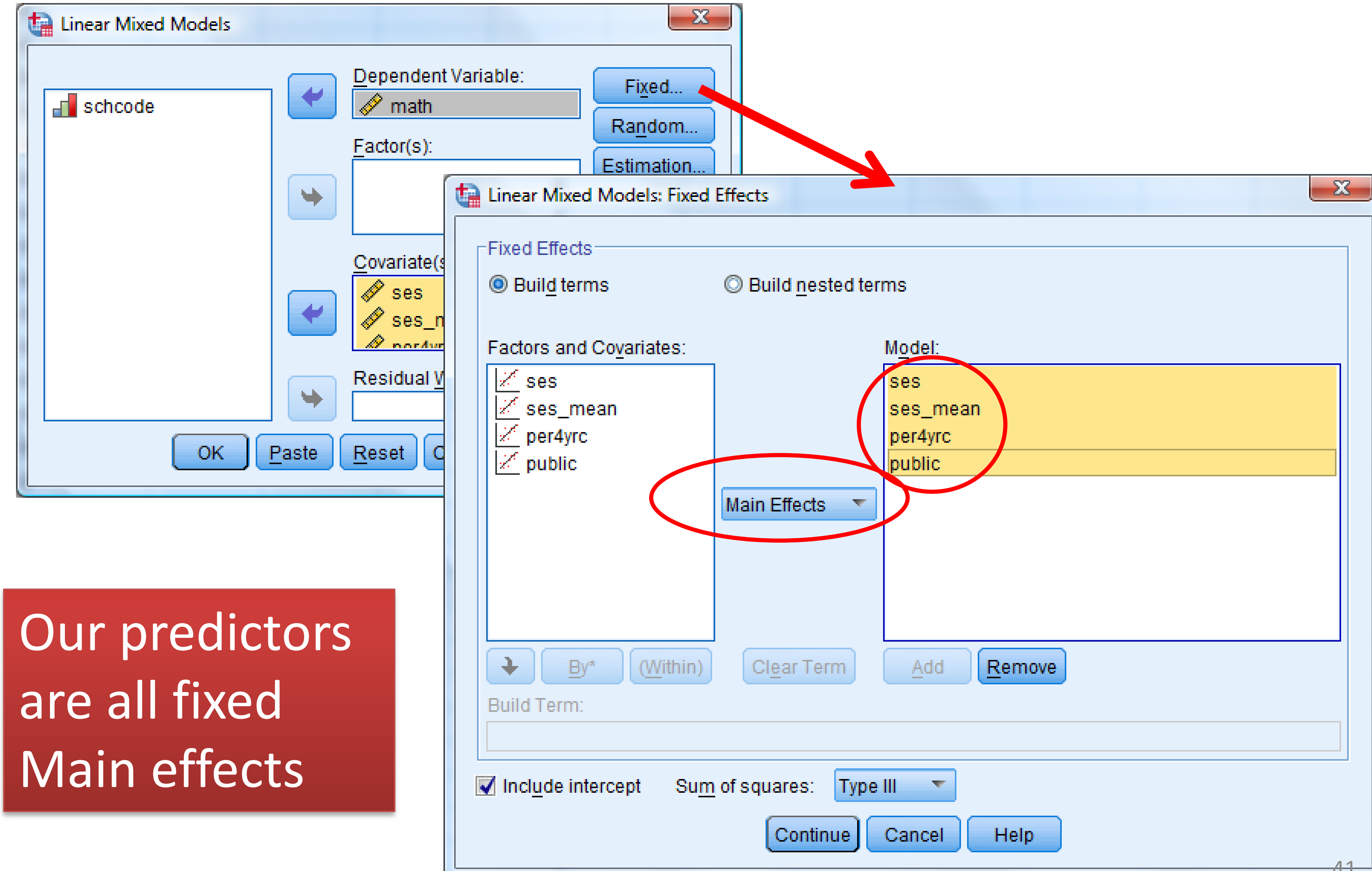
$$= \gamma_{00} + \gamma_{10}\, x_{ij} + \gamma_{01}\, z_j + \boldsymbol{\gamma_{11}\, z_j\, x_{ij}} + u_{0j} + u_{1j}\, x_{ij} + \varepsilon_{ij}$$

