### Mixed Models for Multilevel Data Analysis: An Applied Introduction

Don Hedeker Department of Public Health Sciences Biological Sciences Division University of Chicago

email: hedeker@uchicago.edu

This work was supported by National Institute of Mental Health Contract N44MH32056.

### What are Multilevel Data?

- Data that are hierarchically structured, nested, clustered
- Data collected from units organized or observed within units at a higher level (from which data are also obtained)

data collected on	who are clustered within
students	classrooms
siblings	families
repeated observations	individuals

==> these are examples of two-level data

- **level 1** (students) measurement of primary outcome and important mediating variables
- level 2 (classrooms) provides context or organization of level-1 units which may influence outcome; other mediating variables

### What is Multilevel Data Analysis?

"any set of analytical procedures that involve data gathered from individuals and from the social structure in which they are embedded and are analyzed in a manner that models the multilevel structure"

L. Burstein, Units of Analysis, 1985, Int Ency of Educ

- analysis that models the multilevel structure
- recognizes influence of structure on individual outcome

structure	may influence response from
classroom	students
family	siblings
individual	repeated observations

#### Why do Multilevel Data Analysis?

- assess amount of variability due to each level (*e.g.*, family variance and individual variance)
- model level 1 outcome in terms of effects at both levels

 $individual \ var. = fn(individual \ var. + family \ var.)$ 

- assess interaction between level effects (*e.g.*, individual outcome influenced by family SES for males, not females)
- Responses are not independent individuals within clusters share influencing factors

 $\Rightarrow$  Multilevel analysis - another example of *Golden Rule of Statistics*: "one person's error term is another person's (or many persons') career"

		cluster va	riables	subject v	varia	bles
cluster	subject	tx group	size	outcome	sex	age
1	1				•	•
	:				•	•
	$n_1$				•	•
2	1					•
	÷				•	•
	$n_2$	•	•	•	•	•
	1					
	•	•	•	•		•
	$n_{.}$		•		•	•
N	1	·				•
	•					•
	$n_N$		•			•

 $i = 1 \dots N$  clusters

 $j = 1 \dots n_i$  subjects in cluster i

		time- $inv$	ariant	variables	time-varyin	$g \ variables$
subject	time	tx group	sex	age	outcome	dose
1	1		•			•
	÷	•	•	•		•
	$n_1$					
2	1		•			
	:		•		•	
	$n_2$					
•	1					
	:		•			•
	$n_{\cdot}$					
N	1		•			
	:		•	•		•
	$n_N$	•	•	•		•

 $j = 1 \dots n_i$  timepoints for subject i

# Multilevel models aka

- $\bullet$  random-effects models
- $\bullet$  random-coefficient models
- $\bullet$  mixed-effects models
- $\bullet$  hierarchical linear models

# Useful for analyzing

- Clustered data
  - subjects (level-1) within clusters (level-2)
    \* e.g., clinics, hospitals, families, worksites, schools, classrooms, city wards
- Longitudinal data
  - repeated obs. (level-1) within subjects (level-2)

#### General (2-level) Model for Clustered Data

Consider the model with p covariates for the  $n_i \times 1$  response vector  $\boldsymbol{y}$  for cluster i (i = 1, 2, ..., N):

$$oldsymbol{y}_i = oldsymbol{X}_ioldsymbol{eta} + arepsilon_i + oldsymbol{arepsilon}_i$$

 $\begin{aligned} \boldsymbol{y}_i &= n_i \times 1 \text{ vector of responses for cluster } i \\ \boldsymbol{X}_i &= n_i \times (p+1) \text{ covariate matrix} \\ \boldsymbol{\beta} &= (p+1) \times 1 \text{ vector of regression coefficients} \\ v_i &= \text{cluster effects} \sim \mathcal{NID}(0, \sigma_v^2) \\ \boldsymbol{\varepsilon}_i &= n_i \times 1 \text{ vector of residuals} \sim \mathcal{NID}(0, \sigma^2 \boldsymbol{I}_{n_i}) \end{aligned}$ 

- as cluster subscript i is present for  $\boldsymbol{y}$  and  $\boldsymbol{X}$ , cluster sample size can vary
- $\bullet$  the covariate matrix  ${\boldsymbol X}$  can include
  - covariates measured at subject-level
  - $-\operatorname{covariates}$  measured at cluster-level
  - cross-level interactions
- $\bullet$  the total number of covariates = p
- the number of columns in X is p + 1 to include intercept (first column of X consists only of ones)

- $v_i$  random parameter distributed  $\mathcal{NID}(0, \sigma_v^2)$ 
  - distinguishes model from ususal (fixed-effects) multiple regression model
  - represent effect of subject clustering (one for every cluster)
  - if subject clustering has little effect
    - estimates of  $v_i \approx 0$
    - $-\sigma_v^2$  will approach 0
  - if subject clustering has strong effect

- estimates of 
$$v_i \neq 0$$

 $-\sigma_v^2$  will increase from 0

$$\boldsymbol{y}_i \sim \mathcal{NID}(\boldsymbol{X}_i \boldsymbol{\beta}, \sigma_v^2 \boldsymbol{1}_i \boldsymbol{1}'_i + \sigma^2 \boldsymbol{I}_{n_i})$$

- usual mean from multiple regression model
- var-covar structure accounts for clustering

- within a cluster, variance =  $\sigma^2 + \sigma_v^2$  and covariance =  $\sigma_v^2$ 

- "compound symmetry" structure

- ratio of the cluster variance  $\sigma_v^2$  to the total variance  $\sigma^2 + \sigma_v^2$  is the *intraclass correlation* 

# Intra-"class" correlation $r = \sigma_v^2/(\sigma_v^2 + \sigma^2)$

- "class" is bad term, since in education "class" has meaning
- Goldstein suggests "intra-unit" correlation, replacing "unit" with appropriate term (clinic, school, family, firm *etc.*, )
- takes on values between 0 (when  $\sigma_v^2 = 0$ ) and 1 (when  $\sigma^2 = 0$ )
- degree of similarity of measurements within a cluster
- ratio of variability attributable to cluster over total variability
- proportion of total (unexplained) variability of  $y_{ij}$  which is accounted for the clusters
- tends larger for smaller clusters (Kish, 1965; Donner, 1982)
  - -0.05 to 0.12 for spouse pairs, 0.0016 to 0.0126 for physician practices, 0.0005 to 0.0085 for counties
- can change depending on the dependent variable

Anorexic Women Study (Casper) - 63 sisters in 26 families Maximum Likelihood (ML) estimates

	Height	Psych Factor	BMI
intercept	64.166	0.568	0.352
family variance	2.743	0.031	0.000
residual variance	2.895	0.055	0.005
intra-family correlation	0.487	0.362	0.000
descriptive statistics			
mean	64.16	0.57	0.35
variance	5.66	0.084	0.005

