

## **POLITICAL SCIENCE 676:02: MULTILEVEL MODELING**

Department of Political Science • Stony Brook University  
Spring 2009 • Wednesdays, 2:00-5:00 • SBS N-702

Professor Brandon Bartels

Email: [brandon.bartels@stonybrook.edu](mailto:brandon.bartels@stonybrook.edu)

Office: SBS S-713

Office phone: 632-7452

Office hours: Open-door policy for grad students

Webpage: <http://ms.cc.sunysb.edu/~bbartels>

*Note:* I will post course materials, data, and announcements on Blackboard.

In this course, you will master the foundations of a methodology consisting of a class of models falling under the “multilevel modeling” umbrella. Multilevel data structures consist of more than one level of analysis, with each level consisting of different units of analysis. In the most common multilevel structure, data are hierarchical and consist of a nesting structure. In a two-level hierarchical structure, units of analysis from the lowest level (level one) are nested within units from level two (for example, voters [level-one units] nested within congressional districts [level-two units]). In a three-level model, a third level of analysis is present, within which level-two units are nested (e.g., voters nested within congressional districts nested within states [level-three units]). We will examine additional structures such as cross-classified data, where the data are non-nested and non-hierarchical. In addition, many types of well-known data structures are special forms of multilevel structures, including panel data, time series cross sectional (TSCS) data, and item response models.

Importantly, all varieties of multilevel models (including those discussed above) share common types of statistical and methodological issues, and moreover, they allow unique opportunities for modeling various types of political and behavioral processes. We will examine a wide range of multilevel modeling topics in this course (see the course schedule on p. 5).

### **REQUIRED AND RECOMMENDED BOOKS**

The following book is required for the course. I will not place orders of the books at the University bookstore, as you will be much better off purchasing these online.

Rabe-Hesketh, Sophia, and Anders Skrondal. 2008. *Multilevel and Longitudinal Modeling Using Stata* (Second Edition). Stata Press. [Available for purchase at [www.statapress.com](http://www.statapress.com)]

We will also draw on material from the following textbooks. I will make copies of portions of these books available to allow students to make copies on their own.

Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Raudenbush, Stephen, and Anthony Bryk. 2002. *Hierarchical Linear Modeling*. Sage.

Skrondal, Anders, and Sophia Rabe-Hesketh. 2004. *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. Chapman & Hall/CRC.

If I had to recommend one of these books to purchase, I would recommend Gelman and Hill.

## ADDITIONAL READINGS

We will also draw upon some of the following articles and books throughout the course:

- Bafumi, Joseph, and Andrew Gelman. 2006. "Fitting Multilevel Models When Predictors and Group Effects Correlate." Presented at the Annual Meeting of the MPSA.
- Bafumi, Joseph, Luke Keele, David K. Park, and Boris Shor. 2007. "Examining Time Series Cross Sectional Data with Bayesian Multilevel Models." *Political Analysis*.
- Beck, Nathaniel. 2001. "Time-Series—Cross-Sectional Data: What Have We Learned in the Past Few Years?" *Annual Review of Political Science* 4:271-93.
- Bowler, Shaun, Todd Donovan, and Robert Hanneman. 2003. "Art for Democracy's Sake? Group Membership and Political Engagement in Europe." *JOP* 65:1111-29.
- Cho, Wendy K. Tam, James G. Gimpel, and Joshua J. Dyck. 2006. "Residential Concentration, Political Socialization, and Voter Turnout." *JOP* 68:156-67.
- Congdon, Peter. 2003. *Applied Bayesian Modelling*. Wiley.
- Congdon, Peter. 2005. *Bayesian Models for Categorical Data*. Wiley.
- Gelman, Andrew, John Carlin, Hal Stern, and Donald Rubin. 2003. *Bayesian Data Analysis*. Chapman & Hall/CRC.
- Gelman, Andrew, and Iain Pardoe. 2006. "Bayesian Measures of Explained Variance and Pooling in Multilevel (Hierarchical) Models." *Technometrics* 48:241-51.
- Gelman, Andrew, and Iain Pardoe. 2007. "Average Predictive Comparisons for Models with Nonlinearity, Interactions, and Variance Components." *Sociological Methodology* (Forthcoming).
- Heckman, James J. 1981. "Heterogeneity and State Dependence." In *Studies in Labor Markets*, ed. S. Rosen. Chicago, IL: University of Chicago Press.
- Hsiao, Cheng. 2003. *Analysis of Panel Data* (Second Edition). Cambridge.
- Martin, Andrew D. 2001. "Congressional Decision Making and the Separation of Powers." *APSR* 95:361-78.
- Park, Jong Hee, and Nathan Jensen. 2007. "Electoral Competition and Agricultural Support in OECD Countries." *AJPS*.
- Rabe-Hesketh, Sophia, Anders Skrondal, and Andrew Pickles. 2002. "Reliable Estimation of Generalized Linear Mixed Models Using Adaptive Quadrature." *Stata Journal* 2:1-21.
- Rahn, Wendy M., and Thomas J. Rudolph. 2005. "A Tale of Political Trust in American Cities." *POQ* 69:530-60.
- Rodriguez, German, and Noreen Goldman. 2001. "Improved Estimation Procedures for Multilevel Models with Binary Response: A Case-Study." *Journal of the Royal Statistical Society, Series A* 164:339-55.
- Rudolph, Thomas J. 2003. "Institutional Context and the Assignment of Political Responsibility." *JOP* 65:190-215.

