

Hierarchical Linear Models

Joseph Stevens, Ph.D., University of Oregon
(541) 346-2445, stevensj@uoregon.edu

© Stevens, 2007

1

Overview and resources

- Overview
- Web site and links: www.uoregon.edu/~stevensj/HLM
- Software:
 - HLM
 - MLwinN
 - Mplus
 - SAS
 - R and S-Plus
 - WinBugs

2

Workshop Overview

- Preparing data
- Two Level models
- Testing nested hierarchies of models
- Estimation
- Interpreting results
- Three level models
- Longitudinal models
- Power in multilevel models

3

Hierarchical Data Structures

Many social and natural phenomena have a nested or clustered organization:

- Children within classrooms within schools
- Patients in a medical study grouped within doctors within different clinics
- Children within families within communities
- Employees within departments within business locations

4

Grouping and membership in particular units and clusters are important



For goodness sake, this is a huge field!
Why do we need to huddle like this all the time?

Hierarchical Data Structures

More examples of nested or clustered organization:

- Children within peer groups within neighborhoods
- Respondents within interviewers or raters
- Effect sizes within studies within methods (meta-analysis)
- Multistage sampling
- Time of measurement within persons within organizations

6

Simpson's Paradox:

Clustering Is Important

Well known paradox in which performance of individual groups is reversed when the groups are combined

	Quiz 1	Quiz 2
Gina	60.0%	10.0%
Sam	90.0%	30.0%

	Quiz 1	Quiz 2	Total
Gina	60 / 100	1 / 10	61 / 110
Sam	9 / 10	30 / 100	39 / 110

7

Simpson's Paradox: Other Examples

2006 US School study:

1975 Berkeley sex bias case:

- UCB sued for bias by women applying to grad school

“When the Oakies left Oklahoma and moved to California, it raised the IQ of both states.”
– Will Rogers

- Men applied more to high admission rate departments

8

Hypothetical Data Example from Snijders & Bosker (1999), $n = 2, j=5, n_j = N = 10$

Participant (i)	Cluster (j)	Outcome (Y)	Predictor (X)
1	1	5	1
2	1	7	3
3	2	4	1
4	2	6	4
5	3	3	3
6	3	5	5
7	4	2	4
8	4	4	6
9	5	1	5
10	5	3	7

All 10 cases analyzed without taking cluster membership into account:

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.333 ^a	.111	.000	1.826

a. Predictors: (Constant), X

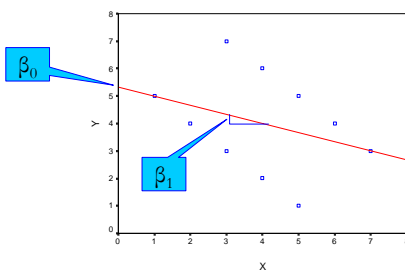
Model		Unstandardized Coefficients		Standardized Coefficients		t	Sig.
		B	Std. Error	Beta			
1	(Constant)	5.333	1.453			3.671	.006
	X	-.333	.333	-.333		-1.000	.347

a. Dependent Variable: Y

$$Y = 5.333 - .333(X) + r$$

Interpretation: There's a negative relationship between the predictor X and the outcome Y, a one unit increase in X results in .333 lower Y

$$Y = 5.333 - .333(X) + r$$



This is an example of a disaggregated analysis

Another alternative is to analyze data at the aggregated group level

Participant (i)	Cluster (j)	Outcome (Y)	Predictor (X)
1	1	5	1
2	1	7	3
3	2	4	2
4	2	6	4
5	3	3	3
6	3	5	5
7	4	2	4
8	4	4	6
9	5	1	5
10	5	3	7

Cluster (j)	Outcome (Y)	Predictor (X)
1	6	2
2	5	3
3	4	4
4	3	5
5	2	6

The clusters are analyzed without taking individuals into account:

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.100 ^a	.100	.099	.690

a. Predictors: (Constant), MEANX

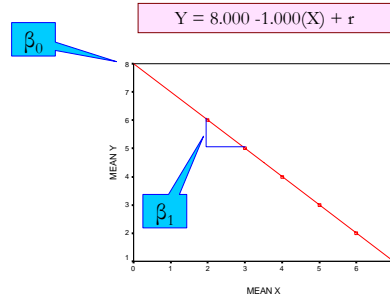
Coefficients ^a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	8.000	.000			
	MEANX	-1.000	.000	-1.000		

a. Dependent Variable: MEANY

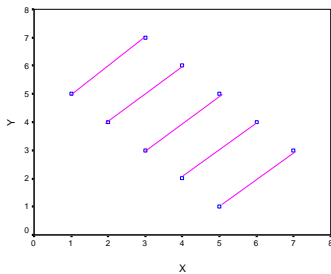
$$Y = 8.000 - 1.000(X) + r$$

Interpretation: There's a negative relationship between the predictor X and the outcome Y, a one unit increase in X results in 1.0 lower Y

This is an example of a disaggregated analysis



A third possibility is to analyze each cluster separately, looking at the regression relationship within each group



$$Y_{ij} = \bar{Y}_j + 1.00(X_{ij} - \bar{X}_j) + r_{ij}$$

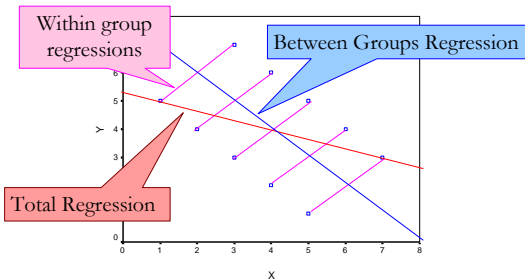
Multilevel regression takes both levels into account:

$$Y = 8.000 - 1.000(X) + r$$

$$Y_{ij} = \bar{Y}_j + 1.00(X_{ij} - \bar{X}_j) + r_{ij}$$

$$Y_{ij} = 8.00 - 1.00(\bar{X}_j) + 1.00(X_{ij} - \bar{X}_j) + r_{ij}$$

Taking the multilevel structure of the data into account:



Why Is Multilevel Analysis Needed?

- Nesting creates dependencies in the data
 - Dependencies violate the assumptions of traditional statistical models (“independence of error”, “homogeneity of regression slopes”)
 - Dependencies result in inaccurate statistical estimates
- Important to understand variation at different levels

Decisions About Multilevel Analysis

- Properly modeling multilevel structure often matters (and sometimes a lot)
- Partitioning variance at different levels is useful
 - tau and sigma ($\sigma^2_y = \tau + \sigma^2$)
 - policy & practice implications

“Randomization by cluster accompanied by analysis appropriate to randomization by individual is an exercise in self-deception and should be discouraged” (Cornfield, 1978, pp.101-2)

19

Preparing Data for HLM Analysis

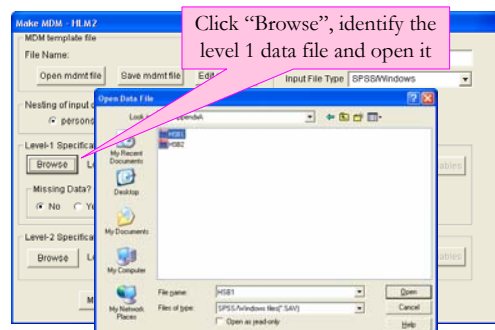
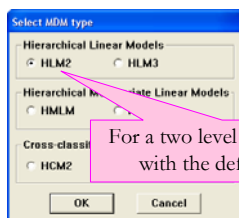
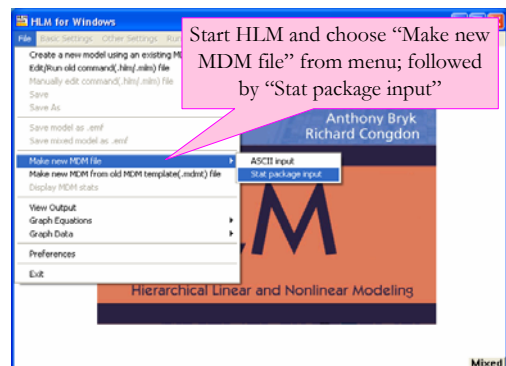
- Use of SPSS as a precursor to HLM assumed
- HLM requires a different data file for each level in the HLM analysis
- Prepare data first in SPSS
 - Clean and screen data
 - Treat missing data
 - ID variables needed to link levels
 - Sort cases on ID
- Then import files into HLM to create an “.mdm” file