# Random-effects Regression Models for Clustered Data: With an Example from Smoking Prevention Research

Hedeker, Gibbons, and Flay

Journal of Consulting and Clinical Psychology, 1994, 62:757-765

The Television School and Family Smoking Prevention and Cessation Project (Flay, *et al.*, 1988); a subsample of this project was chosen with the characteristics:

- *sample* 1600 7th-graders 135 classrooms 28 LA schools
  - between 1 to 13 classrooms per school
  - between 2 to 28 students per classroom
- *outcome* knowledge of the effects of tobacco use
- *timing* students tested at pre and post-intervention
- $\bullet \ design$  schools randomized to
  - -a social-resistance classroom curriculum (CC)
  - -a media (television) intervention (TV)
  - $-\operatorname{CC}$  combined with TV
  - $-\,\mathrm{a}$  no-treatment control group

Tobacco and Health Knowledge Scale							
Subgroup Descript	ive Statist	ics at Prete	est and Pos	st-Intervention			
	CC	= no	CC =	= yes			
	TV = no	TV = yes	TV = no	TV = yes			
$\overline{n}$	421	416	380	383			
Pretest mean	2.152	2.087	2.050	1.979			
sd	1.182	1.288	1.285	1.286			
Post-Int mean	2.361	2.539	2.968	2.823			
sd	1.296	1.437	1.405	1.312			
Difference	0.209	0.452	0.918	0.844			

Within-Cluster / Between-Cluster representation <u>Within-clusters model</u> - level 1  $(j = 1, ..., n_i)$ 

 $PostTHKS_{ij} = b_{0i} + [b_{1i}PreTHKS_{ij}] + \varepsilon_{ij}$ 

<u>Between-clusters model</u> - level 2 (i = 1, ..., N)

$$b_{0i} = \beta_0 + [\beta_2 C C_i] + v_{0i}$$
$$b_{1i} = \beta_1$$

 $\varepsilon_{ij} \sim NID(0, \sigma^2)$  level-1 residuals  $v_{0i} \sim NID(0, \sigma_v^2)$  level-2 residuals TVSFP Study (Flay *et al.*, 1988): Tobacco and Health Knowledge *Posttest* Scores 1600 students in 135 classrooms in 28 schools: ML estimates (and standard errors)

	student	ts in clas	ssrooms	students in schools			
Intercept	2.618	2.007	1.757	2.682	2.047	1.800	
	(0.052)	(0.072)	(0.080)	(0.078)	(0.089)	(0.092)	
Pretest score		0.302	0.310		0.303	0.310	
		(0.026)	(0.026)		(0.026)	(0.026)	
Classroom			0.497			0.470	
curriculum			(0.086)			(0.106)	
Cluster var	0.194	0.157	0.096	0.130	0.101	0.044	
	(0.043)	(0.037)	(0.029)	(0.045)	(0.036)	(0.020)	
Residual var	1.725	1.601	1.601	1.788	1.653	1.653	
	(0.064)	(0.060)	(0.059)	(0.064)	(0.059)	(0.059)	
ICC	0.101	0.090	0.057	0.068	0.057	0.026	
$\log L$	-2760.9	-2696.4	-2681.3	-2756.8	-2692.0	-2684.7	
$\chi^2_1$		129.0	30.2		129.6	14.6	

Within-Cluster / Between-Cluster representation

<u>Within-clusters model</u> - level 1  $(j = 1, ..., n_i)$ 

 $PostTHKS_{ij} = b_{0i} + \varepsilon_{ij}$ 

<u>Between-clusters model</u> - level 2 (i = 1, ..., N)  $b_{0i} = \beta_0 + \beta_1 C C_i + \beta_2 T V_i + \beta_3 (C C_i \times T V_i) + v_{0i}$   $\varepsilon_{ij} \sim NID(0, \sigma^2)$  level-1 residuals  $v_{0i} \sim NID(0, \sigma^2_v)$  level-2 residuals

• If cluster effect is completely explained by condition, then

 $-\upsilon_{0i} = 0$  for all i (thus  $\sigma_{\upsilon}^2 = 0$ )

- model is same as ordinary regression (individual-level analysis)

- If  $n_i = n$  for all clusters (and no level-1 covariates), then
  - model is same as ordinary regression of cluster means (cluster-level analysis)

THKS post-inter-	THKS post-intervention scores - Regression estimates (se)							
	Ordinary	Regression	Multilevel Model					
	Class-level	Student-level	Students in classes					
Intercept	2.342	2.361	2.341					
	(.117)	(.066)	(.092)					
classroom	.507	.607	.589					
curriculum (CC)	(.166)	(.096)	(.133)					
television	082	.177	.120					
$(\mathrm{TV})$	(.150)	(.094)	(.131)					
interaction	.011	323	247					
(CC by TV)	(.236)	(.137)	(.189)					
residual variance	.468	1.860	1.727					
			(.064)					
class variance			.134					
			(.037)					
$p < .05$ $\mathbf{p} <$	.01		ICC = .072					

cian actimates (co) TUKS post intervention of 

#### Within-Cluster / Between-Cluster representation

<u>Within-clusters model</u> - level 1  $(j = 1, ..., n_i)$ 

$$PostTHKS_{ij} = b_{0i} + b_{1i}PreTHKS_{ij} + \varepsilon_{ij}$$

<u>Between-clusters model</u> - level 2 (i = 1, ..., N)

$$b_{0i} = \beta_0 + \beta_2 C C_i + \beta_3 T V_i + \beta_4 (C C_i \times T V_i) + \upsilon_{0i}$$
  
$$b_{1i} = \beta_1$$

 $\varepsilon_{ij} \sim NID(0, \sigma^2)$  level-1 residuals  $v_{0i} \sim NID(0, \sigma_v^2)$  level-2 residuals

#### 3-level model representation

Within-schools, within-classrooms, between-subjects model - level 1  

$$(k = 1, ..., n_{ij})$$

$$PostTHKS_{ijk} = b_{0ij} + b_{1ij}PreTHKS_{ijk} + \varepsilon_{ijk}$$

Within-schools, between-classrooms model - level 2  $(j = 1, ..., n_i)$ 

$$egin{array}{rcl} b_{0ij} &=& b_{0i} \ + \ arpi_{0ij} \ b_{1ij} \ = \ b_{1i} \end{array}$$

<u>Between-schools model</u> - level 3 (i = 1, ..., N)

$$b_{0i} = \beta_0 + \beta_2 C C_i + \beta_3 T V_i + \beta_4 (C C_i \times T V_i) + v_{0i}$$
  
$$b_{1i} = \beta_1$$

$$\varepsilon_{ijk} \sim NID(0, \sigma^2)$$
 level-1 residuals  
 $\upsilon_{0ij} \sim NID(0, \sigma^2_{\upsilon_{(2)}})$  level-2 residuals  
 $\upsilon_{0i} \sim NID(0, \sigma^2_{\upsilon_{(3)}})$  level-3 residuals