A fixed effect that is the interaction of the group and individual effects

So that

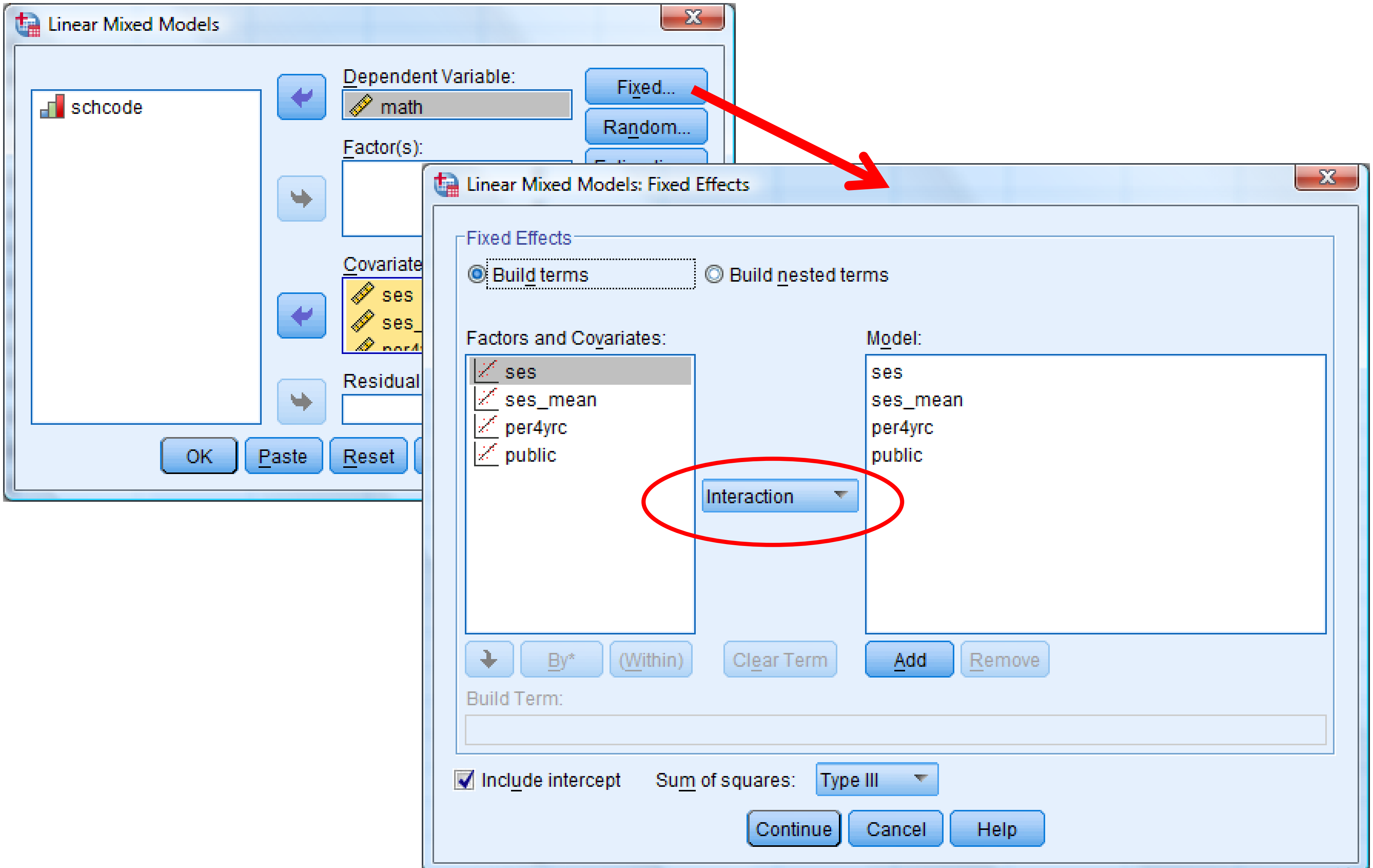
$$\begin{aligned} math_{ij} = {} & \gamma_{00} + \gamma_{10}\, ses_{ij} + \gamma_{01}\, ses\_mean_j + \gamma_{02}\, per4yrc_j + \gamma_{03}\, public_j\ + \\ & + \gamma_{11}\,(ses\_mean * \mathrm{ses}) + \gamma_{12}\,(per4yrc * \mathrm{ses}) + \gamma_{13}\,(public * \mathrm{ses}) + \\ & + u_{0j} + u_{1j}\, ses_{ij} + \varepsilon_{ij} \end{aligned}$$