- Sastry, Narayan. 1997. "A Nested Frailty Model for Survival Data, With an Application to the Study of Child Survival in Northeast Brazil." *JASA* 92:426-35.
- Steenbergen, Marco, and Bradford Jones. 2002. "Modeling Multilevel Data Structures." *AJPS* 46:218-37.
- Western, Bruce. 1998. "Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modelling Approach." *AJPS* 42:1233-59.
- Western, Bruce, and Simon Jackman. 1994. "Bayesian Inference for Comparative Research." *APSR* 88:412-23.
- Wilson, Sven, and Daniel M. Butler. 2007. "A Lot More to Do: The Sensitivity of Time-Series Cross-Section Analyses to Simple Alternative Specifications." *Political Analysis* .
- Wooldridge, Jeffrey M. 2005. "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity." *Journal of Applied Econometrics* 20:39-54.
- Zorn, Christopher. 2000. "Modeling Duration Dependence." *Political Analysis* 8:367-80.

## SOFTWARE

We'll rely on two software programs: *Stata 10* and *WinBUGS*. We will use Stata's built-in commands for estimating multilevel models (the "xt" family and the "xtmixed" command) as well as *GLLAMM*, which is an add-on package to Stata (created by Sophia Rabe-Hesketh) that estimates multilevel models. All of the worked examples in the Rabe-Hesketh and Skrondal book (the second one in the required readings list) use Stata. *WinBUGS*, which can be downloaded for free, estimates models via the Bayesian / MCMC approach (Gibbs sampling). While we'll primarily rely on these two programs, I will show you differences and similarities between these programs and others such as *R*.

## REQUIREMENTS (500 points total)

- 1. Class participation** (50 points): Students are expected to complete all of the assigned readings, to attend each class, and to participate in class sessions.
- 2. Problem sets:** Three major problem sets (100 points each) will constitute a supermajority of your responsibilities. Each problem set will require you to demonstrate your understanding of the material and the ability to make appropriate interpretations. The due dates, also noted in the Course Outline section of the syllabus, are as follows:

**Problem Set 1:** Wednesday, March 4

**Problem Set 2:** Wednesday, April 1

**Problem Set 3:** Tuesday, May 12

Two important notes on problem sets:

- All problem sets will be due at *the beginning of class* on the day they are due. *I will not accept late problem sets*. Also, these are in-depth, intense assignments, so you will not want to wait until the last minute to start them. I strongly suggest working on the assignments progressively as we get through the relevant material.

- Feel free to work with your classmates on the problem sets. Collaboration can be beneficial for mastering the material. However, *you must do your own work*. That is, while you can work together, *the final product that you hand in must be your own work*.

