

Bayesian Modeling for the Social Sciences I: Introduction and Application

ICPSR Summer Program 2019

Ryan Bakker (Weeks 1–2)

Associate Professor

Department of Political Science

University of Georgia

rbakker@uga.edu

Johannes Karreth (Weeks 3–4)

Assistant Professor

Department of Politics & International Relations

Ursinus College

jkarreth@ursinus.edu

This syllabus will be updated prior to the course. Please visit www.jkarreth.net/bayes-icpsr.html for the most recent information on topics, labs, rooms, and assigned readings.

Workshop: 9am–11am / Room: See the Summer Program handbook.

Office: 325 Helen Newberry Building

Office hours: 1pm–2pm, M–F. Feel free to stop by the office any time and come in if the door is open. We're also happy to schedule meetings at most other times during the day.

Teaching assistants

Neil Williams

University of Georgia

snpwill@uga.edu

Office: 324 Helen Newberry Building

Office hours: 12pm–1pm

Myunghee Lee

University of Missouri - Columbia

mlq38@mail.missouri.edu

Office: 326 Helen Newberry Building

Office hours: 3pm–4pm

Course description

This workshop introduces the basic theoretical and applied principles of Bayesian statistical analysis in a manner geared toward students and researchers in the social sciences. The Bayesian paradigm is particularly useful for the type of data that social scientists encounter given its recognition of the mobility of population parameters, its ability to incorporate information from prior research, and its ability to update estimates as new data are observed. The workshop begins with a discussion of the strengths of the Bayesian approach for social science data and the philosophical differences between Bayesian and frequentist analyses. Next, the workshop covers the theoretical underpinnings of Bayesian modeling and provides a brief introduction to the primary estimation algorithms. The bulk of the workshop focuses on estimating and interpreting Bayesian models from an applied perspective. Participants are introduced to the Bayesian forms of the standard statistical models taught in regression and MLE courses (i.e., linear, logit/probit, poisson, etc.). Additional topics include measurement models, model comparison, and an in-depth treatment of multilevel modeling.

Participants should have a solid understanding of the linear model and matrix algebra and some exposure to models with limited dependent variables. The **Mathematics for Social Scientists III** lecture with Don Eckford is an ideal companion lecture for this workshop.

The course relies mostly on R and WinBUGS/JAGS for estimation. Prior experience with R is preferred but not assumed. We offer lab sessions to familiarize participants with R as well as WinBUGS, JAGS, and (briefly) Stan. No prior experience with software for Bayesian estimation is necessary.

Goals. Upon conclusion of this workshop, we aim for participants to be able to:

- appreciate the fundamental differences and similarities between frequentist and Bayesian approaches to inference
- apply Bayes' rule to statistical inference
- formulate linear and generalized linear models in the Bayesian framework
- estimate linear and generalized linear models in the Bayesian framework using customizable code
- exploit the advantages of Bayesian estimation with regard to
 - incorporating prior information
 - estimating uncertainty in parameter estimates
 - dealing with missing data
 - measuring latent concepts
 - incorporating variance at multiple levels of observation
- present and communicate results from Bayesian (and frequentist) estimation in an effective manner
- have fun learning new methods and better understanding familiar ones!

A note on computing. This workshop mostly uses JAGS and WinBUGS (and its open-source twin OpenBUGS) as the primary software options to fit Bayesian models, with one unit toward the end dedicated to Stan. We access JAGS and Win/OpenBUGS through R. Most lectures build on JAGS and Win/OpenBUGS. The languages of these two programs are nearly identical. Win/OpenBUGS run on Macs only with Windows emulation software, but can be a bit buggy. JAGS runs on all platforms, including Macs. We offer special Mac-friendly lab sessions and support both JAGS and Win/OpenBUGS. JAGS code for all models encountered in this workshop and other JAGS-specific code and examples are provided.

Workshop resources

M-Box: All slides, code used in workshop sessions, and problem sets will be posted on M-Box. Details will be provided on the first day of the workshop. The TAs are also always available to point you to these materials.

Google Group: We will ask you to sign you up for a Google Group for this workshop. We will use this Google Group to share information related to the workshop during the ICPSR Summer Program.

Website with additional materials: Additional code, tutorials, and other materials for weeks 3–4 are posted on Johannes' website: <http://www.jkarreth.net/bayes-icpsr.html>.