Before we had …

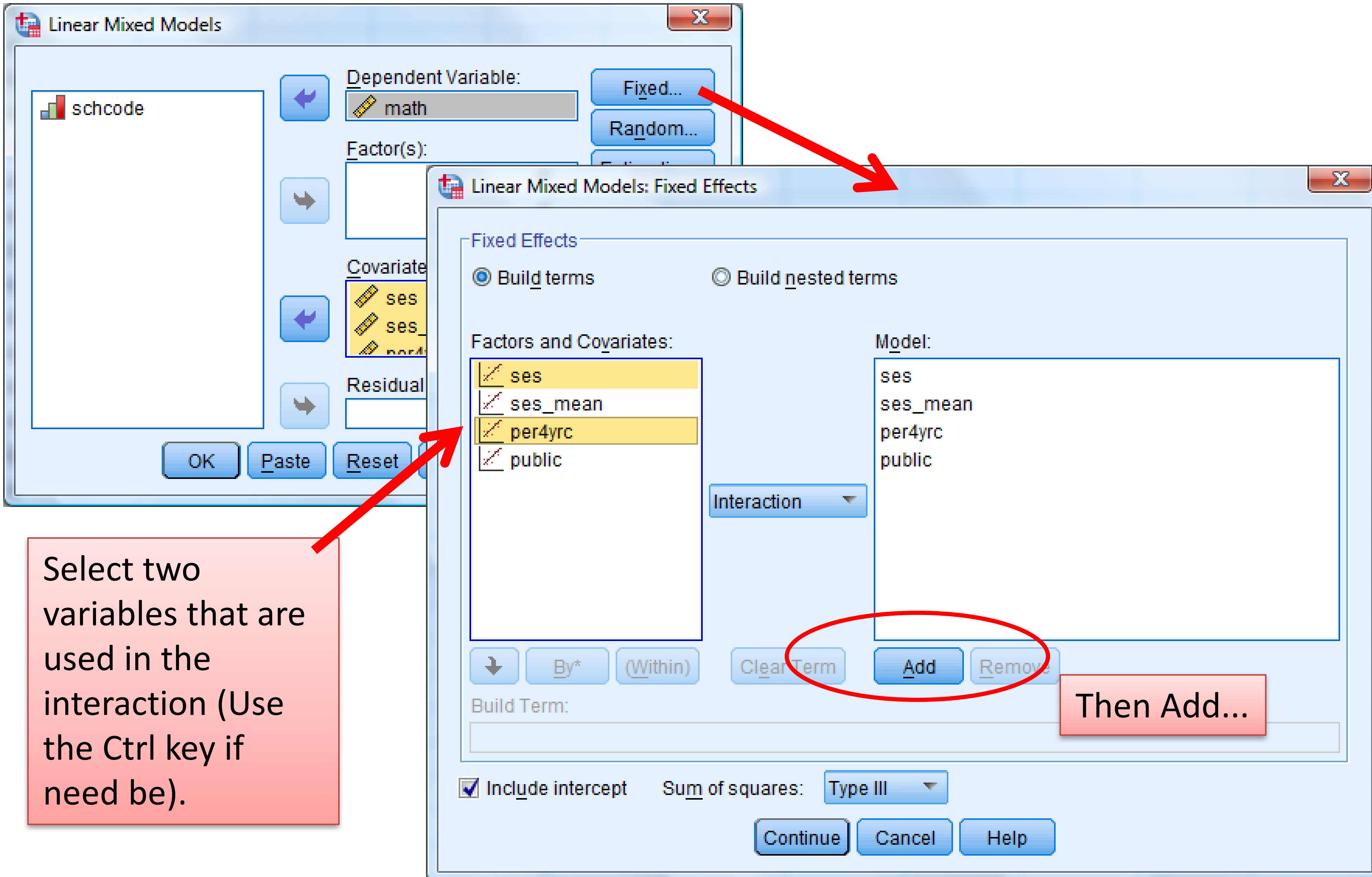


Our predictors are all fixed Main effects

Now we want …

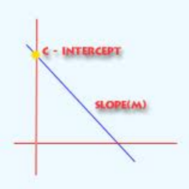

Linear Mixed Models
schcode
Dependent Variable:
math
Fixed...
Random...
Factor(s):
Covariate
ses
ses_
Residual
OK
Paste
Reset
Linear Mixed Models: Fixed Effects
Fixed Effects
Build terms
Build nested terms
Factors and Covariates:
ses
ses_mean
per4yrc
public
Model:
ses
ses_mean
per4yrc
public
Interaction
By*
(Within)
Clear Term
Add
Remove
Build Term:
Include intercept
Sum of squares:
Type III
Continue
Cancel
Help