	Ordinar	y Re	gression Models	5			Multilevel Mode	els		
	Class-level		Student-level		Stu in classes		Stu in schools		Three-level	
Intercept	1.3087	***	1.6613	***	1.6776	***	1.6952	***	1.6970	***
	(0.229)		(0.084)		(0.099)		(0.115)		(0.117)	
pretest	0.4962	***	0.3252	***	0.3116	***	0.3103	***	0.3072	***
THKS	(0.097)		(0.026)		(0.026)		(0.026)		(0.026)	
classroom	0.5749	***	0.6406	***	0.6330	***	0.6601	***	0.6392	***
curriculum	(0.153)		(0.092)		(0.119)		(0.144)		(0.147)	
television	-0.0150		0.1987	**	0.1597		0.2023		0.1781	
	(0.150)		(0.090)		(0.117)		(0.140)		(0.144)	
interaction	-0.0485		-0.3216	**	-0.2747		-0.3696	*	-0.3204	
	(0.216)		(0.130)		(0.168)		(0.201)		(0.206)	
error	0.3924		1.6929		1.6030	***	1.6522	***	1.6020	***
variance					(0.059)		(0.059)		(0.059)	
class					0.0870	***			0.0636	**
variance					(0.028)				(0.028)	
school							0.0372	**	0.0258	
variance							(0.018)		(0.020)	

#### THKS Post-Intervention Scores - Regression Estimates (se)

\*\*\*p < 0.01 \*\*p < 0.05 \*p < 0.10

# Results

- conclusions about CC by TV interaction differ
  - non-significant by class-level analysis, significant by student-level analysis, marginally significant by multilevel
- student-level results close to multilevel, but estimates are more similar than standard errors  $\rightarrow$  underestimation of standard errors by ordinary regression analysis is expected since assumption of independence of observations is violated
- students more homogeneous within classrooms than schools
  - $-\,\mathrm{students}$  within classrooms model, r=0.052
  - students within schools model, r = 0.022
- 3-level model close to students within classrooms model
  - based on 3-level model, classroom and school effects accounted for 3.8% and 1.5% of total variance, respectively

#### **3-level ICCs**

From the three-level model: error var = 1.6020, class var = 0.0636, school var = 0.0258

Similarity of students within the same school

$$ICC = \frac{0.0258}{1.6020 + 0.0636 + 0.0258} = .0153$$

Similarity of students within the same classrooms (and schools)  $ICC = \frac{0.0636 + 0.0258}{1.6020 + 0.0636 + 0.0258} = .0529$ 

Similarity of classes within the same school  $ICC = \frac{0.0258}{0.0636 + 0.0258} = .289$  Explained Variance (Hox, Multilevel Analysis, 2002)

level-1 
$$R_1^2 = 1 - \frac{\hat{\sigma}_p^2}{\hat{\sigma}_0^2}$$
 level-2  $R_2^2 = 1 - \frac{\hat{\sigma}_{v_p}^2}{\hat{\sigma}_{v_0}^2}$ 

subscript 0 refers to a model with no covariates (*i.e.*, null model), subscript p refers to a model with p covariates (*i.e.*, full model)

e.g., students in classrooms models

	models					
level	variance	null	full	$R^2$		
1 (students)	$\hat{\sigma}^2$	1.725	1.603	.071		
2 (classrooms)	$\hat{\sigma}_{v}^{2}$	.194	.087	.552		

#### Explained Variance: 3-level model

$$R_1^2 = 1 - \frac{\hat{\sigma}_p^2}{\hat{\sigma}_0^2} \qquad R_2^2 = 1 - \frac{\hat{\sigma}_{\upsilon(2)p}^2}{\hat{\sigma}_{\upsilon(2)0}^2} \qquad R_3^2 = 1 - \frac{\hat{\sigma}_{\upsilon(3)p}^2}{\hat{\sigma}_{\upsilon(3)0}^2}$$

subscript 0 refers to a model with no covariates (*i.e.*, null model), subscript p refers to a model with p covariates (*i.e.*, full model)

e.g., students in classrooms in schools models

level	variance	null	full	$R^2$
1 (students)	$\hat{\sigma}^2$	1.724	1.602	.071
2  (classrooms)	$\hat{\sigma}^2_{\upsilon_{(2)}}$	.085	.064	.247
3  (schools)	$\hat{\sigma}^{2^{(-)}}_{arvarvar}$	.110	.026	.764

# Likelihood-ratio tests:

suppose Model I is nested within Model II

$$2 \times \log(L_{\mathrm{II}} / L_{\mathrm{I}}) = 2 \times (\log L_{\mathrm{II}} - \log L_{\mathrm{I}}) \sim \chi_q^2$$

where q = number of additional parameters in Model II

 $-2 \log L$  is called the *deviance* (the higher the deviance the poorer the model fit)

$$D_{\mathrm{I}} - D_{\mathrm{II}} \sim \chi_q^2$$

to evaluate the null hypothesis that the additional parameters in Model II jointly equal 0

#### Comparison of models using LR tests

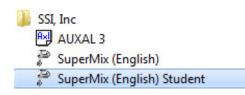
			-			halved
Model	deviance	CM	$\chi^2$	df	p <	p <
1. student-level	5377.90					
2a. students in classes	5359.96	1	17.94	1	.001	.001
2b. students in schools	5366.01	1	11.89	1	.001	.001
3. three-level	5357.36	1 2a	20.54 2.60		.001 .107	.001 .053

LR tests with halved *p*-values (akin to one-tailed *p*-values) for tests of variance parameters is recommended (Snijders & Bosker, *Multilevel Analysis, 2nd edition*, 2012, pps. 98-101)

### Supermix for Mixed Models

- Free student and 15-day trial editions http://www.ssicentral.com/supermix/downloads.html
- Datasets and examples http://www.ssicentral.com/supermix/examples.html
- Manual and documentation in PDF form http://www.ssicentral.com/supermix/resources.html

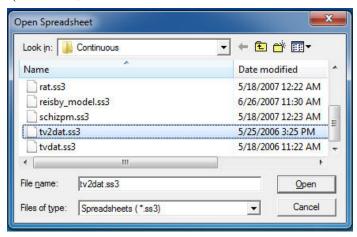
• Under SSI, Inc > "SuperMix (English)" or 'SuperMix (English) Student"



• Under "File" click on "Open Spreadsheet"

File Help	
New Spreadsheet	Ctrl+N
Open Spreadsheet	Ctrl+O
Import Data File	Ctrl+I
Exit	

• Open C:\SuperMixEn Examples\Workshop\Continuous\tv2dat.ss3 (or C:\SuperMixEn Student Examples\Workshop\Continuous\tv2dat.ss3)