Reading materials

Books

The main required texts for this workshop are:

- Gelman, Andrew and Hill, Jennifer. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, NY: Cambridge University Press.
- Gill, Jeff. 2014. *Bayesian Methods: A Social and Behavioral Sciences Approach, Third Edition*. Boca Raton, FL: Chapman / Hall/CRC.

You may also find the following titles useful for many of the topics discussed in this workshop. They are available in the ICPSR Summer Program Library for borrowing:

- Congdon, Peter D. 2003. *Applied Bayesian Modelling*. Chichester: Wiley.
- Congdon, Peter D. 2010. *Applied Bayesian Hierarchical Methods*. Boca Raton, FL: Chapman / Hall/CRC.
- Gelman, Andrew et al. 2013. *Bayesian Data Analysis, Third Edition*. Boca Raton, FL: Chapman & Hall/CRC.
- Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. Chichester: Wiley.
- Kruschke, John. 2014. *Doing Bayesian Data Analysis, Second Edition: A Tutorial with R, JAGS, and STAN*. Oxford: Academic Press / Elsevier.
- Lunn, David et al. 2012. *The BUGS Book: A Practical Introduction to Bayesian Analysis*. Boca Raton, FL: Chapman / Hall/CRC.
- Ntzoufras, Ioannis. 2009. *Bayesian Modeling Using WinBUGS*. Hoboken, NJ: Wiley.

As a general primer for R, we recommend:

- Fox, John and Weisberg, Sanford. 2011. *An R Companion to Applied Regression, Second Edition*. Thousand Oaks: Sage.
- Monogan, James E. 2015. *Political analysis using R*. Cham, Switzerland: Springer.

As a background guide for mathematical concepts discussed in this workshop, we recommend:

- Moore, Will H. and Siegel, David A. 2013. *A Mathematics Course for Political and Social Research*. Princeton, NJ: Princeton University Press.

Articles

All articles listed in the syllabus are made available to you on M-Box and through the University of Michigan library website from the campus network. Please contact the TAs if you have trouble locating any of the articles listed in the syllabus.

Software

This workshop relies mostly on R, JAGS/WinBUGS/OpenBUGS, and Stan. We may also briefly discuss Stata as an alternative for some applications. We provide assistance installing R and JAGS/WinBUGS/OpenBUGS on your computers during labs in the first week of the workshop. There is a learning curve for these programs, but you need not have any computer programming background to learn them rather easily—just patience and desire. Our goal is to make you as comfortable as possible with these programs by the end of this workshop so that you will be able to use them with ease at your home institutions and in your own work.

Mac and JAGS users: See Johannes' website for more information on installing JAGS.

Problem sets

We assign problem sets in class. Our goal is to make sure participants receive sufficient feedback to complete all assignments successfully. We distribute between 2 and 4 assignments per week. They are mostly computer-based with the exception of the first assignment. We ask that you include all code you used to complete your assignments; the TAs will show you how to best do this. The TAs will aim to return graded assignments to you within 5 days with comments via email. We (the instructors and TA) are more than happy to provide help with assignments during office hours: don't be afraid to come by and ask.

Labs

We offer several labs with guided hands-on exercises. Lab sessions will be held at the times in the room(s) listed in the schedule below (watch for updates). Planned topics:

- 1 First steps in R
- 2 Installing and using JAGS/BUGS from R
- 3 Using R and RMarkdown for an integrated and reproducible workflow for Bayesian (and frequentist) statistics
- 4 Postestimation
- 5 Multilevel data management
- 6 Using Stan

Preparing for each workshop meeting

To get the most out of this workshop, we recommend that you read the assigned background & textbook readings for each day in depth and skim at least one of the applied studies if any are listed. Particularly in the second half of the workshop, we provide a larger number of applied works—pick one that is closest to your area of interest.

You should also feel strongly encouraged to come to our office hours on any day of the workshop to follow up on topics discussed during workshop meetings and to discuss how any topic we discussed might relate to your own work.