Now we want …

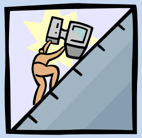


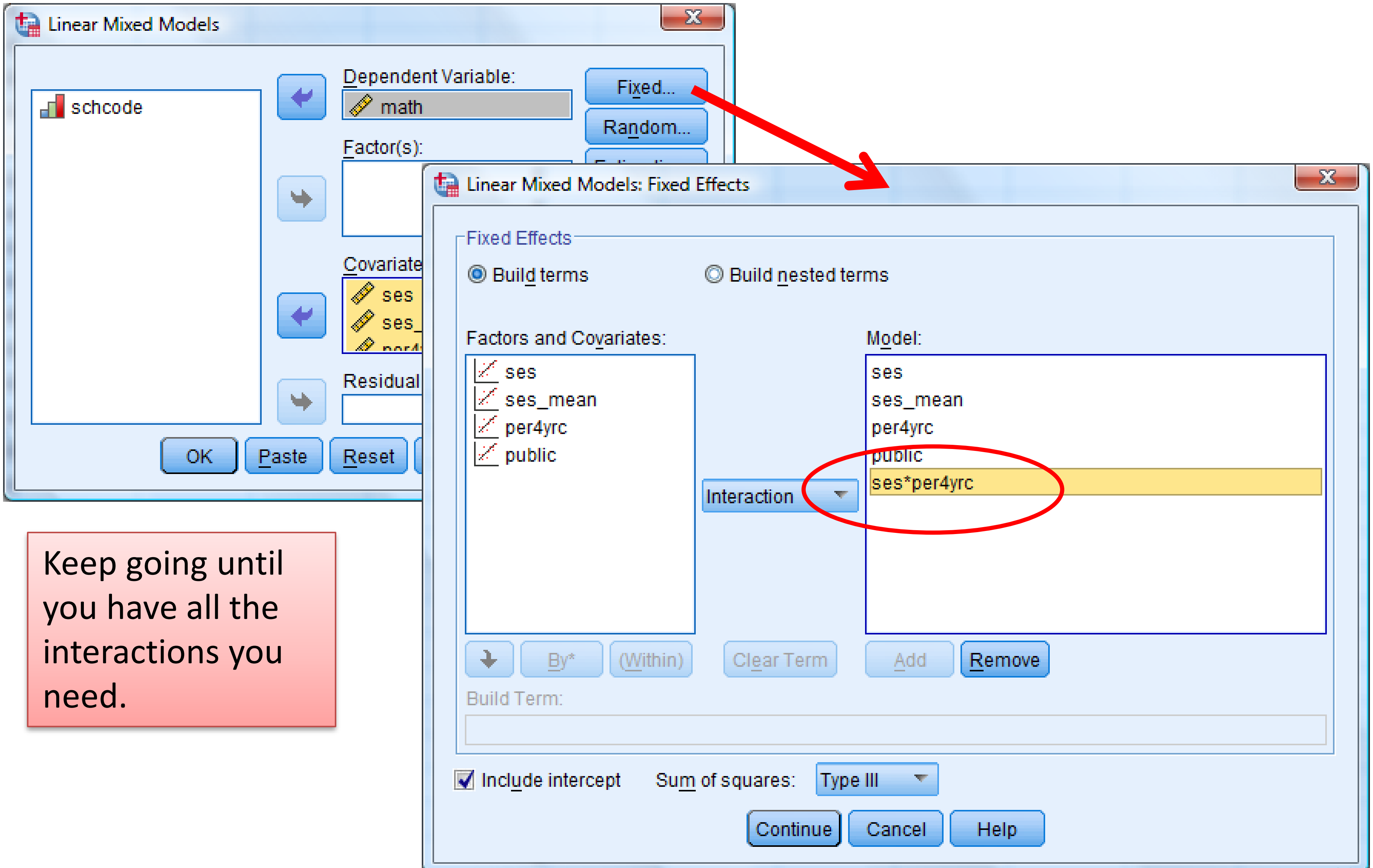
Select two variables that are used in the interaction (Use the Ctrl key if need be).

Then Add…



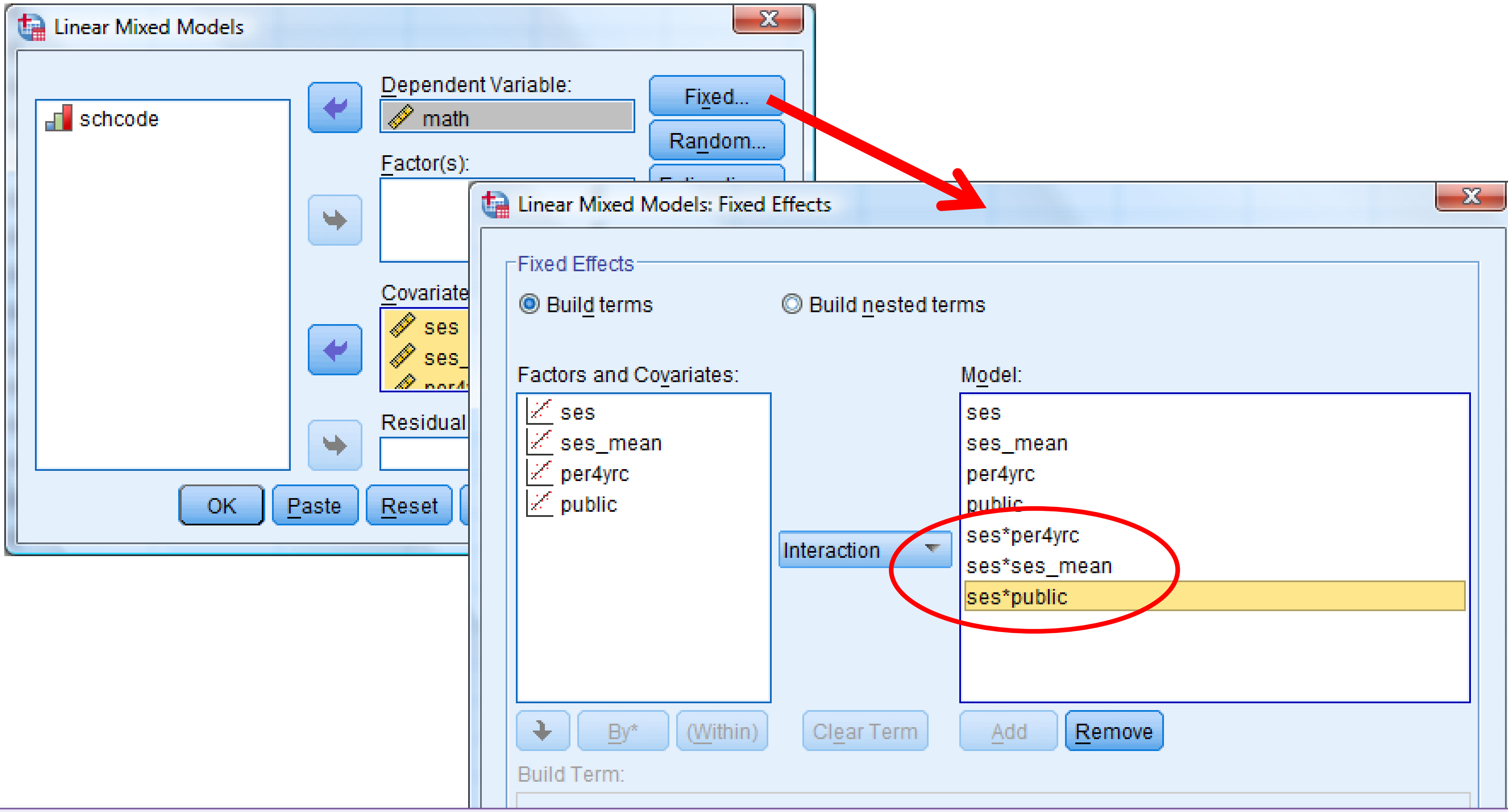
Now we want ...