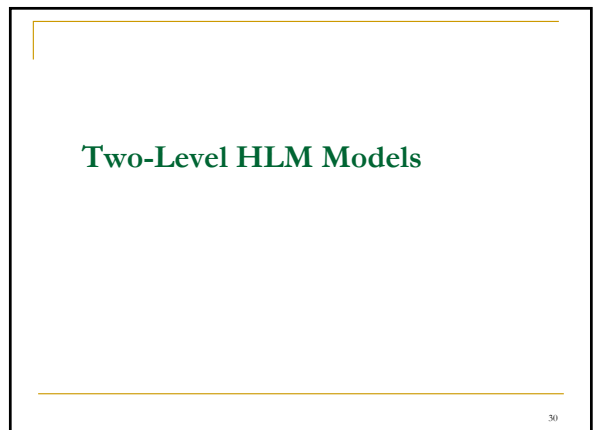
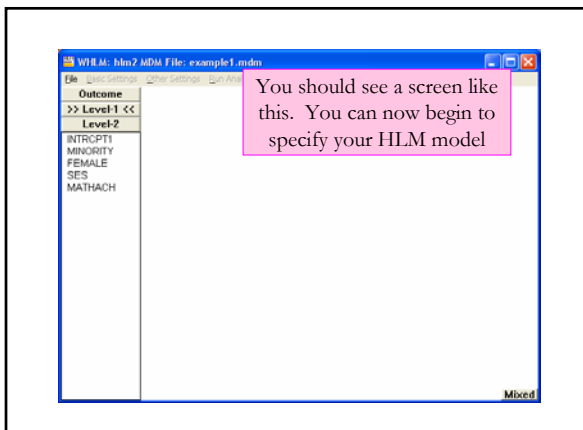
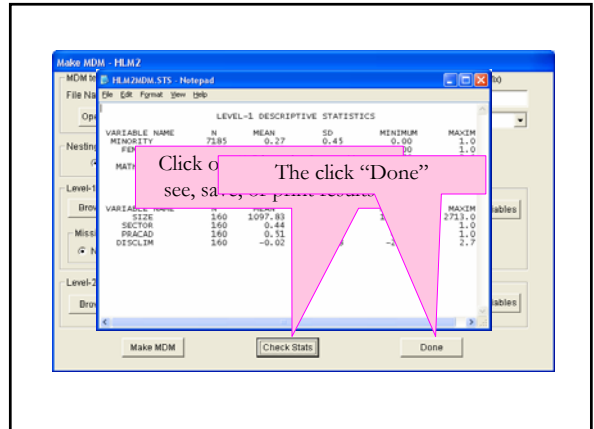
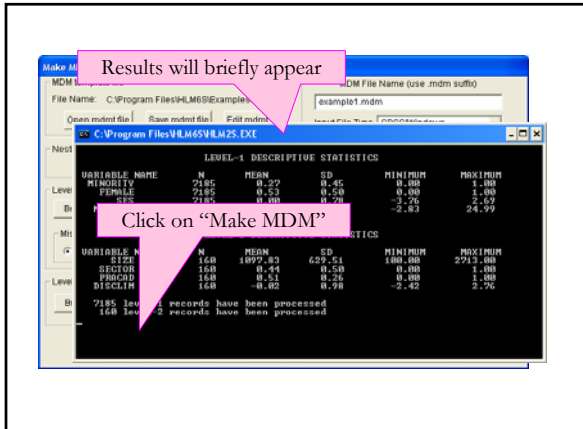
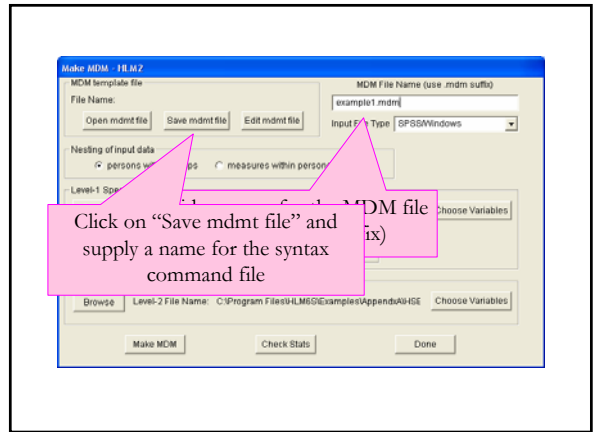
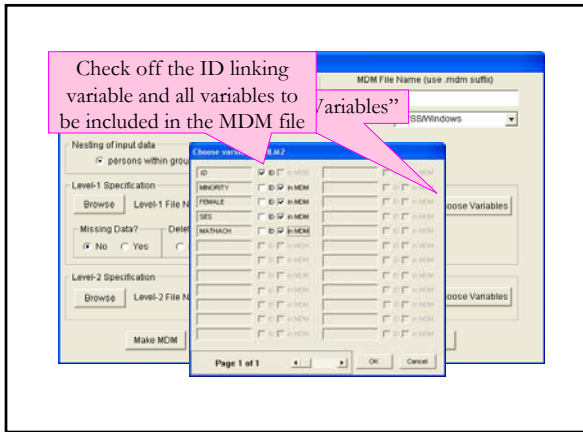
20

Creating an MDM file

- Example:
 - Go to “Examples” folder and then “Appendix A” in the HLM directory
 - Open “HSB1.sav” and “HSB2.sav”

21





The Single-Level, Fixed Effects Regression Model

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + r_i$$

- The parameters β_{ki} are considered fixed
 - One for all and all for one
 - Same values for all i and j ; the single level model
- The r_i 's are random: $r_i \sim N(0, \sigma)$ and independent

31

The Multilevel Model

- Takes groups into account and explicitly models group effects
- How to conceptualize and model group level variation?
- How do groups vary on the model parameters?
- Fixed versus random effects

32

Fixed vs. Random Effects

- Fixed Effects represent discrete, purposefully selected or existing values of a variable or factor
 - Fixed effects exert constant impact on DV
 - Random variability only occurs as a within subjects effect (level 1)
 - Can only generalize to particular values used
- Random Effects represent more continuous or randomly sampled values of a variable or factor
 - Random effects exert variable impact on DV
 - Variability occurs at level 1 and level 2
 - Can study and model variability
 - Can generalize to population of values

33

Fixed vs. Random Effects?