Grading

Participants seeking a grade certificate must complete all required homework assignments (8 out of 10 assignments). For each individual assignment, we use the following grading scheme:

- ✓+ (3 points): a full write-up similar to what one would find in a scholarly article. The text should be properly formatted. Results are presented and discussed in a clear manner. Figures and tables are properly labeled. All code is included in the document. An appendix describing the data should follow after the main text and be modeled after the TIER protocol's guidelines for data appendices.

✓ (2 points): a full write-up addressing all requirements listed in the assignment. Results are presented and discussed in a clear manner. Figures and tables are properly labeled. All code is included in the document.

✓ – (1 point): insufficient assignment; has errors and omissions, did not address all parts of the assignment as required.

The final course grade is roughly determined as follows:

- Submitting all required assignments (7 points or more) and attending the course earns a **B**.
- Submitting all assignments with a ✓ (14 points) and attending the course earns an **A–**; adding active participation in discussion, labs, or office hours earns an **A**.
- Submitting at least a few assignments with a ✓ + (17 points or more) and active participation earns an **A+**.

Workshop content and schedule

The following dates and topics may be modified as the course proceeds. The most recent version of the syllabus will always be at www.jkarreth.net/bayes-icpsr.html.

Monday, June 24

No course meeting

Recommended: *Introduction to the L^AT_EX Text Processing System*, 5:30pm–7:30pm.

Day 1: Tuesday, June 25

Introduction: Background and Basics of Bayesian Inference

Please read:

- Gill: Chapter 1.
 - Siegfried, Tom. 2010. “Odds are, it’s wrong: Science Fails to Face the Shortcomings of Statistics.” *Science News* 177 (7): 26–29.
 - Senn, Stephen. 2003. “Bayesian, Likelihood, and Frequentist Approaches to Statistics.” *Applied Clinical Trials* 12 (8): 35–38.
-

Day 2: Wednesday, June 26

Review of Generalized Linear Models

Refresher:

- Gill: Section 2.2.
 - Gelman & Hill: Chapter 6.
-

Day 3: Thursday, June 27

Probability and Bayes’ Rule

Please read:

- Gill: Chapter 2.
 - Western, Bruce and Jackman, Simon. 1994. “Bayesian Inference for Comparative Research.” *American Political Science Review* 88 (2): 412–423.
-

Day 4: Friday, June 28**Priors**

Please read:

- Gill: Chapter 4.
- Gill, Jeff and Walker, Lee D. 2005. “Elicited Priors for Bayesian Model Specifications in Political Science Research.” *Journal of Politics* 67 (3): 841–872.
- Seaman, John W. III, Seaman, John W. Jr., and Stamey, James D. 2012. “Hidden Dangers of Specifying Noninformative Priors.” *The American Statistician* 66 (2): 77–84.

HW 1 assigned: Prior and posterior distributions.

Day 5: Monday, July 1**Sampling Methods and Introduction to the BUGS/JAGS Language**

Please read:

- Gill: Chapters 9 & 10.
 - Spiegelhalter, David J. et al. 2003. “WinBUGS Version 1.4 User Manual.”
 - Plummer, Martyn. 2013. “JAGS Version 3.4.0 User Manual.”
-

Day 6: Tuesday, July 2**Convergence Diagnostics**

Please read:

- Robert, Christian and Casella, George. 2010. *Introducing Monte Carlo Methods with R*. New York, NY: Springer, Chapter 8
- Plummer, Martyn et al. 2006. “CODA: Convergence Diagnosis and Output Analysis for MCMC.” *R News* 6 (1): 7–11.

Background on specific convergence diagnostics:

- Cowles, Mary Kathryn and Carlin, Bradley P. 1996. “Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review.” *Journal of the American Statistical Association* 91 (434): 883–904.
- Gelman, Andrew and Shirley, Kenneth. 2011. “Inference from Simulations and Monitoring Convergence.” Chap. 6 in *Handbook of Markov Chain Monte Carlo*, ed. by Brooks, Steve et al., 163–174. Chapman / Hall/CRC.

R implementations of different convergence diagnostics:

- Tsai, Tsung-han and Gill, Jeff. 2012. “superdiag: A Comprehensive Test Suite for Markov Chain Non-Convergence.” *The Political Methodologist* 19 (2): 12–18.
- Marín, Xavier Fernández-i. 2016. “ggmcmc: Analysis of MCMC Samples and Bayesian Inference.” *Journal of Statistical Software* 70 (1): 1–20.

