EDMS 769M (0101)
Introduction to Multilevel Modeling

Spring 2001
Thursday, 4:15 – 7:00
Benjamin Building, Rm. 3315
(3 credits)

Instructor
Bob Croninger
2110 Benjamin Building
Phone: 301.405.2927
E-mail: rc164@umail.umd.edu

Office Hours
Thursday, 2 - 4:00
& by appointment

Course Overview

General Background for Course. A major phenomenon of interest in educational research is the learning of individual students. Two educational realities – (1) that learning implies change in individuals over time but (b) learning occurs in organizational settings (classrooms, schools) – cause persistent methodological problems for quantitative research. One problem is how to measure change. Another involves how to accommodate the grouped nature of the phenomenon (often referred to as "the unit of analysis problem"). In each case, the structure of the data poses serious methodological issues that challenge the assumptions of traditional data analysis methods.

The study of education illustrates a clear example in this regard, and much educational research is flawed because of it. Similar methodological issues have been raised in other fields, including psychology, sociology, economics, business, health care, and public policy. Measuring change and estimating the effects on individuals of organizational practices are common phenomena in social science research. But it has not been until recently that we have been able to develop new methodological strategies for modeling complex phenomena that occur in hierarchical structures. Aided by advances in estimation theory and statistical software development, researchers across fields have developed a general set of strategies for modeling change and organizational effects. These strategies – often referred to as multilevel modeling – are the focus of this course. .

Methodological Focus. This course is meant to introduce you to multilevel methodology and to help you become skilled in its use. We begin by discussing the need for multilevel methods in survey research, as well as the circumstances under which these methods are

appropriate and necessary. Although are focus will be on education research (it's what I know best), I'll argue that multilevel methods have a wide range of uses relevant to social science research. I ask you to read articles that employ these methods, as well as more theoretical articles about methodological issues associated with multilevel methods. And I ask you to actually use multilevel statistical software (HLM) to analyze data. By the end of the course, I hope that you will understand the logic of multilevel modeling, some of its "standard" terminology and notation (to the extent that there is a standard), and how to actually use HLM to analyze nested survey data. I hope to facilitate your acquisition of these skills through lectures, labs, and homework assignments.

Course Prerequisites. Although course focuses on quantitative methodology, its philosophical orientation is more toward research applications than statistical derivations. I expect students to have a working knowledge of general linear models and be familiar with non-experimental multivariate methods, but a background in formal mathematical statistics is not required. I do expect, however, that students have more than cursory experience with statistical computing in a social science setting. Generally, this means that students have completed at least two semesters of graduate-level training in quantitative research methods and statistics, and have working skill and full understanding of analysis of variance (ANOVA) and ordinary least squares regression (with multiple independent variables). I also expect that students are reasonably skilled and comfortable with using statistical software packages, such as SAS or SPSS.

I'll provide a NELS database to use in class, as well as ensure students access to HLM software (version 4.02). You may use your own data for course assignments, providing the data has a hierarchical structure and is appropriate for multilevel analyses. A word of caution, though: preparing data can be a lengthy and time-consuming process and HLM has specific requirements regarding file structures and missing data. Although students are welcome to use their own data, please be aware that I will be of limited assistance in helping you with data that I do not know. If you have data that you would like to use for the class, please contact me to discuss their appropriateness for multilevel modeling and class assignments.

Goals & Objectives. My overall goal for the course is to introduce you to multilevel modeling – its use, logic, craft, and interpretation. By the end of the class, I want each of you to be able to:

- Formulate multilevel research questions that can be addressed using hierarchical linear modeling software (HLM).
- Pose and test hypotheses about the parameters of multilevel models, provide meaningful interpretations of results, and write clearly about the substantive meaning of these results.
- Examine critically the application of HLM to particular problems in light of its statistical assumptions.
- Conduct a meaningful two-level hierarchical linear model analysis, write up the results, and draw substantive conclusions from the analysis.
- Gain a reading knowledge of applications of hierarchical linear models beyond the standard two-level organizational model or growth curve model.

Organization

General Format. I teach the course in a workshop format, which alternates between lectures and labs. Both lectures and labs are required, as this is a "hands-on" course. Lectures and labs are informal, meant to provide opportunities for questions and discussions. A culminating activity in the course is the final project (both a paper and an oral presentation). The purpose of presenting students' "work in progress" is to provoke discussion around multilevel methods, as well as enhance your ability to communicate findings, rather than as a "test" of the quality of work. You may be able to develop your final projects into conference presentations or publications.