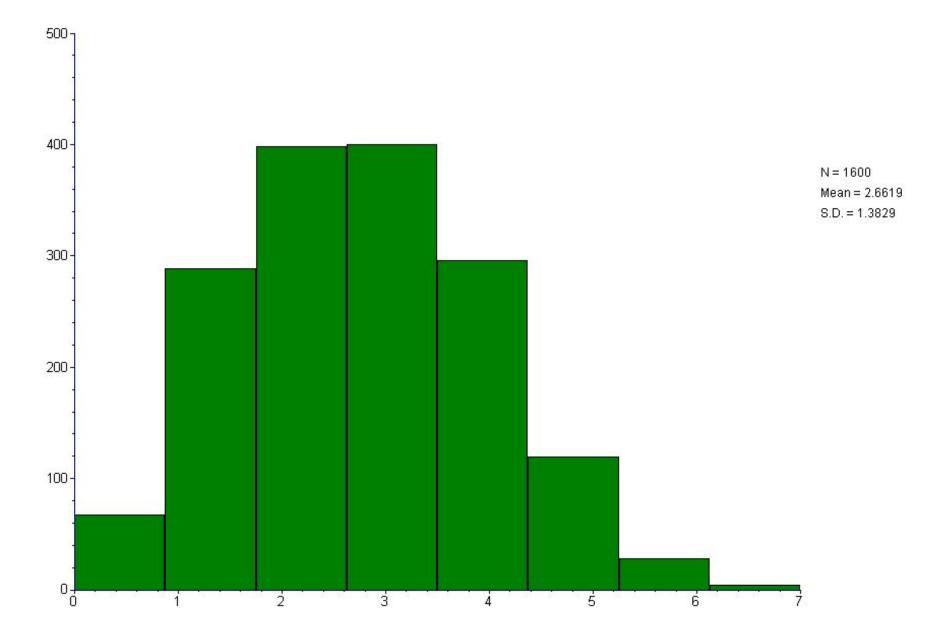
#### c:\SuperMixEn Examples\Workshop\Continuous\tv2dat.ss3

Super	Mix								
Edi	t Window	Help							
tv2d	at.ss3								
1.973 PX	611030976 								
-		(m) (m)		(D) 1 ( )		(T) 00	(0) TH	410 00 TH	
	(A)_School	(- <u>)</u>	(C)_THKSco		(E)_PreTHK	(F)_CC	(G)_TV	(H)_CCxTV	
1	403.00	403101.00	3.00	1.00 1.00	2.00	1.00	0.00 0.00	0.00	
2	403.00	403101.00	3.00	1.00	4.00	1.00	0.00	0.00	
3	403.00	403101.00	4.00	1.00	3.00	1.00	0.00	0.00	
4 5	403.00	403101.00	4.00	1.00	3.00	1.00	0.00	0.00	
6	403.00	403101.00	3.00	1.00	4.00	1.00	0.00	0.00	
7	403.00	403101.00	2.00	1.00	2.00	1.00	0.00	0.00	
8	403.00	403101.00	4.00	1.00	4.00	1.00	0.00	0.00	
9	403.00	403101.00	5.00	1.00	5.00	1.00	0.00	0.00	
10	403.00	403101.00	4.00	1.00	3.00	1.00	0.00	0.00	
11	403.00	403101.00	3.00	1.00	3.00	1.00	0.00	0.00	
12	403.00	403101.00	4.00	1.00	3.00	1.00	0.00	0.00	
13	403.00	403101.00	3.00	1.00	1.00	1.00	0.00	0.00	
14	403.00	403101.00	4.00	1.00	2.00	1.00	0.00	0.00	
15	403.00	403101.00	2.00	1.00	2.00	1.00	0.00	0.00	
16	403.00	403101.00	4.00	1.00	1.00	1.00	0.00	0.00	
17	403.00	403101.00	4.00	1.00	4.00	1.00	0.00	0.00	
18	403.00	403101.00	3.00	1.00	3.00	1.00	0.00	0.00	
19	403.00	403101.00	3.00	1.00	0.00	1.00	0.00	0.00	
20	403.00	403101.00	4.00	1.00	3.00	1.00	0.00	0.00	
21	403.00	403102.00	2.00	1.00	0.00	1.00	0.00	0.00	
22	403.00	403102.00	5.00	1.00	1.00	1.00	0.00	0.00	
23	403.00	403102.00	3.00	1.00	5.00	1.00	0.00	0.00	
24	404.00	404101.00	3.00	1.00	1.00	1.00	1.00	1.00	
25	404.00	404101.00	4.00	1.00	2.00	1.00	1.00	1.00	
26	404.00	404101.00	2.00	1.00	4.00	1.00	1.00	1.00	
27	404.00	404101.00	3.00	1.00	3.00	1.00	1.00	1.00	
28	404.00	404101.00	2.00	1.00	1.00	1.00	1.00	1.00	
29	404.00	404101.00	1.00	1.00	1.00	1.00	1.00	1.00	
30	404.00	404101.00	3.00	1.00	2.00	1.00	1.00	1.00	
31	404.00	404101.00	5.00	1.00	0.00	1.00	1.00	1.00	
32	404.00	404101.00	2.00	1.00	1.00	1.00	1.00	1.00	
33	404.00	404101.00	3.00	1.00	2.00	1.00	1.00	1.00	
34	404.00	404101.00	4.00	1.00	2.00	1.00	1.00	1.00	
35	404.00	404102.00	3.00	1.00	1.00	1.00	1.00	1.00	
36	404.00	404102.00	1.00	1.00	1.00	1.00	1.00	1.00	1
37	404.00	404102.00	0.00	1.00	0.00	1.00	1.00	1.00	l
<u> </u>		101102-00	1 00	1 00	1 00	1 00	1 00	1.00	

#### Select "File" > "Data-based Graphs" > "Univariate"

ist of Variables Name	Plot	
School		
Class	<u> </u>	
THKScore	<b>v</b>	
ntropt	Ē	
PreTHKS	Г	
C	Г	
ſV		
CCxTV		
Bar Chart Pie Chart 3D Pie Chart		-
	intervals: 8	

#### Histogram of THKScore



#### Under "File" click on "Open Existing Model Setup"

File	Edit Window Help	
	New Project	Ctrl+N
	Import Data File	Ctrl+I
	Close	
	New Model Setup	Ctrl+W
	Open Existing Model Setup	Ctrl+E
	New Syntax File	
	Open Syntax File	
	Open Text File	
	Data-based Graphs	
	Open Graph	Ctrl+G
	Save	Ctrl+S
	Save As	
	Exit	

Open C:\SuperMixEn Examples\Workshop\Continuous\TVC.mum (or C:\SuperMixEn Student Examples\Workshop\Continuous\TVC.mum)