- Use fixed effects if
 - The groups are regarded as unique entities
 - If group values are determined by researcher through design or manipulation
 - Small j (< 10); improves power
- Use random effects if
 - Groups regarded as a sample from a larger population
 - Researcher wishes to test effects of group level variables
 - Researcher wishes to understand group level differences
 - Small j (< 10); improves estimation

34

Fixed Intercepts Model

- The simplest HLM model is equivalent to a one-way ANOVA with fixed effects:

$$Y_{ij} = \gamma_{00} + r_{ij}$$

- This model simply estimates the grand mean (γ_{00}) and deviations from the grand mean (r_{ij})
- Presented here simply to demonstrate control of fixed and random effects on all parameters

35

SPSS: Mixed Models File: hml1.mxd Command File: whltemp.hml

File Basic Settings Other Settings Run Analysis Help

Outcome

>> Level-1 <<

Level-2

INTRCPT1
SCORE
GENDER
BILLING
TITLE1
FILL
LEP
MODAD

LEVEL 1 MODEL (bold: group-mean centering, bold italic: grand-mean centering)

SCORE = $\beta_0 + \gamma$

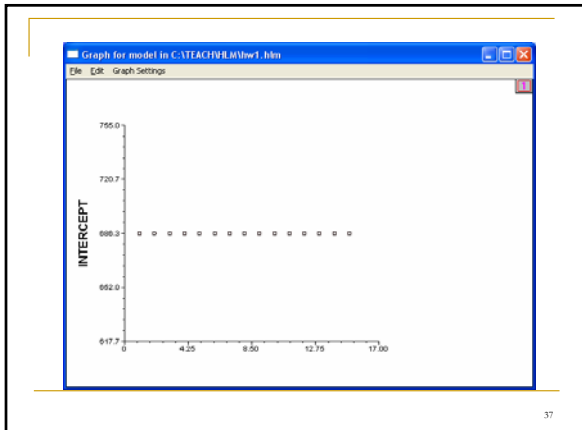
LEVEL 2 MODEL (bold italic: grand-mean centering)

$\beta_0 = \gamma_{00} + \gamma_{0j}$

Note the equation has no "u" residual term; this creates a fixed effect model

Mixed

36



37

ANOVA Model (random intercepts)

- A simple HLM model with random intercepts
- Equivalent to a one-way ANOVA model

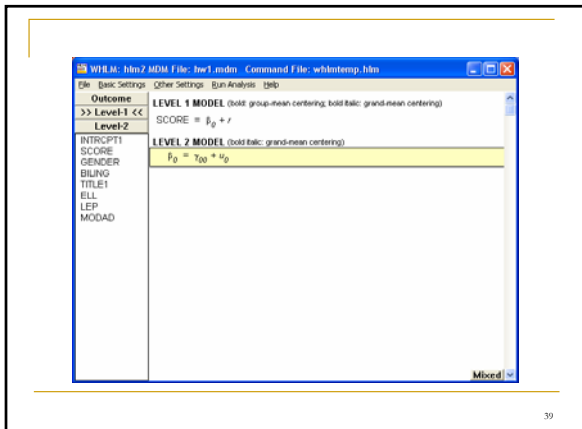
Note the addition of u_{0j} allows different intercepts for each j unit, a random effects model

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

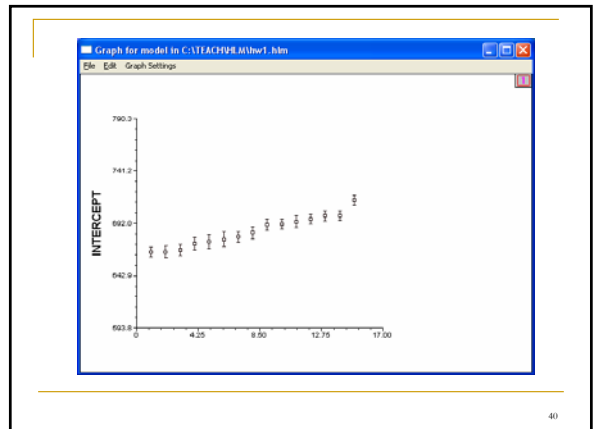
$$\beta_{0j} = \gamma_{00} + u_{0j}$$

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

38



39



40

ANOVA Model

- In addition to providing parameter estimates, the ANOVA model provides information about the presence of level 2 variance (the ICC) and whether there are significant differences between level 2 units
- This model also called the Unconditional Model (because it is not “conditioned” by any predictors) and the “empty” model
- Often used as a baseline model for comparison to more complex models

41

Variables in HLM Models

- Outcome variables
- Predictors
 - Control variables
 - Explanatory variables
- Variables at higher levels
 - Aggregated variables (Is n sufficient for representation?)
 - Contextual variables

42

Conditional Models: ANCOVA

- Adding a predictor to the ANOVA model results in an ANCOVA model with random intercepts:

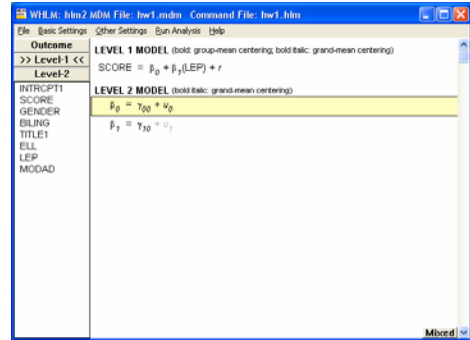
$$Y_{ij} = \beta_{0j} + \beta_1(X_{1j}) + \epsilon_{ij}$$

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

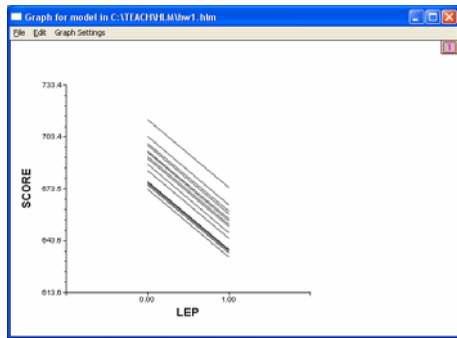
$$\beta_1 = \gamma_{10}$$

- Note that the effect of X is constrained to be the same fixed effect for every j unit (homogeneity of regression slopes)

43



44



45

Conditional Models: Random Coefficients

- An additional parameter results in random variation of the slopes:

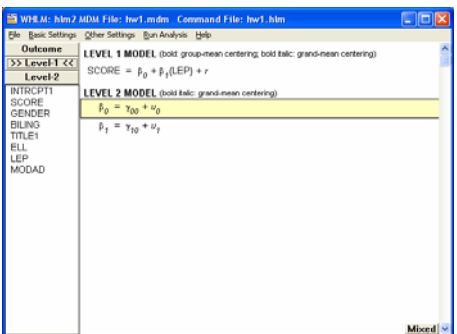
$$Y_{ij} = \beta_{0j} + \beta_1(X_{1j}) + \epsilon_{ij}$$

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

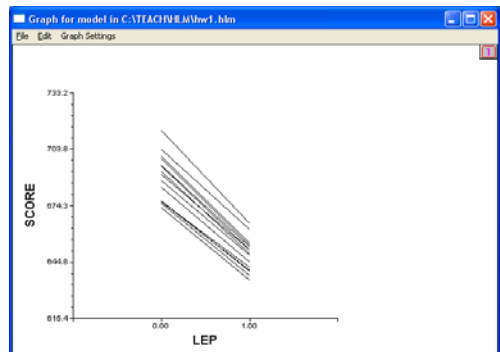
$$\beta_{1j} = \gamma_{10} + u_{1j}$$

- Both intercepts and slopes now vary from group to group

46



47



48

Standardized coefficients

Standardized coefficient at level 1:

$$\beta_{0j} (SD_X / SD_Y)$$

Standardized coefficient at level 2:

$$\gamma_{00} (SD_X / SD_Y)$$

49

Modeling variation at Level 2: Intercepts as Outcomes

$$Y_{ij} = \beta_{1j} X_{1ij} + r_{ij}$$

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

- Predictors (W's) at level 2 are used to model variation in intercepts between the j units

50

Modeling Variation at Level 2: Slopes as Outcomes

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{1ij} + r_{ij}$$

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

- Do slopes vary from one j unit to another?
- W's can be used to predict variation in slopes as well

51

Variance Components Analysis

- VCA allows estimation of the size of random variance components
 - Important issue when unbalanced designs are used
 - Iterative procedures must be used (usually ML estimation)
- Allows significance testing of whether there is variation in the components across units