Lectures provide the theory and purpose of HLM, discussion of analyses using it, as well as a vehicle for understanding what is being done, why it might be done, and what conclusions may be drawn from current work. The labs offer "how-to" experiences, where students receive instruction about using the statistical software, HLM. Out-of-class work involves both readings and a few write-ups of computing projects, but a major bulk of students' time is to be spent in front of a computer terminal, learning to work with HLM and to prepare data and variables to use in HLM analyses. Students should plan on allocating time for actually working with data and performing analyses. It is my firm belief that the logic of HLM is best understood through its application to real data.

Course Requirements. Minimum requirements include attending all lectures and labs, completing assigned readings, and completing four written assignments. Written work involves reporting the results of three short and one longer data analysis project. There is also, as mentioned above, an oral presentation for the final assignment. Because the course is designed around learning to use HLM, all students (including auditors, if any) should complete the assignments, in order to learn the methodology, how to apply it to data, and to receive feedback on their work. The assignments are meant to be progressive. Thus, it is difficult (if not impossible) to do the later assignments without mastering the skills involved in the early ones.

The three smaller papers are typically 3-5 pages in length, not including attachments. Each reports the results of one of the major steps in performing multilevel analyses. Final projects, each focusing on a topic of personal interest, include both a written paper of 10-20 pages and a short presentation to the class (roughly 15 minutes, depending on class size). Students may work individually or in pairs on all assignments. The final project, and its presentation, represents the culminating experience of the course. Thus, students may find it useful to select a topic and follow through with the analysis of the topic in subsequent assignments. Format for the final project should resemble articles from scholarly journals that we review. I'll provide you with more detail about each assignment at least two weeks before an assignment is due.

Lab, Data, & Software. We'll talk more about this in class, but these are the basic facts. The University of Maryland has a site license for HLM Version 4.02. Although the most recent version of HLM (Version 5.02) includes more advances uses (e.g., multinomial logit), Version 4.02 is sufficient for the purposes of the class and most users. If you wish to purchase your own copy of the HLM 5 software, the cost is about $450, with additional charges for upgrades. There is also a free student version of HLM that can be downloaded from the company's website. (Unfortunately, the case restrictions for the student version are too limited for the purposes of the class). For additional information about HLM, contact: Scientific Software Incorporated (SSI), 7383 N. Lincoln Avenue, Suite 100, Lincolnwood, IL 60712-1704. Telephone: (847) 675-0720. Fax: (847) 675-2140. Website: www.ssicentral.com. Of course, it is strictly illegal to copy the

software.

The location of our labs will depend on where the University loads the software. Data for the course come from a nationally representative, random sample of students and schools that participated in the first two waves of NELS:88. This includes data about roughly 9,000 students nested in 800 high schools between 1988 and 1990. The class dataset is in two SPSS systems files (one for students, one for schools). Students who wish to use SAS will need to convert these files. It is usually necessary to do some data manipulation before enter data into HLM. The labs will include only a modest amount of direct instruction about how to use SPSS (and none about SAS). I assume that students have already gained many skills in statistical computing and general linear model techniques in previous courses or work. Although I will assist you with discussion about weighted data, recoding data, transformations, and missing data issues, the focus of the labs will be on HLM and not SPSS or any other statistical software package.

Materials

Required Books. There are two texts required for the course. The first book is the introductory text for HLM written by Tony Bryk and Steve Raudenbush. Although Bryk and Raudenbush will undoubtedly revise the text at some point, this book has been the standard for those using HLM since its publication in 1992. The second text is the manual for HLM 5. If you would prefer the manual for HLM 4, I can provide you with a copy.

- Bryk, A.S. & Raudenbush, S. (1992). Hierarchical Linear Models for Social and Behavioral Research: Applications and Data Analysis Methods. Newbury Park, CA: Sage Publications.

- Raudenbush, S.W., Bryk, A.S., Cheong, Y.F., & Congdon, R.T. (2000). HLM 5: Hierarchical Linear and Nonlinear Modeling. Lincolnwood, IL: Scientific Software International.

Additional Books. Multilevel modeling has become a bit of a "growth industry," so not surprisingly, there are quite a few other books that have been written as introductory texts. The two that I like best are listed below. Neither is required for the course, but each provides a slightly different perspective on multilevel methods than the one presented by Bryk and Raudenbush.

- Kreft, I., & De Leeuw, J. (1998). Introducing Multilievel Modeling. London: Sage.

- Snijders, T., & Bosker, R. (1999). Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. London: Sage.

Additional Readings. Because the class is small, I decided not to create a course pack. Instead, I will bring in a copy of the readings several weeks in advance and you can make copies yourself. In the past, students have found it convenient to divide all of the articles among themselves so as to create their own course packs. We can discuss the possibility of doing so the first week of class.

