Political Science 606
Topics in Research Methods: Hierarchical Modeling

Summer Session, 2004

Monday, Wednesday, 2:00-4:00 (To be changed)

Instructor: Dave Peterson
Office: 2037 Bush Academic Building
Phone: 845-6783
Email: dave@polisci.tamu.edu

The best way to get a hold of me is through email. Please keep in mind that I am unlikely to check my email much later than 8:00 PM on any given night. If you try to contact me after that time, assume that I will not read my mail until the next morning.

Description: Hierarchical, or multi-level, models are techniques designed to handle data that are collected at two or more levels. In this class of models, the dependent variable is measured at one level of analysis and at least some of the independent variables are measured at some higher level of aggregation. The classic example is school settings, where some student level measure is believed to be a function of some individual characteristics (e.g. race or gender), some class level (teacher), some school level (principal), some district level (superintendent) and some state level measures. This type of data produces several difficulties that make standard linear models (OLS or any other OLS derivative such as logit/probit) inappropriate. While lots of data in political science fit this description, there has been little use of the appropriate models to analyze them. Some of this is due to the lack of sound statistical theory, some due to the absence of software, and some sue to the limitations of computation. In the past decade all of these problems have been more or less solved.

Student Responsibilities: Students should come to class having read the material assigned for that day. There will be two sessions each week. One (three hours onWednesday) will be in the classroom covering the readings. The other (one hour on Monday) will be spent in the lab learning how to get the software to do what you want it to. Your grade in the class will be based on class involvement, a number (to be determined) of homework assignments, and a final project.

Readings: I am a believer in multiple approaches to topic, especially in statistics. Thus I am assigning several books in an effort to help you find the one that works best for you. You are only required to buy two books (two others are available for free). The main readings for the class are

Raudenbush, Stephen W. and Anthony S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods, second edition.* Thousand Oaks, CA: Sage

Snijders, Tom A. B. and Roel J. Bosker 1999. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling.* Thousand Oaks: Sage

I am also assigning electronic copies of two other books that are available on line (I will put them on the course WebCT page).

Goldstein, Harvey. 1999. *Multilevel Statistical Models.* Internet edition.

Hox, J. J. 1995. *Applied Multilevel Analysis.* Amsterdam: TT-Publikaties.

There will also be a number of articles and chapters assigned for specific topics.

Project: You have three choices for your final project.

1. An original data project. Develop and carry out an original research paper using hierarchical models

2. A meta analysis. Meta analysis (which we will be learning) is the technique of combining multiple previously published studies on the same topic into a single model to test the robustness of the results. While this will not require as much data collection as option 1, you will need to find an area that has a large number of studies which can be combined.

3. A replication/correction. This is probably the easiest of the options and will require my permission before you can do it. I expect that the more senior graduate students will not take this option.

I am flexible on the due date of the project. This is a ten week seminar, not the standard 15 weeks of a semester. My plan is to give students longer to finish the projects if they want, with the expectation that they will be completed by the start of Fall semester.

Software: The department has purchased several copies of HLM for this class. HLM is one of a handful of specialized software to run hierarchical models (it is possible to run some of these models in either SAS or SPSS, but HLM is more flexible). There are two versions of the software available: the full version (which will be installed on several machines in the lab) and the student version. The student version is freely available for download and you can install it on your personal machines (I have also asked the computer folks to install it on office machines if people ask for it). The main difference between the student version and the full version is that the student version is limited in the number of variables you can include. I don't think this will be a problem for the homework assignments, but it might be for your projects. Students auditing the course will be expected to use machines with the student version during the lab sessions.

Topics

I have not taught this course before and am unsure of how quickly or slowly I will move through the specific topics. Thus, I am not assigning topics to a specific week, but am assigning them in a general order. I will try to communicate clearly about where we are and where we are going. The initials listed in the topics refer to the text books. The numbers are the chapters.

1. Logic and previous treatments

(R&B 1-2; S&B 1-5; H 1-2; G 1, 2)

Additional reading:

Freidrich, Robert J. 1982. In Defense of Multiplicative Terms in Multiple Regression Equations. *American Journal of Political Science.* 26:797-833.