Keep going until you have all the interactions you need.

Now we want ...



$$
\begin{aligned}
math_{ij} = \gamma_{00} &+ \gamma_{10}\, ses_{ij} + \gamma_{01}\, ses\_mean_j + \gamma_{02}\, per4yrc_j + \gamma_{03}\, public_j\ + \\
&+ \gamma_{11}\,(ses\_mean * ses) + \gamma_{12}\,(per4yrc * ses) + \gamma_{13}\,(public * ses) + \\
&+ u_{0j} + u_{1j}\, ses_{ij} + \varepsilon_{ij}
\end{aligned}
$$

Output

Model Dimension[a]

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	ses	1		1	
	ses_mean	1		1	
	per4yrc	1		1	
	public	1		1	
	ses * per4yrc	1		1	
	ses * ses_mean	1		1	
	ses * public	1		1	
Random Effects	Intercept + ses[b]	2	Variance Components	2	schcode
Residual				1	
Total		10		11	

a. Dependent Variable: math.

Model dimension sets out the shape of the data.

It's the same as before EXCEPT that now we have the three interaction effects.

Eleven parameters:

$$math_{ij} = \gamma_{00} + \gamma_{10}\, ses_{ij} + \gamma_{01}\, ses\_mean_{j} + \gamma_{02}\, per4yrc_{j} + \gamma_{03}\, public_{j} + \\ + \gamma_{11}\, (ses\_mean * \mathrm{ses}) + \gamma_{12}\, (per4yrc * \mathrm{ses}) + \gamma_{13}\, (public * \mathrm{ses}) + \\ + u_{0j} + u_{1j}\, ses_{ij} + \varepsilon_{ij}$$

Output

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	56.505254	.485329	450.229	116.427	.000	55.551462	57.459046
ses	3.757343	.605681	518.400	6.203	.000	2.567451	4.947235
ses_mean	2.706473	.324107	759.554	8.351	.000	2.070222	3.342724
per4yrc	1.361887	.479336	439.899	2.841	.005	.419814	2.303960
public	-.119925	.274238	405.467	-.437	.662	-.659031	.419182
ses * per4yrc	-.130132	.592163	479.307	-.220	.826	-1.293688	1.033423
ses * ses_mean	-.136539	.298571	303.798	-.457	.648	-.724069	.450990
ses * public	-.668237	.331145	404.838	-2.018	.044	-1.319216	-.017258