Organize 👻 New folde					) III 🔹 🗖 🌘
🛧 Favorites 🔺	Name	Date modified	Туре	Size	
📃 Desktop	📄 reisby.mum	5/24/2006 12:02 PM	MUM File	4 KB	
🚺 Downloads	reisby2.mum	7/25/2007 12:39 PM	MUM File	4 KB	
🔛 Recent Places	schizpma.mum	5/24/2006 10:11 AM	MUM File	4 KB	
	📄 schizpmb.mum	5/24/2006 10:15 AM	MUM File	4 KB	
🥽 Libraries	schizpmc.mum	5/25/2006 3:19 PM	MUM File	4 KB	
Documents	TVBC.mum	5/15/2006 10:40 AM	MUM File	4 KB	
🎝 Music 😑	TVBSC.mum	5/15/2006 10:51 AM	MUM File	4 KB	
E Pictures	TVC.mum	5/24/2006 10:21 AM	MUM File	4 KB	
😸 Videos	TVS.mum	5/24/2006 10:23 AM	MUM File	4 KB	
	TVSC.mum	5/24/2006 10:23 AM	MUM File	4 KB	
Computer					
🚰 Local Disk (C:)					
🚽 groups (\\phsfile					
🚽 DHedeker (\\phs					
📬 Network 👻					
	me: TVC.mum	S		✓ Mixed Up	Models (*.mum)

Model Setup: TVC.mum					
Configuration	tarting Values   <u>P</u> al	terns   <u>A</u> dvance	ed   <u>L</u> inear Transforms	1	
Title 1: TVSFP data - Cont	inuous Outcome				
Title 2: Students in Classro	oms				
Dependent Variable Type:	continuous		Level-2 IDs:	Class	
Dependent Variable:	THKScore	•	Level-3 IDs:		
		٧	/rite Bayes Estimates:	no	
		С	onvergence Criterion:	0.0001	
			Number of Iterations:	100	
Missing Values Present:	false		Generate Table of	Means: no	
			Output Type:	standard	
Use the arrow ke	eys or click on the c	lesired tab to sel	ect the category of inte	erest for the mo	del.

Available	E 2	Explanatory Variables	E	L-2 Random Effects 2
School				
Class Turko				
THKScore		TV	1	
Intropt		CCxTV		
PreTHKS				
CC				
TV				
CCxTV				
				Include Intercept

```
- 0
                                                                       X
 SuperMix - [TVC.out]
🚰 File Analysis Window Help
                                                                    - 8 ×
                                                                       ٠4.
       0-----0
          SuperMix Version 2 for Continuous Outcomes
            Copyright 2005-2014
            Scientific Software International, Inc.
            Phone: (847)675-0720
            Fax:
                   (847) 675-2140
            Website: www.ssicentral.com
            Support: smix@ssicentral.com
          Date of analysis: December 31, 2014
          Time of analysis: 12H03:24
       0-----0
 Model specifications are as follows:
 Model=Continuous;
 Options Output=standard Converge=0.0001 Maxiter=100 Bayes=No;
 Link=identity;
 Distribution=nor;
 Varnames= School Class THKScore Intrcpt PreTHKS CC TV CCxTV intercept;
 Title1=TVSFP data - Continuous Outcome;
 Title2=Students in Classrooms;
 DataFile=C:\SuperMixEn Examples\Workshop\Continuous\TVC.dat;
 Level2ID= Class;
 Dependent= THKScore;
 Predictors= intercept PreTHKS CC TV CCxTV;
 L1Random= intercept;
 L2Random= intercept;
 FixPatType=Free;
 Cov2PatType=Correlated;
 AutoCor=None;
 Numbers of observations
 -----
 Level 2 observations =
                      135
 Level 1 observations =
                      1600
 N2
      2
            1
                 2
                        3
                                4
                                      5
                                             6
                                                  7
                                                          8
                                            26 11 10
 N1
      23
            20
                 3
                        11
                                9
                                      5
                                                      16
 N2
            9 10
                      11
                               12
                                     13 14 15
      -
 N1
            15 12
                      12
                                      21 10 17
                                                      19
      -
                               10
 N2
      =
            17 18 19
                               20
                                     21 22 23
                                                         24
 N1
             2
                 4 21
                              16
                                     15
                                            13 2
                                                         14
      -
                                                                       ÷
 Save As...
           Close
```

#### SuperMix - [TVC.out]

🊰 File Analysis Window Help

#### - 0 X

- 8 ×

.

E

÷

Descriptive statistics for all variables -----

		Minimum	Maximum	Mean	Stand. Dev
Dependent					
THKScore		0.0000	7.0000	2.6619	1.3829
Random-Effect	s				
	100				
intercept	(2)	1.0000	1.0000	1.0000	0.0000
intercept	(1)	1.0000	1.0000	1.0000	0.000
Fixed Regress	sor(s)				
TIACG REGIESS					
		1 0000	1 0000	1 0000	0.0000
intercept		1.0000	1.0000	1.0000	0.0000
intercept PreTHKS		0.0000	6.0000	2.0694	1.2602
intercept					

0= \_\_\_\_\_ =0 | TVSFP data - Continuous Outcome | | Students in Classrooms 1 0=================================

Parameter starting values \_\_\_\_\_

Fixed regressor(s) ------

Variable	Estimate	Std.Err.	Z-value	p-value
intercept	1.66126	0.08436	19.69278	0.0000
PreTHKS	0.32518	0.02585	12.57835	0.0000
CC	0.64055	0.09210	6.95462	0.00000
TV	0.19871	0.08996	2.20876	0.02719
CCxTV	-0.32162	0.13025	-2.46922	0.01354
Log Likelihood	=	-2819.4696		
Number of free parameter	- =	7		

Nu er of free parameters =

Save As...

File Analysis Wind	ow Help					- 5
Convergence attain	ed in 6 iteratio	ons				*
)   TVSFP data - Con   Students in Clas	srooms					
Maximum likelihood						
Fixed regressor(s)						
Variable		Std.Err.		p-value		
intercept	1.67763	0.09881	16.97799			
PreTHKS	0.31157	0.02580				
CC	0.63298	0.11863			104.0	
TV	0.15966	0.11670				
CCXTV	-0.27469	0.16780				
Log Likelihood	= -	-2679.9821				
	(Deviance) =					
Akaike's Informatio	on Criterion =	5373.9641				
Schwarz's Bayesian	Criterion =	5394.3010				
Number of free par	ameters =	7				
						E
Jariance/covariance	e components					
Level 2		Estimate	Std.Err.	Z-value	p-value	
intercept /	intercept	0.08697	0.02765	3.14588	0.00166	
Level 1		Estimate	Std.Err.	Z-value	p-value	
			0.05000	27 20007	0.00000	
intercept /	intercept	1.60301	0.05893	21.20001		