Robinson, W. S. 1950. Ecological Correlations and the Behavior of Individuals. *American Sociological Review*. 15(3): 351-357.

Steenbergen, Marco R. and Bradford S. Jones. 2002. Modeling Multilevel Data Structures. *American Journal of Political Science.* 46(1): 218-237.

Suggested: There are a number of introductory pieces out there that may be of some additional help:

Raudenbush, Stephen and Anthony S. Bryk. 1986. A Hierarchical Model for Studying School Effects. *Sociology of Education.* 59(1): 1-17.

De Leeuw, Jan and Ita Kreft. 1986. Random OCefficeint Models for Multilevel Analysis. *Journal of Educational Statistics* 11(1): 57-85.

Hedeker, Doland, Robert D. Gibbons, and Brian R. Flay. 1994. Random-Effects Regression Models for Clustered Data with an Example from Smoking Prevention Research. *Journalof Consulting and Clinical Psychology*. 62(4): 757-765.

2. Principles of estimation
(R&B 3 S&B 4-6)
Additional reading:

Kreft, Ita G. G., Jan de Leeuw, and Leona S. Aiken. 1995. The Effect of Different Forms of Centering in Hierarchical Linear Models. *Multivariate Behavioral Research.* 30(1): 1-21.

3. Examples of the estimations.
(R&B 4, 5)

4. Individual change—growth models.
(R&B 6, G 6)

Plutzer, Eric 2002. Becoming a Habitual Voter: Inertia, Resources, and Growth in Young Adulthood. *American Political Science Review*. 96(1): 41-56.

5. Meta analysis
(R&B 7, H 4.1)

Lau, Richard R., Lee Sigelman, Caroline Heldman, and Paul Babbit. 1999. The Effects of Negative Political Advertisements: A Meta-Analytic Assessment. *American Political Science Review*. 93(4): 851-876.

Hox, Joop J. and Edith D. de Leeuw. 2003. Multilevel Models for Meta-Analysis in Steven P. Reise and Naihua Duan eds. *Multilevel Modeling: Methodological Advances, Issues, and Applications.* Mahwah, NJ: Lawrence Erlbaum.

The following is available as an ebook through the library and is a decent introduction to meta-analysis:

Wolf, Frederic M. 1986. *Meta Analysis: Quantitative Methods for Research Synthesis*. Thousand Oaks, CA: Sage.

6. Three (or more) level models
(R&B 8)

7. Goodness of fit
(R&B 9)

8. HGLM
(R&B 10; H 4.2; G 7; G 9; S&B 14)
Rudolph, Thomas J. 2003. Institutional Context and the Assignment of Political Responsibility. *Journal of Politics*. 65(1): 190-215.

9. Latent models
(R&B 11; H 5)

10. Cross categorization
(R&B 12; G 8; S&B 11)

11. MCMC
(R&B 13)

Jackman, Simon 2000. Estimation and Inference via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo. *American Journal of Political Science*. 44(2): 375-404.
Western, Bruce. 1998. Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modeling Approach. *American Journal of Political Science.* 42(4): 1233-1259.

12. Estimation theory
(R&B 14)

13. Complex variance functions
(G 3)

I expect that all students will conduct themselves in a manner that is consistent with the Aggie Code. Any lying or cheating in this class will be handled in accordance with Texas A&M policy.

ADA Statement:

The Americans with Disabilities Act (ADA) is a federal anti-discrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities. If you believe you have a disability requiring an accommodation, please contact the Department of Student Life, Services for Students with Disabilities in Room 126 of the Koldus Building, or call 845-1637.

Plagiarism Statement

As commonly defined, plagiarism consists of passing off as one's own the ideas, words, writings, etc., which belong to another. In accordance with the definition, you are committing plagiarism if you copy the work of another person and turn it in as your own, even if

you should have the permission of the person. Plagiarism is one of the worst academic sins, for the plagiarist destroys the trust among colleagues without which research cannot be safely communicated. If you have any questions regarding plagiarism, please consult the latest issue of the Texas A&M University Student Rules, under the section “Scholastic Dishonesty.”