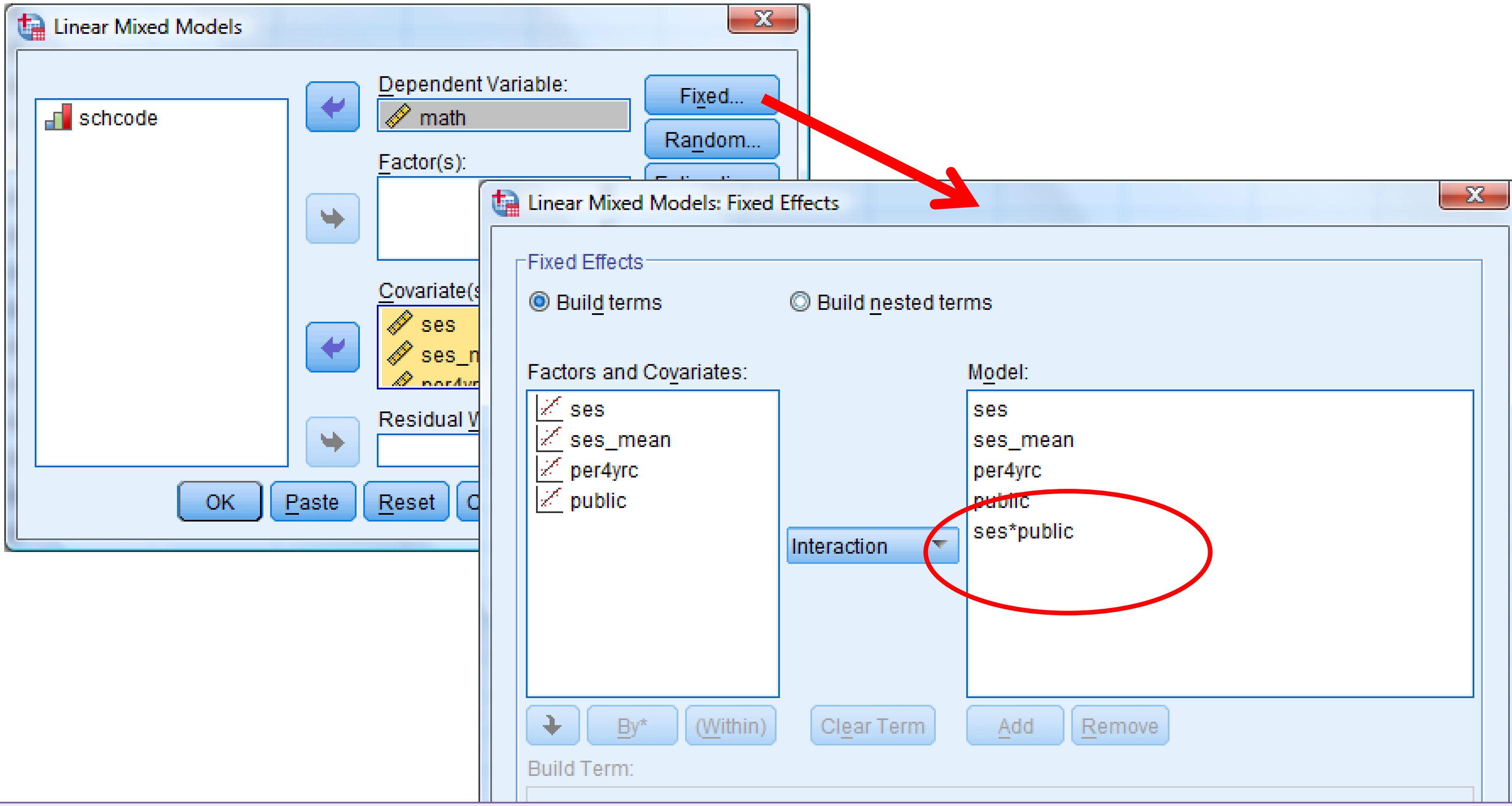
a. Dependent Variable: math.

Estimates for fixed effects:

Same significant main effects as before, but only one significant interaction – ses*public

We can remove the non-significant interactions for a more parsimonious model

Simpler model



$$
\begin{aligned}
math_{ij} = {} & \gamma_{00} + \gamma_{10}\, ses_{\,ij} + \gamma_{01}\, ses\_mean_{\,j} + \gamma_{02}\, per4yrc_{j} + \gamma_{03}\, public_{j} \; + \\
& + \gamma_{13}\, (public{*}ses) + \\
& + u_{0j} + u_{1j}\, ses_{\,ij} + \varepsilon_{ij}
\end{aligned}
$$

Output 2

Model Dimension[a]

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	ses	1		1	
	ses_mean	1		1	
	per4yrc	1		1	
	public	1		1	
	ses * public	1		1	
Random Effects	Intercept + ses[b]	2	Variance Components	2	schcode
Residual				1	
Total		8		9	

a. Dependent Variable: math.

Model dimension sets out the shape of the data.

Nine parameters:

$$math_{ij} = \gamma_{00} + \gamma_{10}\, ses_{ij} + \gamma_{01}\, ses\_mean_{j} + \gamma_{02}\, per4yrc_{j} + \gamma_{03}\, public_{j} + \\ + \gamma_{13}\,(public{*}ses) + \\ + u_{0j} + u_{1j}\, ses_{ij} + \varepsilon_{ij}$$

Output 2

Estimates of Fixed Effects[a]

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	56.440337	.470907	418.862	119.855	.000	55.514701	57.365972
ses	3.659485	.292123	461.438	12.527	.000	3.085429	4.233542
ses_mean	2.659142	.313166	697.326	8.491	.000	2.044280	3.274004
per4yrc	1.404177	.467536	410.165	3.003	.003	.485112	2.323242
public	-.123093	.273888	406.624	-.449	.653	-.661507	.415320
ses * public	-.682460	.328354	395.850	-2.078	.038	-1.327996	-.036925