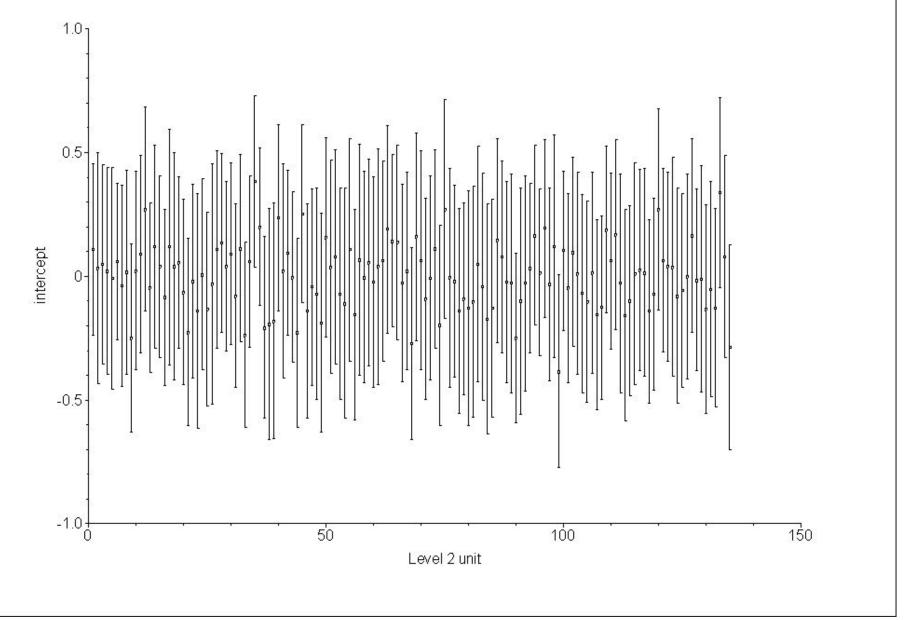
## **Empirical Bayes Estimates of Random Effects** Select "Analysis" > "View Level-2 Bayes Results"

File	Analysis Win	dow He	lp			- 8
	403101.00	1	0.16497	0.41712E-01	intercept	
	403102.00	1	0.55937E-01	0.74801E-01		
	404101.00	1	0.65535E-01	0.54468E-01	intercept	
	404102.00	1	0.31999E-01	0.58439E-01	E 1.5.5	
	404103.00	1	-0.67832E-02	0.68416E-01	Contraction of the second second	
	193101.00	1	0.88539E-01	0.36079E-01		
	194101.00	1	-0.25504E-01	0.54468E-01		
	194102.00	1	0.39498E-01	0.56384E-01		
	194103.00	1	-0.27520	0.47951E-01		1
	194104.00	1	0.49997E-01	0.52678E-01		
	194105.00	1	0.13191	0.52678E-01		
	194106.00	1	0.35606	0.56384E-01		
	196101.00	1	-0.32519E-01	0.40654E-01		
	196102.00	1	0.16400	0.56384E-01		
	197101.00	1	0.74533E-01	0.45244E-01		
	197102.00	1	-0.70779E-01	0.42827E-01		
	197103.00	1	0.16315	0.78463E-01		
	197104.00	1	0.63311E-01	0.71466E-01		3
	198101.00	1	0.96253E-02	0.40654E-01		
	198102.00	1	-0.12286	0.46558E-01		
	198103.00	1	-0.31685	0.47951E-01	Construction of the second	
	199101.00	1	-0.45426E-01	0.51002E-01		
	199102.00	1	-0.19191	0.78463E-01	intercept	
	199103.00	1	-0.13753E-01	0.49429E-01		
	199104.00	1	-0.18261	0.51002E-01		
	199105.00	1	-0.43093E-01	0.82501E-01		
	199106.00	1	0.11204	0.52678E-01		
	401101.00	1	0.18902	0.44002E-01		
	401102.00	1	0.77807E-01	0.40654E-01		
	402101.00	1	0.11524	0.45244E-01		
	402102.00	1	-0.86936E-01	0.46558E-01		
	405101.00	1	0.13763	0.47951E-01	C 1. 3.5	
	405102.00	1	-0.28110	0.46558E-01		
	405103.00	1	0.70330E-01	0.40654E-01	A CONTRACTOR OF	
	407101.00	1	0.54723	0.40654E-01	10 10 4 10 10 12 10 10 A	
	407102.00	1	0.33838	0.35285E-01		
	407103.00	1	-0.15894	0.45244E-01		
	408101.00	1	-0.25617	0.74801E-01		
	408102.00	1	-0.24145	0.78463E-01	intercept	
	408103.00	1	0.26948	0.47951E-01		
	408104.00	1	0.18236E-01	0.63036E-01		
	409101.00	1	0.14360	0.37780E-01		
	409102.00	1	0.28120E-01	0.39649E-01	intercept	
	409103.00		-0.24724	0.47951E-01	C	
		1	0.32688	0.42827E-01		
	410101.00	1	-0.22793 -0.12783	0.63036E-01		
	410102.00	1	-0.12783	0.52678E-01	intercept	

ID, random effect number, estimate, variance, name

## Select "File" > "Model-based Graphs" > "Confidence Intervals"

	Name	Predictor	Group	Mark
1	intercept		-	



#### 95% Confidence Intervals - Classroom effects (in order)

### Under "File" click on "Open Existing Model Setup"

File	Edit Window Help	
	New Project	Ctrl+N
	Import Data File	Ctrl+I
	Close	
	New Model Setup	Ctrl+W
	Open Existing Model Setup	Ctrl+E
	New Syntax File	
	Open Syntax File	
	Open Text File	
	Data-based Graphs	
	Open Graph	Ctrl+G
	Save	Ctrl+S
	Save As	
	Exit	

Open C:\SuperMixEn Examples\Workshop\Continuous\TVSC.mum (or C:\SuperMixEn Student Examples\Workshop\Continuous\TVSC.mum)

Organize 👻 New fol	der				i 🕶 🗖
🖈 Favorites 🤺	Name	Date modified	Туре	Size	
E Desktop	reisby.mum	5/24/2006 12:02 PM	MUM File	4 KB	
📕 Downloads	reisby2.mum	7/25/2007 12:39 PM	MUM File	4 KB	
💹 Recent Places	schizpma.mum	5/24/2006 10:11 AM	MUM File	4 KB	
	schizpmb.mum	5/24/2006 10:15 AM	MUM File	4 KB	
📜 Libraries	schizpmc.mum	5/25/2006 3:19 PM	MUM File	4 KB	
Documents	TVBC.mum	5/15/2006 10:40 AM	MUM File	4 KB	
🌙 Music 🛛 ≣	TVBSC.mum	5/15/2006 10:51 AM	MUM File	4 KB	
E Pictures	TVC.mum	5/24/2006 10:21 AM	MUM File	4 KB	
🛃 Videos	TVS.mum	5/24/2006 10:23 AM	MUM File	4 KB	
	TVSC.mum	5/24/2006 10:23 AM	MUM File	4 KB	
Somputer					
Local Disk (C:)					
🖵 groups (\\phsfile					
🖵 DHedeker (\\phs					
Network					

Model Setup: TVSC.mur	n				•
Configuration	tarting Values   <u>P</u> ati	terns   <u>A</u> dva	nced   Linear Transforms		
Title 1: TVSFP data - Cont	inuous Outcome			n:	
Title 2: Students in School	s and Classrooms				
Dependent Variable Type:	continuous	-	Level-2 IDs:	Class	Ī
Dependent Variable:	THKScore	<u> </u>	Level-3 IDs:	School	Ī
			Write Bayes Estimates:	no	•
			Convergence Criterion:	0.0001	
			Number of Iterations:	100	
Missing Values Present:	false	-	Generate Table of	Means: no	T
					4
			Output Type:	standard	Ţ
Use the arrow ke	eys or click on the d	esired tab to	select the category of inte	erest for the model.	