52

Estimating Variance Components: Unconditional Model

$$\begin{aligned} \text{Var}(Y_{ij}) &= \text{Var}(u_{0j}) + \text{Var}(r_{ij}) \\ &= \tau_0 + \sigma^2 \end{aligned}$$

53

HLM Output

Final estimation of variance components:

Random Effect		Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1,	U0	14.38267	206.86106	14	457.32201	0.000
level-1,	R	32.58453	1061.75172			

Statistics for current covariance components model

Deviance = 21940.853702
Number of estimated parameters = 2

54

Variance explained

R^2 at level 1 =

$$1 - (\sigma_{\text{cond}}^2 + \tau_{\text{cond}}) / (\sigma_{\text{uncond}}^2 + \tau_{\text{uncond}})$$

R^2 at level 2 =

$$1 - [(\sigma_{\text{cond}}^2 / n_h) + \tau_{\text{cond}}] / [(\sigma_{\text{uncond}}^2 / n_h) + \tau_{\text{uncond}}]$$

Where n_h = the harmonic mean of n for the level 2 units
 $(k / [1/n_1 + 1/n_2 + \dots + 1/n_k])$

55

Comparing models

■ Deviance tests

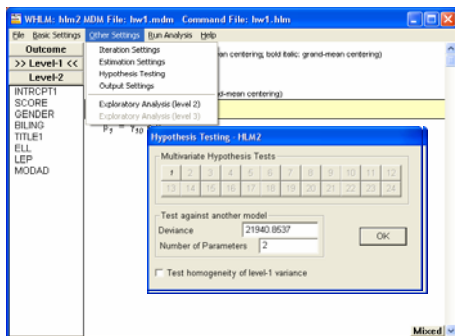
- Under “Other Settings” on HLM tool bar, choose “hypothesis testing”
- Enter deviance and number of parameters from baseline model

■ Variance explained

- Examine reduction in unconditional model variance as predictors added, a simpler level 2 formula:

$$R^2 = (\tau_{\text{baseline}} - \tau_{\text{conditional}}) / \tau_{\text{baseline}}$$

56



57

Deviance Test Results

Statistics for current covariance components model

 Deviance = 21615.283709
 Number of estimated parameters = 2

Variance-Covariance components test

 Chi-square statistic = 325.56999
 Number of degrees of freedom = 0
 P-value = >.500

58

Testing a Nested Sequence of HLM Models

1. Test unconditional model
2. Add level 1 predictors
 - Determine if there is variation across groups
 - If not, fix parameter
 - Decide whether to drop nonsignificant predictors
 - Test deviance, compute R^2 if so desired
3. Add level 2 predictors
 - Evaluate for significance
 - Test deviance, compute R^2 if so desired

59

Example

- Use the HSB MDM file previously created to practice running HLM models:
 - Unconditional
 - Level 1 predictor fixed, then random
 - Level 2 predictor

60

Statistical Estimation in HLM Models

- Estimation Methods
 - FML
 - RML
 - Empirical Bayes estimation
- Parameter estimation
 - Coefficients and standard errors
 - Variance Components
- Parameter reliability
- Centering
- Residual files

61

Estimation Methods: Maximum Likelihood Estimation (MLE) Methods

- MLE estimates model parameters by estimating a set of population parameters that maximize a likelihood function
- The likelihood function provides the probabilities of observing the sample data given particular parameter estimates
- MLE methods produce parameters that maximize the probability of finding the observed sample data

62

Estimation Methods

RML – Restricted Maximum Likelihood, only
FML – Full Maximum Likelihood, both the regression coefficients and the variance components are included in the likelihood function

components

Goodness of fit statistics (deviance tests) apply only to the random effects

RML only tests hypotheses about the VCs (and the models being compared must have identical fixed effects)

fixed effects and the variance components.

Goodness of fit statistics apply to the entire model

(both fixed and random effects)

Check on software default

63

Estimation Methods

- RML expected to lead to better estimates especially when j is small
- FML has two advantages:
 - Computationally easier
 - With FML, overall chi-square tests both regression coefficients and variance components, with RML only variance components are tested
 - Therefore if fixed portion of two models differ, must use FML for nested deviance tests

64

Computational Algorithms

- Several algorithms exist for existing HLM models:
 - Expectation-Maximization (EM)
 - Fisher scoring
 - Iterative Generalized Least Squares (IGLS)
 - Restricted IGLS (RIGLS)
- All are iterative search and evaluation procedures

65

Model Estimation

- Iterative estimation methods usually begin with a set of start values
- Start values are tentative values for the parameters in the model
 - Program begins with starting values (usually based on OLS regression at level 1)
 - Resulting parameter estimates are used as initial values for estimating the HLM model

66

Model Estimation

- Start values are used to solve model equations on first iteration
- This solution is used to compute initial model fit
- Next iteration involves search for better parameter values
- New values evaluated for fit, then a new set of parameter values tried
- When additional changes produce no appreciable improvement, iteration process terminates (convergence)
- Note that convergence and model fit are very different issues

67

Intermission



"Frankly, Harold, you're beginning to bore everyone with your statistics."

Centering

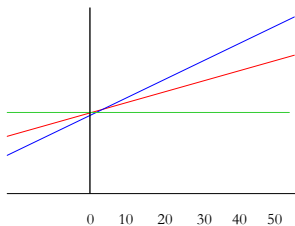
- No centering (common practice in single level regression)
- Centering around the group mean (\bar{X}_j)
- Centering around the grand mean (M)
- A known population mean
- A specific meaningful time point

69

Centering: The Original Metric

- Sensible when 0 is a meaningful point on the original scale of the predictor
 - For example, amount of training ranging from 0 to 14 days
 - Dosage of a drug where 0 represents placebo or no treatment
- Not sensible or interpretable in many other contexts, i.e. SAT scores (which range from 200 to 800)

70



$$\beta_{0j} = E(Y_{ij} | X_{ij} = 0)$$

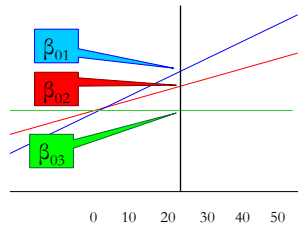
71

Centering Around the Grand Mean

- Predictors at level 1 (X 's) are expressed as deviations from the grand mean (M): $(X_{ij} - M)$
- Intercept now expresses the expected outcome value (Y) for someone whose value on predictor X is the same as the grand mean on that predictor
- Centering is computationally more efficient
- Intercept represents the group mean adjusted for the grand mean $\bar{X}_j - M$
- Variance of $\beta_{0j} = \tau_{0j}$, the variance among the level-2 unit means adjusted for the grand mean

72

Centering Around the Grand Mean



$$\beta_{0j} = E(Y_{ij} | X_{ij} = \gamma_{00})$$

73

Centering Around the Group Mean

- Individual scores are interpreted relative to their group mean ($X_{ij} - \bar{X}_j$)
- The individual deviation scores are orthogonal to the group means
- Intercept represents the unadjusted mean achievement for the group
- Unbiased estimates of within-group effects
- May be necessary in random coefficient models if level 1 predictor affects the outcome at both level 1 and level 2
- Can control for unmeasured between group differences
- But can mask between group effects; interpretation is more complex

74

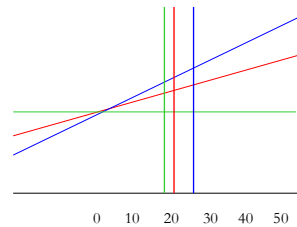
Centering Around the Group Mean

- Level 1 results are relative to group membership
- Intercept becomes the unadjusted mean for group j
- Should include level 2 mean of level 1 variables to fully disentangle individual and compositional effects
- Variance β_{0j} is now the variance among the level 2 unit means