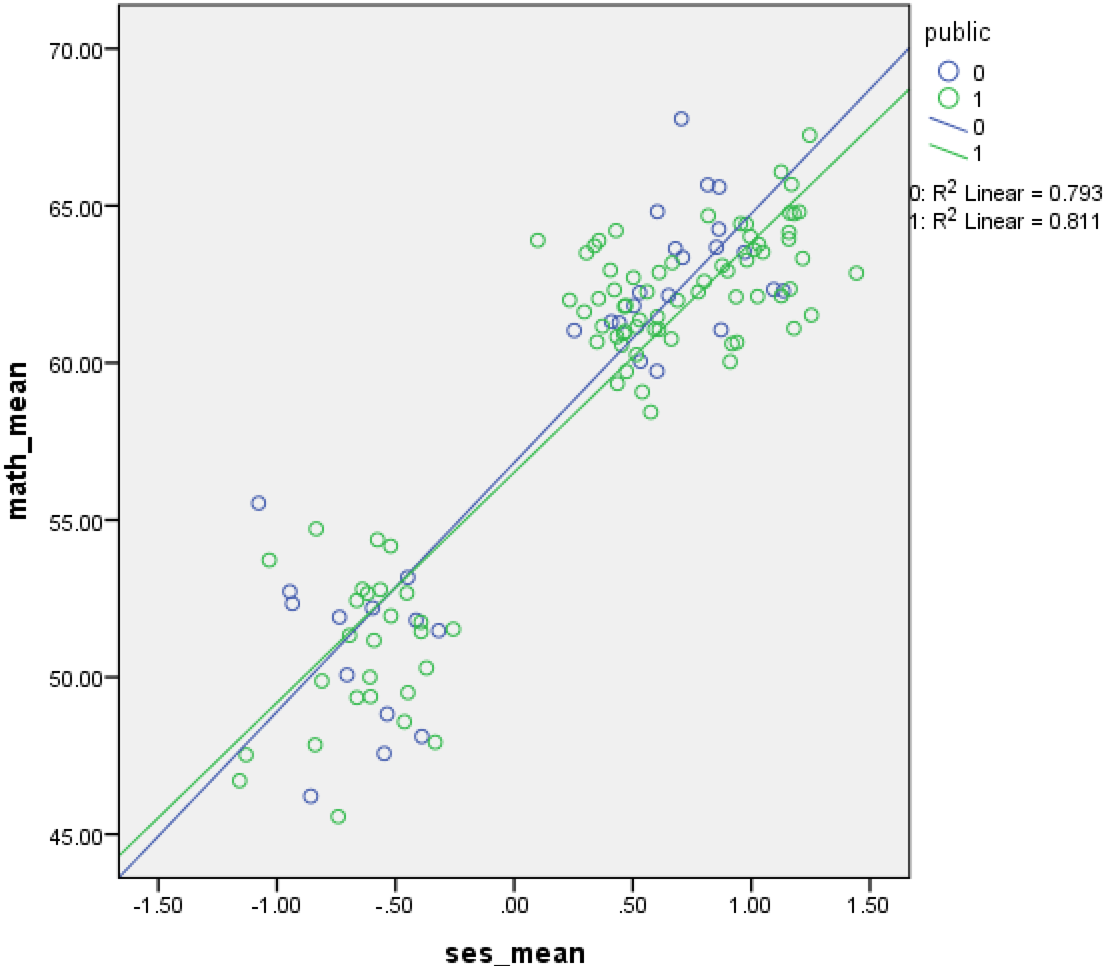
a. Dependent Variable: math.

Estimates for fixed effects:

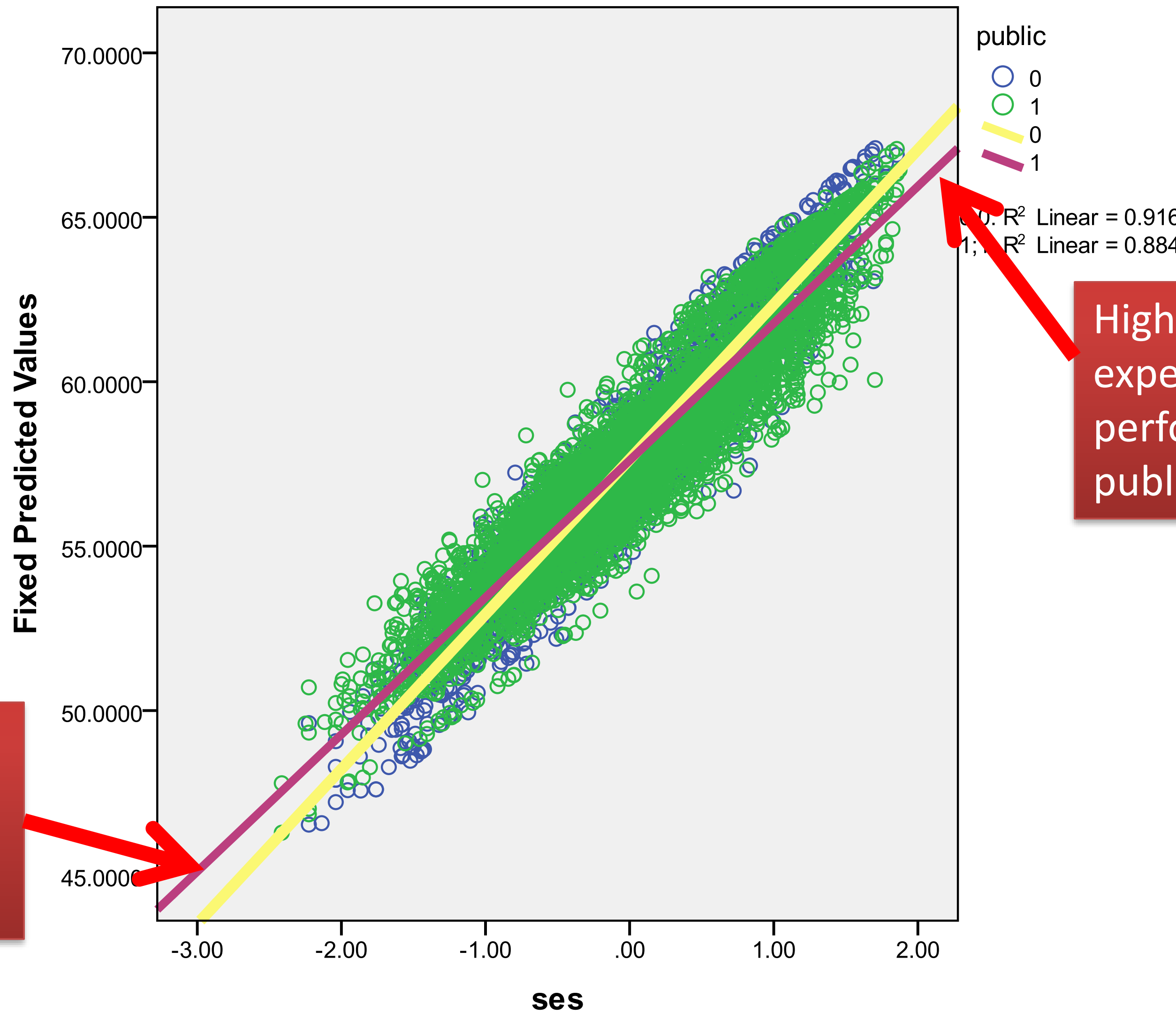
Same significant main effects as before, and a significant slope effect for public, estimated at – 0.68.