Available	E 2 3	Explanatory Variables	E	L-2 Random Effects 2
ichool		PreTHKS	1	
llass		CC		
HKScore		TV	2	
ntropt		CCxTV		
'reTHKS				
iC.				
V				
CxTV				
				Include Intercept
				l∙ include intercept
				L-3 Random Effects 3
			N	
		Include Intercept	13	Include Intercept

```
TVSC.out
                                                                 *
       SuperMix Version 2 for Continuous Outcomes
       10
         Copyright 2005-2014
           Scientific Software International, Inc.
          Phone: (847)675-0720
          Fax:
                (847) 675-2140
         Website: www.ssicentral.com
       Support: smix@ssicentral.com
       | Date of analysis: December 31, 2014
         Time of analysis: 13H15:19
       0-----0
 Model specifications are as follows:
 Model=Continuous;
 Options Output=standard Converge=0.0001 Maxiter=100 Bayes=No;
 Link=identity;
 Distribution=nor;
 Varnames= School Class THKScore Intropt PreTHKS CC TV CCxTV intercept;
 Title1=TVSFP data - Continuous Outcome;
 Title2=Students in Schools and Classrooms;
 DataFile=C:\SuperMixEn Examples\Workshop\Continuous\TVSC.dat;
 Level2ID= Class;
 Level3ID= School;
 Dependent= THKScore;
 Predictors= intercept PreTHKS CC TV CCxTV;
 L1Random= intercept;
 L2Random= intercept;
 L3Random= intercept;
 FixPatType=Free;
 Cov2PatType=Correlated;
 Cov3PatType=Correlated;
 AutoCor=None;
 Numbers of observations
 -----
 Level 3 observations =
                     28
 Level 2 observations =
                    135
 Level 1 observations = 1600
 LEVEL3 :
           1 2
                           4 5 6 7 8
                     3
 N2 :
          2 3 1 6 2 4 3 6
 N1 : 23 25 26 70 31 42 52 55
 LEVEL3 :
                                      14
         9
                 10
                     11
                           12
                                   13
                                             15
                                                     16
 N2 :
           2
                 2
                       3
                            3
                                   4 4
                                              4
                                                     2
 N1 : 39
               33
                     52 65
                                 27 80
                                                     18
                                             33
 Save As...
           Close
```

TVSC.out						
						-
Convergence attained i	n 6 iterati					
		=0				
TVSFP data - Continu   Students in Schools						
Students in Schools						
Maximum likelihood est	imates					
Fixed regressor(s)						
Variable	Estimate	Std.Err.	Z-value	p-value		
intercept	1.69700					
PreTHKS	0.30720					
CC TV	0.63919 0.17811					
CCxTV	-0.32042		-1.55910			
COATV	0.02042	0.20001	1.00010	0.1103		
Log Likelihood	=	-2678.6793				
-2 Log Likelihood (Dev						
Akaike's Information C						
Schwarz's Bayesian Cri		5384.0163 8				
Number of free paramet	ers =	•				
						E
Variance/covariance co	mponents					
	1990 - 1990 - 1990 - 1990 - 1990 - 1990 - 1990 - 1990 - 1990 - 1990 - 1990 - 1990 - 1990 - 1990 - 1990 - 1990 -					
Level 3		Estimate	Std.Err.	Z-value	p-value	
					-	
intercept /inte	rcept	0.02575	0.01971	1.30665	0.19133	
		-	<b>a</b> . 1 <b>a</b>			
Level 2		Estimate	Std.Err.	Z-value	p-value	
intercept /inte	rcept	0.06358	0.02767	2.29768		
, 1100	7850000					
Level 1		Estimate	Std.Err.	Z-value	p-value	
intercept /inte	rcept	1.60201	0.05889	27.20544	0.00000	
to devidence of the second						
						-

## **Empirical Bayes Estimates of Random Effects** Select "Analysis" > "View Level-2 Bayes Results"

403.00	403101.00	20	1	0.10870	0.31321E-01	intercept
403.00	403102.00	3	1	0.34552E-01	0.56582E-01	intercept
404.00	404101.00	11	1	0.48502E-01	0.42427E-01	1987 - Yan
404.00	404102.00	9	1	0.21327E-01	0.45475E-01	intercept
404.00	404103.00	5	1	-0.81081E-02	0.52513E-01	intercept
193.00	193101.00	26	1	0.59534E-01	0.25784E-01	intercept
194.00	194101.00	11	1	-0.38507E-01	0.42915E-01	intercept
194.00	194102.00	10	1	0.16603E-01	0.44342E-01	intercept
194.00	194103.00	15	1	-0.25007	0.37826E-01	intercept
194.00	194104.00	12	1	0.23386E-01	0.41554E-01	intercept
194.00	194105.00	12	1	0.89461E-01	0.41554E-01	intercept
194.00	194106.00	10	1	0.27233	0.44342E-01	intercept
196.00	196101.00	21	1	-0.47277E-01	0.30584E-01	intercept
196.00	196102.00	10	1	0.11939	0.43919E-01	intercept
197.00	197101.00	17	1	0.40053E-01	0.35029E-01	intercept
197.00	197102.00	19	1	-0.84134E-01	0.32904E-01	intercept
197.00	197102.00	2	1	0.11790	0.58812E-01	intercept
197.00	197104.00	4	1	0.40540E-01	0.54535E-01	intercept
198.00	198101.00	21	1	0.56531E-01	0.31114E-01	1997 X H-
198.00	198102.00	16	1	-0.62934E-01	0.36287E-01	intercept
198.00	198103.00	15	1	-0.22461	0.37454E-01	intercept
199.00	199101.00	13	1	-0.19525E-01	0.40075E-01	intercept
199.00	199102.00	2	1	-0.14020	0.58823E-01	199 X 14 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
199.00	199103.00	14	1	0.79596E-02	0.38815E-01	intercept
199.00	199104.00	13	1	-0.13230	0.40075E-01	intercept
199.00	199105.00	1	1	-0.30231E-01	0.61136E-01	intercept
199.00	199106.00	12	1	0.10878	0.41392E-01	1997 - Yan - Carlos -
401.00	401101.00	18	1	0.13415	0.33783E-01	intercept
401.00	401102.00	21	1	0.40196E-01	0.30746E-01	intercept
402.00	402101.00	17	1	0.91264E-01	0.34802E-01	intercept
402.00	402102.00	16	1	-0.78120E-01	0.35947E-01	
405.00	405101.00	15	1	0.11404	0.37454E-01	intercept
405.00	405102.00	16	1	-0.23619	0.36287E-01	intercept
405.00	405103.00	21	1	0.58924E-01	0.31114E-01	intercept
407.00	407101.00	21	1	0.38397	0.31261E-01	1991 X H
407.00	407102.00	27	1	0.19960	0.26267E-01	intercept
407.00	407103.00	17	1	-0.20588	0.35282E-01	intercept
408.00	408101.00	3	1	-0.19175	0.56599E-01	intercept
408.00	408102.00	2	1	-0.17972	0.58804E-01	intercept
408.00	408102.00	15	1	0.23679	0.37123E-01	intercept
408.00	408103.00	15	1	0.23679 0.20771E-01	0.48833E-01	intercept
409.00	409101.00	24	1	0.95309E-01	0.28984E-01	intercept
409.00	409102.00	24	1	-0.27363E-02	0.30676E-01	intercept
409.00	409102.00	15	1	-0.22784	0.37762E-01	intercept