**3. Class presentation** (50 points): Each student will give a brief presentation to the class on a multilevel modeling topic, preferably one in which you have a particular interest. Students can present on an article that performs a multilevel modeling application or on a methodological article that explores an advanced topic. Please consult with me about the article or reading you would like to present on. I would like these presentations to commence around week 3 or 4. I would like to spread them out throughout the semester so please let me know your topic as soon as you can so that I can plan accordingly. The presentation, and associated discussion, should be around 30-45 minutes, so you should plan a presentation that is about 15-20 minutes.

**4. Final Paper** (100 points): Due *Tuesday, May 19*. For the final paper, you will use the skills you have learned throughout the class to analyze and write up results from a multilevel model (or models) using your own data. Importantly, the paper project is also intended to give you practice toward mastering the art of writing a research paper (in particular, discussing results and how they relate back to the substantive research question(s), the theory, and hypotheses). The paper should be roughly 12-15 pages (double-spaced) and it should resemble the second half of a journal article. The paper will be written as if the intro, literature assessment, and theory sections have already been written. Use a small amount of space giving me a rough outline of the research question, literature assessment, and the theory. List the hypotheses to be tested, and then write up the data and results section as you would do for a journal article. I strongly encourage you to think about your research question very early in the semester and to make progress on the paper gradually throughout the semester. Students should touch base with me about their papers—particularly regarding the data they are interested in analyzing—sometime before the middle of the semester.

#### **AMERICANS WITH DISABILITIES ACT**

If you have a physical, psychological, medical or learning disability that may impact your course work, please contact Disability Support Services, ECC (Educational Communications Center) Building, room 128, (631) 632-6748. They will determine with you what accommodations are necessary and appropriate. All information and documentation is confidential.

#### **ACADEMIC INTEGRITY**

Each student must pursue his or her academic goals honestly and be personally accountable for all submitted work. Representing another person's work as your own is always wrong. Faculty are required to report and suspected instances of academic dishonesty to the Academic Judiciary. For more comprehensive information on academic integrity, including categories of academic dishonesty, please refer to the academic judiciary website at <http://www.stonybrook.edu/uaa/academicjudiciary/>.

#### **CRITICAL INCIDENT MANAGEMENT**

Stony Brook University expects students to respect the rights, privileges, and property of other people. Faculty are required to report to the Office of Judicial Affairs any disruptive behavior

that interrupts their ability to teach, compromises the safety of the learning environment, or inhibits students' ability to learn.

## **COURSE OUTLINE AND SCHEDULE**

Note: Below, I include the general topics that we'll cover for each session. I will announce readings the week before each topic.

1. **Jan. 28:** Course introduction and introduction to multilevel modeling
2. **Feb. 4** (*Note: We will meet after the job talk, at around 2:15*): Pooling, unobserved heterogeneity, causal heterogeneity, and general model specification
3. **Feb. 11:** The random intercept model (aka, random effects) and alternatives (e.g., the fixed effects model); specification and estimation
4. **Feb. 18:** Shrinkage and pooling in the random intercept model
5. **Feb. 25:** Estimation and interpretation of the random intercept model; introduction to the random coefficient model
6. **Mar. 4:** Estimation and interpretation of the random coefficient model; maximum likelihood and Bayesian estimation  
*\*\*Problem Set 1 Due Wednesday, March 4*
7. **Mar. 11:** Continue estimation and interpretation of the random coefficient model (ML and Bayesian approaches)
8. **Mar. 18:** Model fit, model comparison, computing substantive quantities of interest (ML and Bayesian approaches)
9. **Mar. 25:** Continue model fit, model comparison, computing substantive quantities of interest (ML and Bayesian approaches)
10. **Apr. 1:** Multilevel models for binary responses  
*\*\*Problem Set 2 Due Wednesday, April 1*  
*\*\*No class on Apr. 8 – Spring Break*
11. **Apr. 15:** Multilevel models for ordinal, count, and nominal responses
12. **Apr. 22:** Multilevel models for panel and time-series cross-sectional (TSCS) data
13. **Apr. 29:** Multilevel models for duration data
14. **May 6:** Cross-classified, non-nested models  
*\*\* Problem Set 3 due Tuesday, May 12*  
*\*\* Final Paper due Tuesday, May 19*