HW 2 assigned: Becoming familiar with WinBUGS/JAGS.

Wednesday, July 3

The Normal Distribution; Priors (ctd.)

Please read:

- Gill: Chapter 3
 - Kerman, Jouni. 2011. “Neutral noninformative and informative conjugate beta and gamma prior distributions.” *Electronic Journal of Statistics* 5:1450–1470 (if you want to know more about noninformative priors).
 - Kass, Robert E. and Wasserman, Larry. 1996. “The Selection of Prior Distributions by Formal Rules.” *Journal of the American Statistical Association* 91 (435): 1343–1370 (if you want to know more about how to select priors).
-

Day 7: Thursday, July 4

No lecture

Instructor & TAs are available for office hours by appointment.

Day 8: Friday, July 5

The Bayesian Linear Model

Please read:

- Gill: Chapter 5.
- Efron, Brad. 1986. “Why Isn’t Everyone a Bayesian?” *American Statistician* 40 (1): 1–5.

Sample application:

- Blais, André, Guntermann, Eric, and Bodet, Marc A. 2017. “Linking Party Preferences and the Composition of Government: A New Standard for Evaluating the Performance of Electoral Democracy.” *Political Science Research and Methods* 5 (2): 315–331.

On missing data:

- Jackman, Simon. 2000. “Estimation and Inference Are Missing Data Problems: Unifying Social Science Statistics via Bayesian Simulation.” *Political Analysis* 8 (4): 307–332.
- Zhou, Xiang and Reiter, Jerome P. 2010. “A Note on Bayesian Inference After Multiple Imputation.” *The American Statistician* 64 (2): 159–163.

HW 3 assigned: Linear model.

HW 4 assigned: Debugging BUGS/JAGS code.

Day 9: Monday, July 8

Binary Outcomes

If you'd like a refresher for generalized linear models and their interpretation, please read:

- Gelman & Hill, Chapter 5.
- King, Gary, Tomz, Michael, and Wittenberg, Jason. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44 (2): 347–361.
- Hanmer, Michael J. and Kalkan, Kerem Ozan. 2013. "Behind the Curve: Clarifying the Best Approach to Calculating Predicted Probabilities and Marginal Effects from Limited Dependent Variable Models." *American Journal of Political Science* 57 (1): 263–277.
- Gelman, Andrew et al. 2008. "A weakly informative default prior distribution for logistic and other regression models." *Annals of Applied Statistics* 2 (4): 1360–1383.

Some background on Bayesian estimation for binary outcomes:

- Albert, James H. and Chib, Siddhartha. 1993. "Bayesian Analysis of Binary and Polychotomous Response Data." *Journal of the American Statistical Association* 88 (422): 669–679.
- Hollenbach, Florian M., Montgomery, Jacob M., and Crespo-Tenorio, Adriana. 2019. "Bayesian Versus Maximum Likelihood Estimation of Treatment Effects in Bivariate Probit Instrumental Variable Models." *Political Science Research and Methods* 7 (3): 651–659.

Sample application:

- Karreth, Johannes. 2018. "The Economic Leverage of International Organizations in Interstate Disputes." *International Interactions* 44 (3): 463–490.

HW 5 assigned: Logistic regression model (due Wednesday, July 10).

Day 10: Tuesday, July 9

Ordered and Categorical Outcomes

If you'd like a refresher on today's models, please read:

- Gelman & Hill, section 6.5.

Sample applications (please read one):

- Duch, Raymond M., May, Jeff, and Armstrong, David A. 2010. "Coalition-directed Voting in Multiparty Democracies." *American Political Science Review* 104 (4): 698–719.
 - Stegmueller, Daniel. 2013b. "Modeling Dynamic Preferences: A Bayesian Robust Dynamic Latent Ordered Probit Model." *Political Analysis* 21 (3): 314–333.
 - Stegmueller, Daniel et al. 2012. "Support for Redistribution in Western Europe: Assessing the Role of Religion." *European Sociological Review* 28 (4): 482–497.
 - Alvarez, R. Michael and Nagler, Jonathan. 1998. "When Politics and Models Collide: Estimating Models of Multiparty Elections." *American Journal of Political Science* 42 (1): 55–96.
 - Lacy, Dean and Burden, Barry C. 1999. "The Vote-Stealing and Turnout Effects of Ross Perot in the 1992 U.S. Presidential Election." *American Journal of Political Science* 43 (1): 233–255.
 - Imai, Kosuke and Dyk, David A. van. 2005. "A Bayesian analysis of the multinomial probit model using marginal data augmentation." *Journal of Econometrics* 124 (2): 311–334.
-