School ID, Class ID, number of observations in the class, random effect number, estimate, variance, name

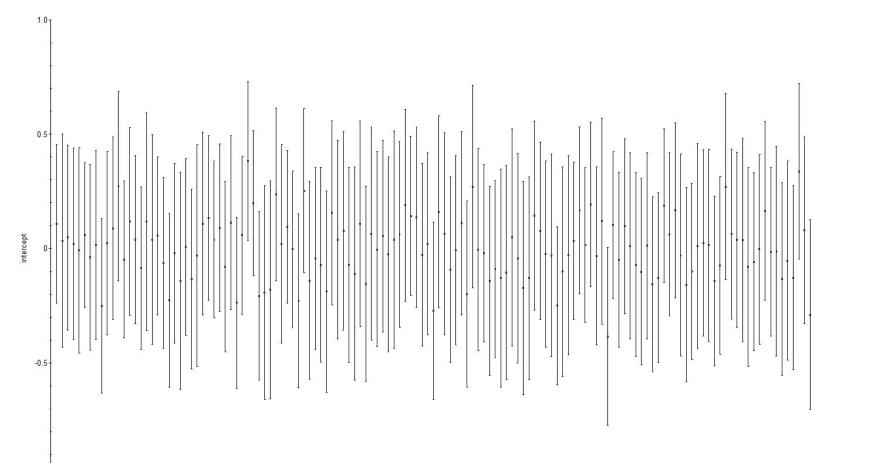
### **Empirical Bayes Estimates of Random Effects** Select "Analysis" > "View Level-3 Bayes Results"

403.00	1	0.58013E-01	0.21068E-01	intercept
404.00	1	0.24995E-01	0.19858E-01	intercept
193.00	1	0.24110E-01	0.21358E-01	intercept
194.00	1	0.45842E-01	0.14584E-01	intercept
196.00	1	0.29206E-01	0.19821E-01	intercept
197.00	1	0.46312E-01	0.18102E-01	intercept
198.00	1	-0.93556E-01	0.17253E-01	intercept
199.00	1	-0.83225E-01	0.16132E-01	intercept
401.00	1	0.70605E-01	0.19034E-01	intercept
402.00	1	0.53231E-02	0.19502E-01	intercept
405.00	1	-0.25606E-01	0.17253E-01	intercept
407.00	1	0.15296	0.16541E-01	intercept
408.00	1	-0.46129E-01	0.19626E-01	intercept
409.00	1	0.47700E-01	0.15045E-01	intercept
410.00	1	-0.17879	0.18470E-01	intercept
411.00	1	0.79235E-01	0.21262E-01	intercept
412.00	1	-0.33035E-01	0.17926E-01	intercept
414.00	1	0.52747E-01	0.17703E-01	intercept
415.00	1	0.18885	0.15239E-01	intercept
505.00	1	-0.93654E-01	0.14141E-01	intercept
506.00	1	-0.20558	0.14000E-01	intercept
507.00	1	-0.83834E-01	0.14362E-01	intercept
508.00	1	0.16605	0.14929E-01	intercept
509.00	1	-0.80506E-01	0.12115E-01	intercept
510.00	1	0.19704E-01	0.12331E-01	intercept
513.00	1	-0.11123	0.18563E-01	intercept
514.00	1	0.81555E-01	0.13039E-01	intercept
515.00	1	-0.58058E-01	0.10263E-01	intercept

School ID, random effect number, estimate, variance, name

# Select "File" > "Model-based Graphs" > "Confidence Intervals"

	Name	Predictor	Group		Mark	
1	Class intercept			-		Ŧ
2	School intercept			+		*



-1.0

150

Level 2 unit

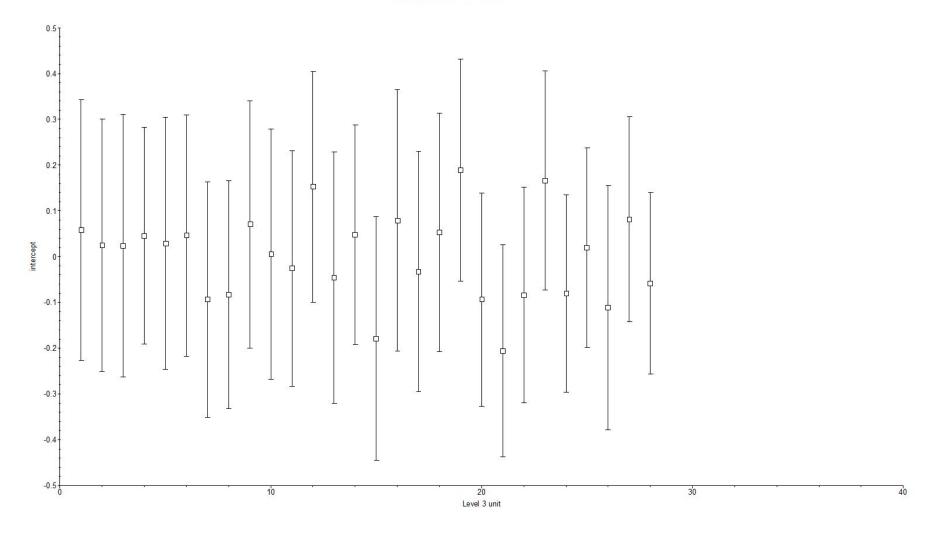
100

50

## Select "File" > "Model-based Graphs" > "Confidence Intervals"

	Name	Predictor	Group	Mark	
1	Class intercept			-	
2	School intercept			-	

95% Confidence Intervals



## Summary

- Spreadsheet allows some data manipulation
  - add/delete columns or rows
  - $-\operatorname{transformations}$  of variables (abs, exp, ln, sqrt, square)
  - summary statistics of variables (average, median, min, max, mode)
  - can create interaction terms and grand-mean centered variables
- Various kinds of data-based and model-based plots
- Up to 3-level models with full likelihood estimation (and empirical Bayes estimation of random effects)
- Linear transforms of parameter estimates
- Non-normal outcomes: binary, ordinal, nominal, and counts