Course Schedule

The topics and reading assignments listed for each class are approximate and based on my expectations regarding pace and coverage of materials. Although the progression will be the same as listed in the syllabus, the days on which each topic is covered depends, to some extent, on how the course progresses. Assignment deadlines will not change, however. The schedule describes assignments on the date that they are due.

Week 1, February 1st

Orientation and review of the syllabus. We'll spend the first class reviewing the course requirements and establishing general guidelines for assignments, lectures, and labs. Read the following for next week and be prepared to discuss the texts.

Burstein, L. & Miller, M.D. (1981). Regression-based analyses of multilevel educational data. In Borusch, R.F., Wortman, P.M.,Corday, D.S. and Associates. Reanalyzing Program Evaluations. San Francisco: Jossey-Bass.

Haney, W. (1980). Units and levels of analysis in large-scale evaluation. In New Directions for Methodology of Social and Behavioral Science, 6, 1-15.

Bidwell, C.E. & Kasarda, J.D. (1980). Conceptualizing and measuring the effects of school and schooling. American Journal of Education, 88, 401-430.

Week 2, February 8th

Why do we need multilevel models? We will focus on the need for multilevel models during the first part of the course. More specifically, we'll discuss the "unit of analysis problem," what we sacrifice when we don't consider the nested nature of data, and the need for statistical methods that can model nested data. Read the following for next week.

Chapter 1 & 2 of Bryk and Raudenbush (1992). Hierarchical Linear Models: Applications and Data Analysis Methods.

Week 3, February 15th

Logic of HLM. We'll discuss the basic logic of HLM and how it addresses issues associated with the "unit of analysis problem." We'll also explore a series of simpler models that can be derived from the basic two-level model. Read the following from the HLM 5 manual. It will provide reasonable background for next week's lab.

Read pages 1-41 of the HLM 5 manual.

Week 4, February 22nd

Lab: Introduction to HLM Software. I'll introduce you to the HLM software and discuss how to build a Sufficient Statistics Matrix (SSM), the basic system's file for HLM analyses. We'll examine how to create an SSM using the Transys interface program and ascii records. We'll also specify the fully unconditional model and calculate the ICC (Type I) and lambda (ICC-Type II). Read the following for next week.

Chapter 3 & 4 of Bryk and Raudenbush (1992). Hierarchical Linear Models: Applications and Data Analysis Methods.

Arnold, C. (1992). An introduction to hierarchical linear models. Measurement and Evaluation in Counseling and Development, 25 (July), 58 – 90.

Week 5, March 1st

Model building strategies. We'll continue our discussion of the logic of HLM, focusing on modeling building strategies. We'll discuss statistical estimation, the consequences of centering, and decisions associated with level-1 models. Read the following for next week.

Review HLM 5 Manual (1-41).

Kreft, I. G., de Leeuw, J., & Aiken, L. (1995). The effect of different forms of centering in hierarchical linear models. Multivariate Behavioral Research, 30, 1-22.

Cohen, J., & Cohen, P. (1983). Missing data. In Applied multiple regression/correlation analysis for the behavioral sciences (Chapter 7, pp. 275-300). Hillsdale, NJ: Lawrence Earlbaum

Week 6, March 8th

Lab: Building the within-group model. We'll discuss strategies for specifying a level-1 model using the NELS:88 data. I'll review software protocol and demonstrate how to build a level-1 model with a random intercept and at least one random slope. Note that your first written assignment is due today. For next week read the following.

Chapter 5 of Bryk and Raudenbush (1992). Hierarchical Linear Models: Applications and Data Analysis Methods.

Hofman, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. Journal of Management, 24 (5): 623-641.

Assignment #1 due at the beginning of this class: Computer printout (SPSS-X [or other statistical program] command file for writing out data, log file for HLM run), which illustrates your ability to read data into HLM and create a sufficient statistics file. Also include HLM output showing you know how to create an HLM file. Please write a brief guide to this output, documenting the procedure that you have followed and the logic involved in the procedure. If your choice of variables was guided by a research question, please spell it out. Add a cover page on all assignments -- including a substantive title, your name, course, and date.

Week 7, March 15th

Applications of the two-level model: Intercepts- and slopes-as outcomes. We'll shift our focus for class lectures to examining actual applications of HLM. We'll begin with the examples in the Bryk and Raudenbush text on organizational studies. We'll consider the estimation of

contextual effects, as well as compare HLM results with results from more traditional analytic methods. Read for next week.

Lee, V. E., & Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. Sociology of Education, 62 (2), 172-192.

Rowan, B., Raudenbush, S. W., & Kang, S. J. (1991). Organizational design in high schools: A multilevel analysis. American Journal of Education, (February), 238 – 266.