Remember *public* = 1 indicates a public school.
So, the slope is *lower* for public schools.
Higher SES students in public schools get lower marks
Lower SES students in public schools get higher marks

At school level
Extreme school performance



At individual level



Higher SES kids expected to perform worse in public schools

Lower SES kids expected to perform better in public schools

Interpreting the results

Fixed effects in the model

$$math = 56.4 + 3.7\, ses + 2.7\, ses\_mean + 1.4\, per4yrc - 0.12\, public - 0.68\, (public*ses)$$

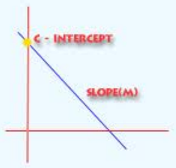
Descriptive Statistics

	Minimum	Maximum
ses	-2.41	1.87
ses_mean	-1.30	1.44
per4yrc	.00	1.00

Overall, estimated average performance on the test is 56.4.

Suppose everything else is equal:

1. Compare two students from lowest and highest ses-mean school:
   Expected difference in scores = 2.7 X 1.44 – 2.7 X (-1.3) = 7.4
2. Two students from lowest and highest *per4yrc* schools
   Expected difference in scores = 1.4 X 1.0 – 1.4 X (0) = 1.4
3. Two students with lowest and highest *ses* scores
   - Both in public school
     3.7 X 1.87 – 0.68 X 1.87 – [3.7X (-2.41) – 0.68 X (-2.41)] = 12.9
   - Both in private school
     3.7 X 1.87 – 3.7X (-2.41) = 15.8



Summary of results

- The final model suggests the following effects
- Student maths performance:
  - is positively predicted by student Socio Economic Status
  - and positively predicted by average SES of the school
  - as well as by the percentage of students in the school with ambitions towards obtaining more prestigious tertiary degrees
    - (as measured by intention to attend universities with a 4 year undergraduate program)
- The significant random slope effect suggests that higher SES students will perform better in non-public schools and that lower SES students will perform better in public schools.

Section 3: Lecture 10

ADDITIONAL ISSUES

- ❑ Centring
- ❑ Assumptions
- ❑ Estimation
- ❑ Covariance structures

Additional issues for MLM

- ***Centring***: It is often recommended that continuous predictors in MLMs be mean-centred, especially if they are going to appear in an interaction.
  - *Grand mean centring*: Compute a new variable by subtracting the overall mean from it.
  - *Group mean centring*: Compute a new variable by subtracting the group mean from it.
    - (You can do this by calculating the group mean for each group – see last week, re aggregation – and then subtract one variable from the other.)
- One point of centring is easier interpretation. A zero score is then the mean value.

Additional issues for MLM

- ***Assumptions***: (Pretty much) all the assumptions of regression still apply. Additionally, the random effects are assumed normally distributed.

- ***Estimation***: In the examples we have used Restricted maximum likelihood (REML), rather than maximum likelihood (ML). In most circumstances this won't make much difference, but ML needs to be used if you are going to compare fit across different models.

- ***Covariance structures:*** In this and previous lectures, we have not looked closely at different covariance structures for the random effects.
  - In effect we have assumed that the random effects are uncorrelated with each other. This may not be the case.
  - If you are using MLMs for your own research, you may need to consider this. Check out Field, 2009, p. 737, or Heck et al, p. 94.

Additional issues for MLM

- ***Covariance structures***: *4 popular flavours (of 17 offered by SPSS)*

Variance components

$$\sigma^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Diagonal

$$\begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}$$

AR1 – autoregressive

$$\sigma_\varepsilon^2 \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}$$

Unstructured

$$\begin{bmatrix} \sigma_1^2 & \sigma_{21} & \sigma_{31} \\ \sigma_{21} & \sigma_2^2 & \sigma_{32} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}$$

Summary

IN THIS LECTURE, you

- extended knowledge of random intercept models from the previous lecture
- learnt about random slope models
- learnt how to fit these models in SPSS
- were provided with summary information on some additional issues