Day 11: Wednesday, July 10

Count Outcomes

If you'd like a refresher on today's models, please read one of the following:

- Gelman & Hill, section 6.2.
- Ntzoufras, sections 7.4 and 8.3

Sample applications (please read one):

- Martin, Andrew D. 2003. "Bayesian Inference for Heterogeneous Event Counts." *Sociological Methods & Research* 32 (1): 30–63.
- Ghosh, Sujit K., Mukhopadhyay, Pabak, and Lu, Jye-Chyi. 2006. "Bayesian analysis of zero-inflated regression models." *Journal of Statistical Planning and Inference* 136 (4): 1360–1375.
- Neelon, Brian H, O'Malley, A James, and Normand, Sharon-Lise T. 2010. "A Bayesian model for repeated measures zero-inflated count data with application to outpatient psychiatric service use." *Statistical Modelling* 10 (4): 421–439.

HW6 (Ordered or multinomial logit model) and HW 7 (Poisson model) assigned (due Sunday, July 14).
 Complete one of HW6, HW7, or HW8 for full credit.

Day 12: Thursday, July 11

Measurement Models

For background, please read:

- On factor models:
 - Lopes, Hedibert Freitas. 2014. "Modern Bayesian Factor Analysis." In *Bayesian Inference in the Social Sciences*, ed. by Jeliaskov, Ivan and Yang, Xin-She, 115–153. John Wiley & Sons, Inc., sections 5.1 and 5.2.
- On IRT models:
 - Jackman, Simon. 2001. "Multidimensional Analysis of Roll Call Data via Bayesian Simulation: Identification, Estimation, Inference, and Model Checking." *Political Analysis* 9 (3): 227.
 - Clinton, Joshua D. and Jackman, Simon. 2009. "To Simulate or NOMINATE?" *Legislative Studies Quarterly* 34 (4): 593–621.

Sample applications (please read one):

- Bakker, Ryan. 2009. "Re-measuring Left–Right: A Comparison of SEM and Bayesian Measurement Models for Extracting Left–Right Party Placements." *Electoral Studies* 28 (3): 413–421.
- Bakker, Ryan and Poole, Keith T. 2013. "Bayesian Metric Multidimensional Scaling." *Political Analysis* 21 (1): 125–140.
- Benson, Brett V. and Clinton, Joshua D. 2016. "Assessing the Variation of Formal Military Alliances." *Journal of Conflict Resolution* 60 (5): 866–898.
- Campbell, Susanna P., Findley, Michael G., and Kikuta, Kyosuke. Forthcoming. "An Ontology of Peace: Landscapes of Conflict and Cooperation with Application to Colombia." *International Studies Review*.
- Caughey, Devin and Warsaw, Christopher. 2015. "Dynamic Estimation of Latent Opinion Using a Hierarchical Group-Level IRT Model." *Political Analysis* 23 (2): 197–211.