75

Centering Around the Group Mean

If $\bar{X}_j = 25$, then $\beta_{0j} = E(Y_{ij} | X_{ij} = 25)$
 If $\bar{X}_j = 30$, then $\beta_{0j} = E(Y_{ij} | X_{ij} = 30)$
 If $\bar{X}_3 = 18$, then $\beta_{03} = E(Y_{i3} | X_{i3} = 18)$



76

Parameter estimation

- Coefficients and standard errors estimated through maximum likelihood procedures (usually)
 - The ratio of the parameter to its standard error produces a Wald test evaluated through comparison to the normal distribution (z)
 - In HLM software, a more conservative approach is used:
 - t-tests are used for significance testing
 - t-tests more accurate for fixed effects, small n , and nonnormal distributions
- Standard errors
- Variance components

77

Parameter reliability

- Analogous to score reliability: ratio of true score variance to total variance (true score + error)
- In HLM, ratio of true parameter variance to total variability
- For example, $\lambda_j = \text{True variance of the sample means (estimated)} / \text{Variance of error of the sample means}$ is:

$$\lambda_j = \text{Var}(\beta_{0j}) / \text{Var}(\bar{Y}_j) = \tau_{00}^2 / (\tau_{00}^2 + \sigma^2 / n_j)$$

True variance of the sample means (estimated) Variance of error of the sample means

True variance of the sample means (estimated) Total variance of the sample means (observed)

78

Parameter reliability

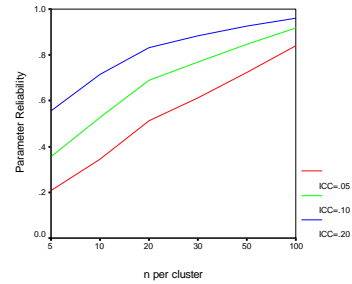
$$\lambda_j = \frac{n_j \rho_I}{1 + (n_j - 1) \rho_I}$$

n_j	ICC (ρ_I)		
	.05	.10	.20
5	.21	.36	.56
10	.34	.53	.71
20	.51	.69	.83
30	.61	.77	.88
50	.72	.85	.93
100	.84	.92	.96

79

Parameter reliability

$$\lambda_j = \frac{n_j \rho_I}{1 + (n_j - 1) \rho_I}$$



80

Predicting Group Effects

- It is often of interest to estimate the random group effects (β_{0j}, β_{1j})
- This is accomplished using Empirical Bayes (EB) estimation
- The basic idea of EB estimation is to predict group values using two kinds of information:
 - Group j data
 - Population data obtained from the estimation of the regression model

81

Empirical Bayes

- If information from only group j is used to estimate then we have the OLS estimate:

$$\beta_{0j} = \bar{Y}_j$$

- If information from only the population is used to estimate then the group is estimated from the grand mean:

$$\gamma_{00} = \bar{Y}_{..} = \sum_{j=1}^N \frac{n_j}{N} \bar{Y}_j$$

82

Empirical Bayes

- A third possibility is to combine the two sources of information in an average weighted by the parameter's reliability:

$$\beta_{0j}^{EB} = \lambda_j \beta_{0j} + (1 - \lambda_j) \gamma_{00}$$

- The results in the "posterior means" or EB estimates

The smaller the reliability, the greater the weight of the grand mean

The larger the reliability, the greater the weight of the group mean

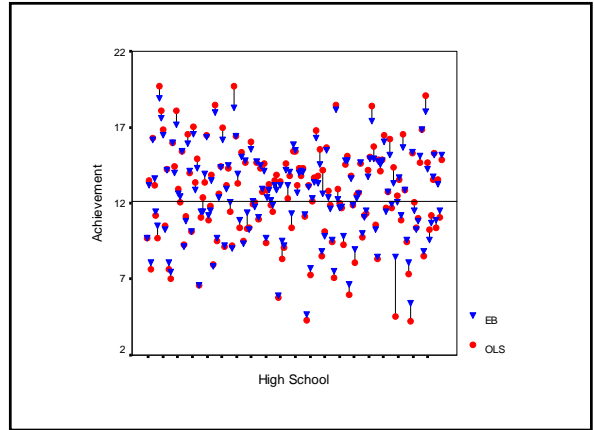
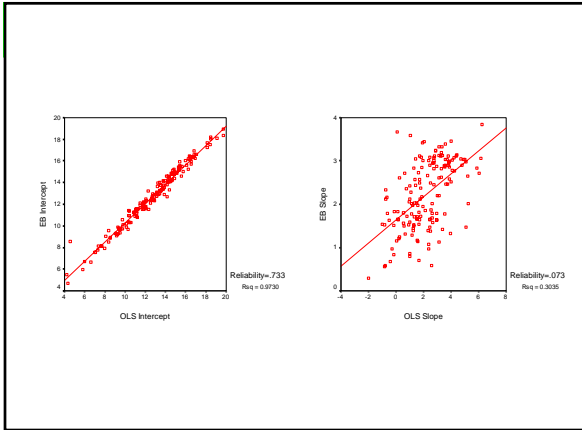
83

Bayesian Estimation

- Use of prior and posterior information improves estimation (depending on purpose)
- Estimates "shrink" toward the grand mean as shown in formula
- Amount of shrinkage depends on the "badness" of the unit estimate
 - Low reliability results in greater shrinkage (if $\lambda = 1$, there is no shrinkage; if $\lambda = 0$, shrinkage is complete, γ_{00})
 - Small n -size within a j unit results in greater shrinkage, "borrowing" from larger units

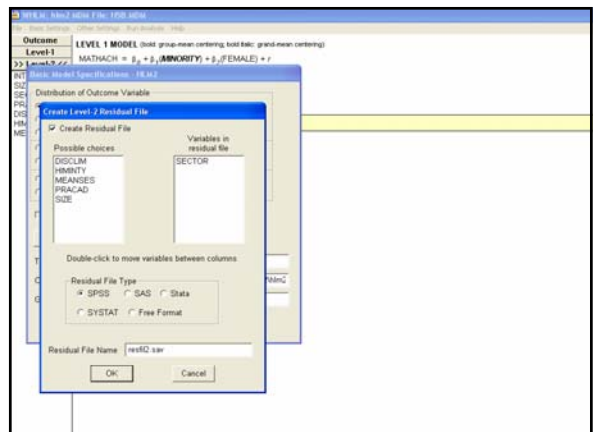
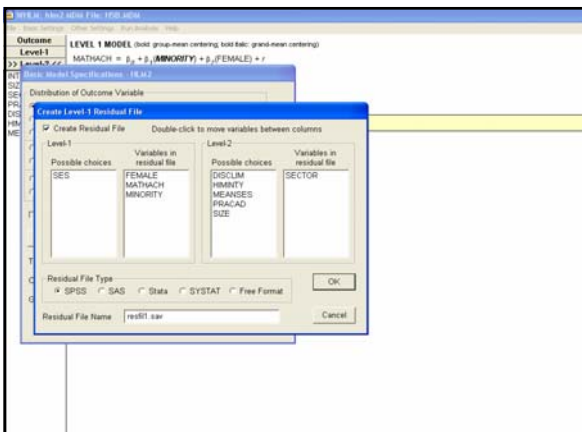
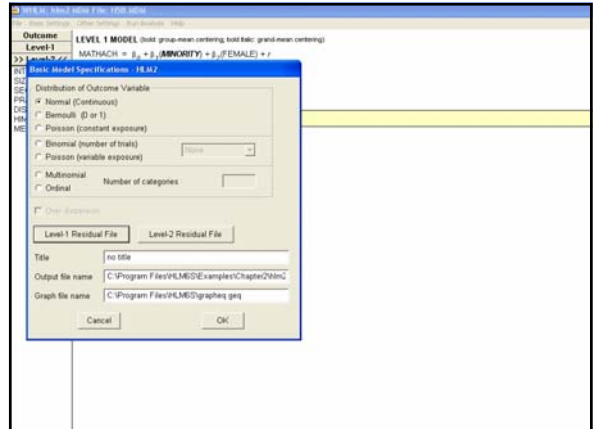
$$\beta_{0j}^{EB} = \lambda_j \beta_{0j} + (1 - \lambda_j) \gamma_{00}$$