Raudenbush, S.W., & Willms, J.D. (1995). The estimation of school effects. Journal of Educational and Behavioral Statistics, 20(4), 307-337.

Week 8, March 22nd

Spring break, No class.

Week 9, March 29th

Applications of the two-level model continued: Intercepts- and slopes-as outcomes. We'll continue our examination of HLM applications. This week we'll focus on the theory of estimating school effects, as well as two early articles written as primers for investigating organizational effects. Read the following for week 10.

Willett, J.B. (1994). Measuring change more efficiently by modeling individual growth over time. In T. Husen & T.N. Postlehwaite (Eds.) The International Encyclopedia of Education (2nd Edition). Oxford, UK: Pergamon Press.

Bryk, A.S. & Raudenbush, S.W. (1987). Application of hierarchical linear models to assessing change. Psychological Bulletin, 101(1), 147-158.

Chapter 6 & 9 of Bryk and Raudenbush (1992). Hierarchical Linear Models: Applications and Data Analysis Methods.

Assignment #2 due at the beginning of the class: Within-group HLM data analysis and write-up. Include at least 3 independent variables in the within-group model. Between group model should be unconditional. You should write up what you have done, along with the substantive meaning and interpretation of results. Be sure to compute the intraclass correlation, discuss its substantive meaning, and discuss possible second-stage outcomes. Again, please put a cover page on this assignment, with a title, your name, and the date. I would prefer that you make and number tables from your HLM output (or read the output into your text), and refer to tables or output in your write-up. If you choose to include HLM output instead of making tables, then you should edit it down to include only the meaningful output.

Week 10, April 5th

Lab: Building the between -group model. We'll discuss strategies for specifying a level-1 model using the NELS:88 data. I'll review software protocol and demonstrate how to build a

fully conditional model, including assessing the model for "fit" and "adequacy". If possible we will also discuss how to use two-level models to assess change. Read the following for next week.

Chapter 8 of Bryk and Raudenbush (1992). Hierarchical Linear Models: Applications and Data Analysis Methods.

Skim Chapter 3 & 4 in the HLM 5 Manual.

Raudenbush, S.W., Rowan, B., & Cheong, Y-F. (1993). Higher-order instructional goals in secondary schools: Class, teacher, and school influences. American Educational Research Journal, 30(3), 523-553.

Ungraded assignment due at the beginning of the class: One-page prospectus of plan for final project. Include research question, data, and some description of HLM usage. Projects may be done in pairs or individually. Try to discuss your project with me beforehand, to make sure the idea is appropriate for HLM and feasible within the timeframe.

Week 12, April 12th

AERA, No class.

Week 13, April 19th

Introduction to the study of change. We'll discuss the study of change and how multilevel modeling strategies address methodological problems associated with longitudinal studies of change. Read for next week the following.

Review pp. 1-72; skim Chapter 5 & 6 in the HLM 5 Manual.

Rumberger, R.W. (1995). Dropping out of middle school: A multi-level analysis of students and schools. American Educational Research Journal, 32(2), 583-625.

Rumberger, R. W., & Thomas, S. L. (2000). The distribution of dropout and turnover rates among urban and suburban high schools. Sociology of Education, 73, 39-67.

Assignment #3 due: Prepare and write up an HLM model which includes no more than 3 significant group-level predictors. You should model the intercept, and if you feel comfortable, try modeling one slope as an outcome. Be sure you understand the meaning of the effects of group predictors on slopes, and describe this in your write-up. Attach relevant computer output. You may work individually or in pairs. It is a good idea to think of this assignment as a test run for your final project idea. Form of write-up should follow that described for Assignment #2.

Week 14, April 26th

Advanced methods: Introduction to three-level models and dichotomous outcomes. We'll expand the notion of two-level models to include studies that examine phenomena across three

units of analysis (e.g., students within teachers within schools) and examinations of phenomena measured as dichotomous outcomes. We'll look at published articles that use these techniques, and we will discuss some of the analytic decisions required when you use each technique.

Week 15, May 3rd

Presentation of final projects. Students will present their final projects in class. I'd like to a lot about 15 minutes per presentation with the expectation of another 15 minutes of class discussion. I'll provide a data projector for anyone who wants to do power point presentation. Final projects due next week.

Week 16, May 10th

Presentation of final projects (continued) and wrap up for class. We'll finish the class presentations, evaluate the course, and discuss future analysis projects using HLM.

Assignment #4 due at the beginning of class: Final projects due. This is a final write-up of your work for the class, which may be an expansion of your work for Assignment 3. If you use citations to support the theory on which your work rests, please use APA or ASA format, and include all citations. Cite course readings as appropriate. Presentation and paper should include tables and graphs as appropriate.