- Caughey, Devin and Warshaw, Christopher. 2016. "The Dynamics of State Policy Liberalism, 1936-2014." *American Journal of Political Science* 60 (4): 899–913.
- Clinton, Joshua D. and Jackman, Simon. 2009. "To Simulate or NOMINATE?" *Legislative Studies Quarterly* 34 (4): 593–621.
- Copelovitch, Mark S., Gandrud, Christopher, and Hallerberg, Mark. "Financial Regulatory Transparency and Sovereign Borrowing Costs."
- Fariss, Christopher J. 2014. "Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability." *American Political Science Review* 108 (2): 297–318.
- Fox, Jean-Paul and Glas, Cees. 2001. "Bayesian Estimation of a Multilevel IRT Model Using Gibbs Sampling." *Psychometrika* 66 (2): 271–288.
- Fox, Jean-Paul and Glas, Cees A.W. 2003. "Bayesian modeling of measurement error in predictor variables using item response theory." *Psychometrika* 68 (2): 169–191.
- Garrett, Elizabeth S. and Zeger, Scott L. 2000. "Latent Class Model Diagnosis." *Biometrics* 56 (4): 1055–1067.
- Gray, Julia and Slapin, Jonathan B. 2012. "How Effective are Preferential Trade Agreements? Ask the Experts." *Review of International Organizations* 7 (3): 309–333.
- Hare, Christopher et al. 2015. "Using Bayesian Aldrich-McKelvey Scaling to Study Citizens' Ideological Preferences and Perceptions." *American Journal of Political Science* 59 (3): 759–774.
- Hollyer, James R., Rosendorff, B. Peter, and Vreeland, James Raymond. 2014. "Measuring Transparency." *Political Analysis* 22 (4): 413–434.
- Linzer, Drew A. and Staton, Jeffrey K. 2015. "A Global Measure of Judicial Independence, 1948-2012." *Journal of Law and Courts* 3 (2): 223–256.
- Manatschal, Anita and Bernauer, Julian. 2016. "Consenting to Exclude? Empirical Patterns of Democracy and Immigrant Integration Policy." *West European Politics* 39 (2): 183–204.
- Rosas, Guillermo, Shomer, Yael, and Haptonstahl, Stephen R. 2015. "No News Is News: Nonignorable Nonresponse in Roll-Call Data Analysis." *American Journal of Political Science* 59 (2): 511–528.
- Selin, Jennifer L. 2015. "What Makes an Agency Independent?" *American Journal of Political Science* 59 (4): 971–987.
- Slapin, Jonathan B. and Proksch, Sven-Oliver. 2008. "A Scaling Model for Estimating Time-Series Party Positions from Texts." *American Journal of Political Science* 52 (3): 705–722.
- Treier, Shawn and Jackman, Simon. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 52 (1): 201–217.
- Juhl, Sebastian. Forthcoming. "Measurement Uncertainty in Spatial Models: A Bayesian Dynamic Measurement Model." *Political Analysis*.
- Caughey, Devin, O'Grady, Tom, and Warshaw, Christopher. Forthcoming. "Policy Ideology in European Mass Publics, 1981–2016." *American Political Science Review*.
- Juhl, Sebastian. 2019. "Measurement Uncertainty in Spatial Models: A Bayesian Dynamic Measurement Model." *Political Analysis* 27 (3): 302–319.
- Claassen, Christopher. 2019. "Estimating Smooth Country–Year Panels of Public Opinion." *Political Analysis* 27 (1): 1–20.
- Williams, Rob et al. Forthcoming. "A latent variable approach to measuring and explaining peace agreement strength." *Political Science Research and Methods*.

HW 8 assigned: Factor or IRT model (due Sunday, July 14). Complete one of HW6, HW7, or HW8 for full credit.

Day 13: Friday, July 12**Bayes Factors and Bayesian Model Averaging**

Please read:

- Montgomery, Jacob M. and Nyhan, Brendan. 2010. “Bayesian Model Averaging: Theoretical Developments and Practical Applications.” *Political Analysis* 18 (2): 245–270.
 - Warren, T. Camber. 2014. “Not by the Sword Alone: Soft Power, Mass Media, and the Production of State Sovereignty.” *International Organization* 68 (1): 111–141 (skim as an example of an application of BMA).
 - Pepinsky, Thomas B. 2014. “The Politics of Capital Flight in the Global Economic Crisis.” *Economics & Politics* 26 (3): 431–456 (skim as an example of an application of BMA).
 - Raftery, Adrian E. 1995. “Bayesian Model Selection in Social Research.” *Sociological Methodology* 25:111–163 (Background on BMA, read if you’re interested)
 - Gelman, Andrew and Rubin, Donald B. 1995. “Avoiding Model Selection in Bayesian Social Research.” *Sociological Methodology* 25:165–173 (Background on BMA, read if you’re interested)
 - Bartels, Larry M. 1997. “Specification Uncertainty and Model Averaging.” *American Journal of Political Science* 41 (2): 641–674 (Background on BMA, read if you’re interested)
 - Montgomery, Jacob M., Hollenbach, Florian M., and Ward, Michael D. 2012. “Improving Predictions Using Ensemble Bayesian Model Averaging.” *Political Analysis* 20 (3): 271–291 (if you are interested prediction & forecasting).
 - Cranmer, Skyler J., Rice, Douglas R., and Siverson, Randolph M. 2017. “What To Do About Atheoretic Lags.” *Political Science Research and Methods* 5 (4): 641–665 (BMA as an approach to atheoretic lags in regression).
-