84



HLM Residual Files

- Important outcome information from an HLM analysis can be saved for each level of the analysis in a “residual file”
 - Residual files contain parameter estimates and other variables from the analysis
 - Residual files can be save in statistical package format (SPSS, SAS, etc.)
- Residual files can be used for diagnostic evaluation of statistical model assumptions
- Residual files can be used to estimate and further describe or analyze effects among the units at each level



Cid	fressal	Bhal	sigma	minority	female	mathach	sector	var	var	var	var
1 1224	-4.180	10.066	6.081	-.275	1.000	5.076	-.438				
2 1224	9.662	10.066	6.081	-.275	1.000	19.700	-.438				
3 1224	8.813	11.536	6.081	-.275	0.000	20.349	-.438				
4 1224	-2.755	11.536	6.081	-.275	0.000	8.781	-.438				
5 1224	6.362	11.536	6.081	-.275	0.000	17.880	-.438				
6 1224	4.963	11.536	6.081	-.275	0.000	4.583	-.438				
7 1224	-12.088	10.066	6.081	-.275	1.000	-2.032	-.438				
8 1224	-11.013	11.536	6.081	-.275	0.000	5.23	-.438				
9 1224	-8.529	10.066	6.081	-.275	1.000	1.527	-.438				
10 1224	9.985	11.536	6.081	-.275	0.000	21.521	-.438				
11 1224	-.581	10.066	6.081	-.275	1.000	8.475	-.438				
12 1224	6.031	10.066	6.081	-.275	1.000	16.667	-.438				
13 1224	9.642	11.536	6.081	-.275	0.000	21.178	-.438				
14 1224	10.122	10.066	6.081	-.275	1.000	20.178	-.438				
15 1224	8.813	11.536	6.081	-.275	0.000	20.349	-.438				
16 1224	10.452	10.066	6.081	-.275	1.000	20.500	-.438				
17 1224	7.802	11.536	6.081	-.275	0.000	19.338	-.438				
18 1224	-1.872	6.017	6.081	.725	0.000	4.148	-.438				
19 1224	-7.129	10.066	6.081	-.275	1.000	2.927	-.438				
20 1224	4.869	11.536	6.081	-.275	0.000	16.405	-.438				
21 1224	3.597	10.066	6.081	-.275	1.000	13.653	-.438				
22 1224	-3.498	10.066	6.081	-.275	1.000	6.588	-.438				
23 1224	-.403	10.066	6.081	-.275	1.000	9.653	-.438				
24 1224	-8.382	11.536	6.081	-.275	0.000	3.154	-.438				
25 1224	-.217	4.536	6.081	.725	1.000	4.793	-.438				
26 1224	10.166	11.536	6.081	-.275	0.000	21.721	-.438				
27 1224	-4.890	10.066	6.081	-.275	1.000	5.166	-.438				
28 1224	-3.657	10.066	6.081	-.275	1.000	6.399	-.438				
29 1224	-3.236	10.066	6.081	-.275	1.000	6.821	-.438				

Cid	n	chgpt	midst	mltvtr	slstrvr	mndvtr	slbtrtp	slmtrvr	slbtrvr	slmtrvr	
1 1224	47	4.812	3.672	2.027	1.975	1.963	-2.098	-.331	.024	-2.436	-4.601
2 1208	25	2.061	9.223	1.849	1.960	1.862	.856	-.050	.062	.457	-.127
3 1296	26	4.688	3.969	1.870	1.870	2.443	-.124	-.362	.100	-3.418	11.865
4 1308	20	-1.0E+36	-1.0E+36	1.811	-1.0E+36	1.913	1.201	-.071	-.262	-1.0E+36	-1.0E+36
5 1317	48	-1.0E+36	-1.0E+36	1.688	-1.0E+36	2.300	.615	-.005	-.032	-1.0E+36	-1.0E+36
6 1360	30	4.974	3.871	1.771	1.513	1.384	-.941	-.525	.052	-.197	-6.892
7 1374	28	3.095	2.291	2.123	2.033	2.170	-.304	-.465	-.146	-.121	-2.868
8 1433	35	-1.0E+36	-1.0E+36	1.366	-1.0E+36	-1.0E+36	3.410	-.000	-.703	-1.0E+36	-1.0E+36
9 1436	44	3.136	2.293	1.515	-1.0E+36	1.176	2.684	-.287	-.644	-1.0E+36	-1.0E+36
10 1461	33	10.373	18.322	1.839	1.800	1.675	3.448	-.792	-.440	.895	-16.588
11 1462	57	-1.0E+36	-1.0E+36	1.842	-1.0E+36	2.188	-1.945	-.576	.336	-1.0E+36	-1.0E+36
12 1477	62	1.267	184	1.968	1.980	1.972	-.641	.034	.196	-.957	-.369
13 1489	53	4.379	3.141	1.848	1.806	2.106	-2.284	-.417	.044	-2.198	1.732
14 1537	27	2.783	1.917	1.862	1.887	2.286	-1.619	-.244	.400	-2.086	-1.757
15 1506	53	4.446	3.193	1.874	1.839	1.929	.960	-.656	.039	.025	-4.881
16 1509	28	1.647	6.66	1.817	-1.0E+36	1.312	1.105	-.087	-.062	-1.0E+36	-1.0E+36
17 1642	29	8.613	8.480	1.700	1.724	1.401	4.440	-.101	-.514	7.404	9.128
18 1646	39	8.240	8.049	1.843	1.868	1.910	1.148	.533	-.787	3.896	5.418
19 2030	47	1.181	1.56	1.839	1.800	1.634	-.334	.642	-.046	-.147	-.303
20 2208	60	1.991	.907	1.812	-1.0E+36	1.825	.765	-.107	-.417	-1.0E+36	-1.0E+36
21 2277	61	-1.0E+36	-1.0E+36	1.717	-1.0E+36	2.231	-2.979	.413	.201	-1.0E+36	-1.0E+36
22 2305	67	-1.0E+36	-1.0E+36	1.612	-1.0E+36	2.314	.867	-.228	.072	-1.0E+36	-1.0E+36
23 2336	47	4.734	3.548	1.766	1.778	1.300	3.297	-.177	-.374	4.586	3.071
24 2458	57	-1.0E+36	-1.0E+36	1.766	-1.0E+36	2.291	1.088	-.008	.067	-1.0E+36	-1.0E+36
25 2467	62	3.211	2.369	1.916	-1.0E+36	1.792	-2.612	.278	.677	-1.0E+36	-1.0E+36
26 2526	57	-1.0E+36	-1.0E+36	1.564	-1.0E+36	1.922	2.508	.022	-.125	-1.0E+36	-1.0E+36
27 2636	38	2.431	1.491	1.832	-1.0E+36	1.426	.738	-.109	-.452	-1.0E+36	-1.0E+36
28 2639	57	-1.0E+36	-1.0E+36	1.642	-1.0E+36	1.682	-.073	1.089	1.28	-1.0E+36	-1.0E+36
29 2639	42	1.374	.276	1.787	-1.0E+36	2.364	-.519	-.178	-.065	-1.0E+36	-1.0E+36

Example: Creating Residual Files

Three level models

Level-1 (p students)

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}(a_{pijk}) + e_{ijk}$$

Level-2 (j classrooms)

$$\pi_{0jk} = \beta_{p0k} + \beta_{p1k}(X_{qjk}) + r_{p0k}$$

$$\pi_{1jk} = \beta_{p1k} + \beta_{p1k}(X_{qjk}) + r_{p1k}$$

Level-3 (k schools)

$$\beta_{p0k} = \gamma_{pq0} + \gamma_{pq5}(W_{sk}) + u_{pqk}$$

$$\beta_{p1k} = \gamma_{pq1} + \gamma_{pq5}(W_{sk}) + u_{pqk}$$

Partitioning variance in the three level model

Proportion of variance within classrooms (individual student differences) = $\sigma^2 / (\sigma^2 + \tau_\pi + \tau_\beta)$

Proportion of variance between classrooms within schools = $\tau_\pi / (\sigma^2 + \tau_\pi + \tau_\beta)$

Proportion of variance between schools = $\tau_\beta / (\sigma^2 + \tau_\pi + \tau_\beta)$

97

Three level example

- Example:
 - Go to “Examples” folder and then “Chapter 4” in the HLM directory
 - Open “EG1.sav”, “EG2.sav”, and “EG3.sav”

98

Longitudinal models

- Level 1 defined as repeated measurement occasions
- Levels 2 and 3 defined as higher levels in the nested structure
- For example, longitudinal analysis of student achievement
 - Level 1 = achievement scores at times 1 – t
 - Level 2 = student characteristics
 - Level 3 = school characteristics

99

Longitudinal models

- Two important advantages of the MLM approach to repeated measures:
 - Times of measurement can vary from one person to another
 - Data do not need to be complete on all measurement occasions

100

Longitudinal models

Level-1

$$Y_{ij} = \pi_{0ij} + \pi_{1ij}(\text{time}) + e_{ij}$$

Level-2

$$\pi_{0ij} = \beta_{00j} + \beta_{01j}(X_{ij}) + r_{0ij}$$

$$\pi_{1ij} = \beta_{10j} + \beta_{11j}(X_{ij}) + r_{1ij}$$

Level-3

$$\beta_{00j} = \gamma_{000} + \gamma_{001}(W_{1j}) + u_{00j}$$

$$\beta_{10j} = \gamma_{100} + \gamma_{101}(W_{1j}) + u_{10j}$$

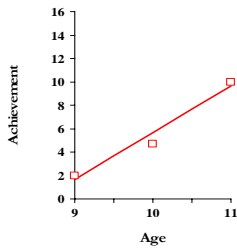
101

A variable is also needed to indicate the time of occurrence of each measurement occasion link the level 2 file to level 3

The set of rows represent the repeated measures for one participant

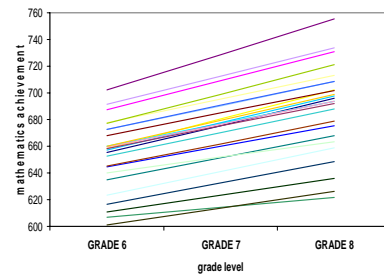
id	schnumb	math	ye	link
1	63	1417	697	1
2	63	1417	671	1
3	63	1417	707	2
4	171	1417	646	0
5	171	1417	660	1
6	171	1417	669	2
7	190	1417	652	0
8	190	1417	674	1
9	190	1417	691	2
10	246	1417	644	0
11	246	1417	650	1
12	246	1417	680	2
13	267	1417	649	0
14	267	1417	645	1
15	267	1417	651	2
16	279	1417	654	0
17	279	1417	661	1
18	279	1417	690	2
19	332	1417	639	0
20	332	1417	673	1
21	332	1417	723	2
22	451	1417	669	0
23	451	1417	685	1

Fitting a growth trajectory



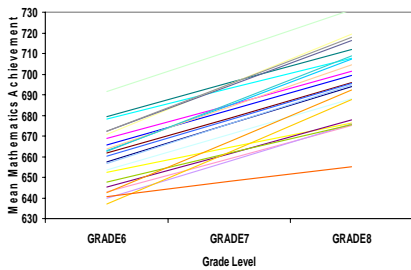
103

Linear growth for individual students



104

Average linear growth by school



105

Curvilinear Longitudinal models

Level-1

$$Y_{ij} = \pi_{0ij} + \pi_{1ij}(\text{time}) + e_{tij}$$

Level-2

$$\pi_{0ij} = \beta_{00j} + \beta_{01j}(X_{ij}) + r_{0ij}$$

$$\pi_{1ij} = \beta_{10j} + \beta_{11j}(X_{ij}) + r_{1ij}$$

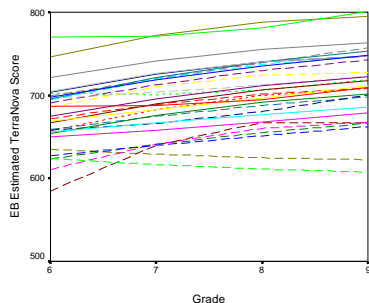
Level-3

$$\beta_{00j} = \gamma_{000} + \gamma_{001}(W_{1j}) + u_{00j}$$

$$\beta_{10j} = \gamma_{100} + \gamma_{101}(W_{1j}) + u_{10j}$$

106

Curvilinear growth for individual students



107

Testing a Nested Sequence of HLM Longitudinal Models

1. Test unconditional model
2. Test Level 1 growth model
3. After establishing the level 1 growth model, use it as the baseline for succeeding model comparisons
4. Add level 2 predictors
 - Determine if there is variation across groups
 - If not, fix parameter
 - Decide whether to drop nonsignificant predictors
 - Test deviance, compute R^2 if so desired
5. Add level 3 predictors
 - Evaluate for significance
 - Test deviance, compute R^2 if so desired

108

Regression Discontinuity and Interrupted Time Series Designs: Change in Intercept

$$Y_{ij} = \pi_{0i} + \pi_{1i}Time_{ij} + \pi_{2i}Treatment_{ij} + \varepsilon_{ij}$$

When Treatment = 0:

$$Y_{ij} = \pi_{0i} + \pi_{1i}Time_{ij} + \varepsilon_{ij}$$

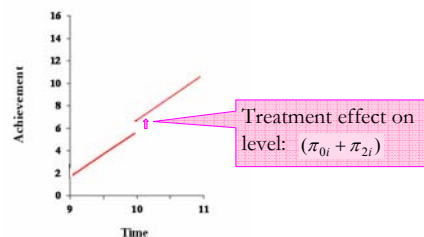
When Treatment = 1:

$$Y_{ij} = (\pi_{0i} + \pi_{2i}) + \pi_{1i}Time_{ij} + \varepsilon_{ij}$$

Treatment is coded 0 or 1

109

Regression Discontinuity and Interrupted Time Series Designs



110

Regression Discontinuity and Interrupted Time Series Designs: Change in Slope

When Treatment = 1:

$$Y_{ij} = \pi_{0i} + \pi_{1i}Time_{ij} + \pi_{3i}TreatmentTime_{ij} + \varepsilon_{ij}$$

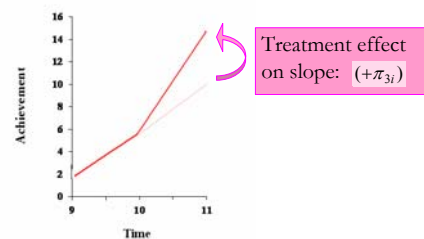
When Treatment = 0:

$$Y_{ij} = \pi_{0i} + \pi_{1i}Time_{ij} + \varepsilon_{ij}$$

Treatment time expressed as 0's before treatment and time intervals post-treatment (i.e., 0, 0, 0, 1, 2, 3)

111

Regression Discontinuity and Interrupted Time Series Designs



112

Change in Intercept and Slope

$$Y_{ij} = \pi_{0i} + \pi_{1i}Time_{ij} + \pi_{2i}Treatment + \pi_{3i}TreatmentTime_{ij} + \varepsilon_{ij}$$