Day 14: rescheduled to Thursday, July 18 (see below)**Day 15: Monday, July 15****Multilevel Models (Fundamentals)**

Please read:

- Gelman & Hill: Chapter 16 or/and Gill: Chapter 10
- Gelman & Hill: Chapter 11 (for a refresher on multilevel models).
- Steenbergen, Marco R. and Jones, Bradford S. 2002. “Modeling Multilevel Data Structures.” *American Journal of Political Science* 46 (1): 218–237 (for a refresher on multilevel models).
- Shor, Boris et al. 2007. “A Bayesian Multilevel Modeling Approach to Time-Series Cross-Sectional Data.” *Political Analysis* 15 (2): 165–181 (if you work with TSCS data).
- Bell, Andrew and Jones, Kelvyn. 2015. “Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data.” *Political Science Research and Methods* 3 (1): 133–153 (if you work with TSCS data).
- Plümper, Thomas and Troeger, Vera E. 2019. “Not so Harmless After All: The Fixed-Effects Model.” *Political Analysis* 27 (1): 21–45 (if you work with TSCS data).
- Greenland, Sander. 2007. “Bayesian perspectives for epidemiological research. II. Regression analysis.” *International Journal of Epidemiology* 36 (1): 195–202 (if you work with more complex nesting structures).
- Gelman, Andrew, Hill, Jennifer, and Yajima, Masanao. 2012. “Why We (Usually) Don’t Have to Worry About Multiple Comparisons.” *Journal of Research on Educational Effectiveness* 5 (2): 189–211.

- Gelman, Andrew and Pardoe, Iain. 2006. "Bayesian Measures of Explained Variance and Pooling in Multilevel (Hierarchical) Models." *Technometrics* 48 (2): 241–251.
-

Tuesday, July 16

Lab (Location: MLB Lecture 2)

Constructing and processing multilevel data

Day 16: Wednesday, July 17

Multilevel Models (non-continuous outcomes; time-series cross-sectional data as multilevel data)

Multilevel regression with poststratification (MRP)

Please continue to read:

- Gelman & Hill: Chapter 17 (Chapter 15 for a refresher).

as well as any of these empirical articles using MLMs that is/are in your area of interest:

- Pang, Xun. 2010. "Modeling Heterogeneity and Serial Correlation in Binary Time-Series Cross-sectional Data: A Bayesian Multilevel Model with AR(p) Errors." *Political Analysis* 18:470–498.
- Pang, Xun. 2014. "Varying Responses to Common Shocks and Complex Cross-Sectional Dependence: Dynamic Multilevel Modeling with Multifactor Error Structures for Time-Series Cross-Sectional Data." *Political Analysis* 22 (4): 464–496.
- Ward, Michael D., Siverson, Randolph M., and Cao, Xun. 2007. "Disputes, Democracies, and Dependencies: A Reexamination of the Kantian Peace." *American Journal of Political Science* 51 (3): 583–601.
- Blaydes, Lisa and Linzer, Drew A. 2012. "Elite Competition, Religiosity and Anti-Americanism in the Islamic World." *American Political Science Review* 106 (2): 225–243.
- Lock, Kari and Gelman, Andrew. 2010. "Bayesian Combination of State Polls and Election Forecasts." *Political Analysis* 18 (3): 337–348.
- Masters, Ryan K., Hummer, Robert A., and Powers, Daniel A. 2012. "Educational Differences in U.S. Adult Mortality." *American Sociological Review* 77 (4): 548–572.
- Stegmueller, Daniel. 2013a. "How Many Countries for Multilevel Modeling? A Comparison of Frequentist and Bayesian Approaches." *American Journal of Political Science* 57 (3): 748–761.
- Chaudoin, Stephen, Milner, Helen V., and Pang, Xun. 2015. "International Systems and Domestic Politics: Linking Complex Theories with Empirical Models in International Relations." *International Organization* 69 (2): 275–309.
- Beazer, Quintin H. and Woo, Byungwon. 2016. "IMF Conditionality, Government Partisanship, and the Progress of Economic Reforms." *American Journal of Political Science* 60 (2): 304–321.
- Danneman, Nathan and Ritter, Emily Hencken. 2014. "Contagious Rebellion and Preemptive Repression." *Journal of Conflict Resolution* 58 (2): 254–279.
- Quaranta, Mario and Martini, Sergio. 2016. "Does the economy really matter for satisfaction with democracy? Longitudinal and cross-country evidence from the European Union." *Electoral Studies* 42:164–174.
- Eagle, David. 2016. "The Negative Relationship between Size and the Probability of Weekly Attendance in Churches in the United States." *Socius* 2.
- Cao, Xun and Ward, Hugh. 2017. "Transnational Climate Governance Networks and Domestic Regulatory Action." *International Interactions* 43 (1): 76–102.