Effect of treatment on intercept

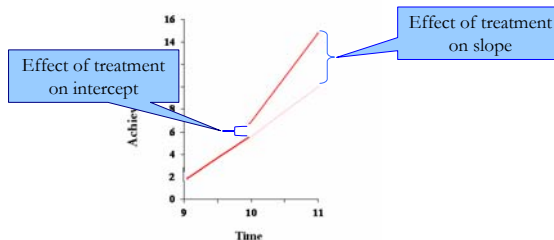
Effect of treatment on slope

When Treatment = 0:

$$Y_{ij} = \pi_{0i} + \pi_{1i}Time_{ij} + \varepsilon_{ij}$$

113

Regression Discontinuity and Interrupted Time Series Designs



114

Analyzing Randomized Trials (RCT) with HLM

- When random assignment is accomplished at the participant level, treatment group is dummy coded and included in the participant level data file
- When random assignment is accomplished at the cluster level, treatment group is dummy coded and included in the cluster level data file
 - Treatment can be used to predict intercepts or slopes as outcomes
 - Another strength of this approach is the ability to empirically model treatment variation across clusters (i.e., replication)

115

Power in HLM Models

116

Using the Optimal Design Software

- The Optimal Design Software can also be used to estimate power in a variety of situations
- The particular strength of this software is its application to multilevel situations involving cluster randomization or multisite designs
- Available at:
http://sitemaker.umich.edu/group-based/optimal_design_software
- Optimal Design

117

Factors Affecting Power in CRCT

- Sample Size
 - Number of participants per cluster (N)
 - Number of clusters (J)
- Effect Size
- Alpha level
- Unexplained/residual variance
- Design Effects
 - Intraclass correlation (ICC)
 - Between vs. within cluster variance
 - Treatment variability across clusters
 - Repeated measures
 - Blocking and matching
- Statistical control

118

Effect of Unexplained Variance on Power

- Terminology: “error” versus unexplained or residual
- Residual variance reduces power
 - Anything that decreases residual variance, increases power (e.g., more homogeneous participants, additional explanatory variables, etc.)
- Unreliability of measurement contributes to residual variance
- Treatment infidelity contributes to residual variance
- Consider factors that may contribute to residual between cluster variance

119

Effect of Design Features on Statistical Power

- Multicollinearity (and restriction of range)

$$s_{b_{y1.2}} = \sqrt{\frac{s_{y12}^2}{\text{[redacted]}}}$$

- Statistical model misspecification
 - Linearity, curvilinearity, ...
 - Omission of relevant variables
 - Inclusion of irrelevant variables

120

The number of clusters has a stronger influence on power than the cluster size as ICC departs from 0

- The standard error of the main effect of treatment is:

$$SE(\hat{\gamma}_{01}) = \sqrt{\frac{4(\rho + (1-\rho)/n)}{J}}$$

- As ρ increases, the effect of n decreases
- If clusters are variable (ρ is large), more power is gained by increasing the number of clusters sampled than by increasing n

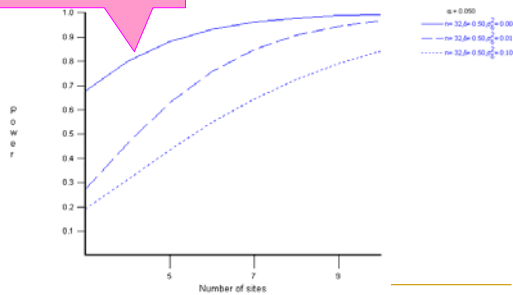
121

Power in Studies with a Small Number of Clusters

122

Fixed vs. Random Effects

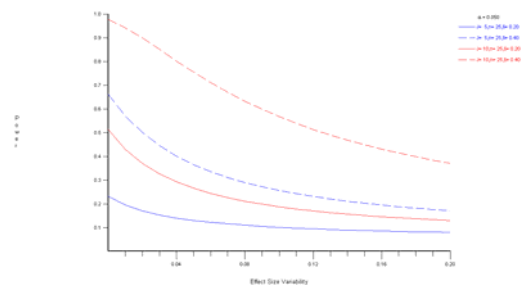
Fixed Effects Model



123

Effect of Effect Size Variability (σ_δ^2)

Variance of the treatment effect across clusters



124

Randomization as a Control Tactic

- What does randomization accomplish?
 - Controls bias in assignment of treatment (works OK with small J)
 - Turns confounding factors into randomly related effects (equivalence vs. randomness; does not work well with small J)
- Applying an underpowered, small CRCT may not be sufficient to achieve rigor
 - Consider other design approaches (e.g., interrupted time series, regression discontinuity designs)
 - Aggressively apply other tactics for experimental or statistical control
- Not all designs are created equal
- No one design is best (e.g., randomized trials)

125

Improving Power Through Planned Design

- Evaluate the validity of inferences for the planned design
- Design to address most important potential study weaknesses
- Realistic appraisal of study purpose, context, and odds of success
- Importance of fostering better understanding of the factors influencing power
- Planning that tailors design to study context and setting
 - Strategies for cluster recruitment
 - Prevention of missing data
 - Planning for use of realistic designs and use of other strategies like blocking, matching, and use of covariates

126

Design For Statistical Power

- Stronger treatments!
- Treatment fidelity
- Blocking and matching
- Repeated measures
- Focused tests ($df = 1$)
- Intraclass correlation
- Statistical control, use of covariates
- Restriction of range (IV and DV)
- Measurement reliability and validity (IV and DV)

127

Bibliography

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park: Sage.
- Bloom, H. S. (2006). *Learning More from Social Experiments: Evolving Analytic Approaches*. New York, NY: Russell Sage Foundation Publications.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Erlbaum.
- Darlington, R. B. (1990). *Regression and linear models*. New York: McGraw-Hill. Fox, J. (1991). *Regression diagnostics*. Thousand Oaks: Sage.
- Goldstein, H. (1995). *Multilevel Statistical Models* (2nd ed.). London: Edward Arnold. Available in electronic form at <http://www.arnoldpublishers.com/support/goldstein.htm>.
- Hedeker, D. (2004). An introduction to growth modeling. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences*. Thousand Oaks, CA: Sage Publications.
- Hedges, L. V., & Hedberg, E. C. (in press). Intraclass correlation values for planning group randomized trials in education. *Educational Evaluation and Policy Analysis*.
- Hox, J. J. (1994). *Applied Multilevel Analysis*. Amsterdam: TT-Publikaties. Available in electronic form at <http://www.ies.ac.uk/multilevel/amabook.pdf>.

Bibliography

- Jaccard, J., Turrisi, R., & Wan, C. K. (1990). *Interaction effects in multiple regression*. Thousand Oaks: Sage.
- Kreft, I. G., & de Leeuw, J. (2002). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research* (3rd Ed.). Orlando, FL: Harcourt Brace & Company.
- Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials, *Psychological Methods*, 2(2), 173-185.
- Raudenbush, S. W., & Bryk, A. S. (2001). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park: Sage.
- Raudenbush, S.W., Bryk, A.S., Cheong, Y.F., & Congdon, R.T. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*. Chicago, IL: Scientific Software International.
- Rumberger, R.W., & Palardy, G. J. (2004). Multilevel models for school effectiveness research. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences*. Thousand Oaks, CA: Sage Publications.
- Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.
- Seltzer, M. (2004). The use of hierarchical models in analyzing data from experiments and quasi-experiments conducted in field settings. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences*. Thousand Oaks, CA: Sage Publications.
- Spybrook, J., Raudenbush, S., & Liu, X.-f. (2006). *Optimal design for longitudinal and multilevel research: Documentation for the Optimal Design Software*. New York: William T. Grant Foundation.