- Helgason, Agnar Freyr and Mérola, Vittorio. 2017. “Employment Insecurity, Incumbent Partisanship, and Voting Behavior in Comparative Perspective.” *Comparative Political Studies* 50 (11): 1489–1523.
- Mummolo, Jonathan and Peterson, Erik. 2018. “Improving the Interpretation of Fixed Effects Regression Results.” *Political Science Research and Methods*.

Overview and applications of multilevel regression with poststratification (MRP):

- Park, David K., Gelman, Andrew, and Bafumi, Joseph. 2004. “Bayesian Multilevel Estimation with Post-stratification: State-Level Estimates from National Polls.” *Political Analysis* 12 (4): 375–385.
- Lax, Jeffrey R. and Phillips, Justin H. 2009. “How Should We Estimate Public Opinion in The States?” *American Journal of Political Science* 53 (1): 107–121.
- Kastellec, J. P., Lax, J. R., and Phillips, J. H. (2016). Estimating State Public Opinion With Multi-Level Regression and Poststratification using R. *Working paper*. Available at http://www.princeton.edu/~jkastell/MRP_primer/mrp_primer.pdf.
- Selb, Peter and Munzert, Simon. 2011. “Estimating Constituency Preferences from Sparse Survey Data Using Auxiliary Geographic Information.” *Political Analysis* 19 (4): 455–470.
- Warshaw, Christopher and Rodden, Jonathan. 2012. “How Should We Measure District-Level Public Opinion on Individual Issues?” *Journal of Politics* 74 (1): 203–219.
- Buttice, Matthew K. and Highton, Benjamin. 2013. “How Does Multilevel Regression and Poststratification Perform with Conventional National Surveys?” *Political Analysis* 21 (4): 449–467.
- Toshkov, Dimiter. 2015. “Exploring the Performance of Multilevel Modeling and Poststratification with Eurobarometer Data.” *Political Analysis* 23 (3): 455–460.
- Flores, Andrew R, Herman, Jody L, and Mallory, Christy. 2015. “Transgender inclusion in state non-discrimination policies: The democratic deficit and political powerlessness.” *Research & Politics* 2 (4).
- Leemann, Lucas and Wasserfallen, Fabio. Forthcoming. “Extending the Use and Prediction Precision of Subnational Public Opinion Estimation.” *American Journal of Political Science*.
- Broockman, David E. and Skovron, Christopher. Forthcoming. “Bias in Perceptions of Public Opinion among Political Elites.” *American Political Science Review*.
- Kiewiet De Jonge, Chad P, Langer, Gary, and Sinozich, Sofi. Forthcoming. “Predicting State Presidential Election Results Using National Tracking Polls and Multilevel Regression With Poststratification (MRP).” *Public Opinion Quarterly*.

HW 10 assigned: Multilevel model (due Friday, July 19).

Extra lecture: Thursday, July 18, 8:00am–9:00am (Location: MLB Lecture 2) Model Checking and Model Presentation

Please read:

- Gill: Chapters 6 & 7.
- Gelman, Andrew et al. 2000. “Diagnostic Checks for Discrete Data Regression Models Using Posterior Predictive Simulations.” *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 49 (2): 247–268.
- Vehtari, Aki, Gelman, Andrew, and Gabry, Jonah. 2016. “Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC.” *Statistics and Computing*: 1–20.
- Quinn, Kevin M., Martin, Andrew D., and Whitford, Andrew B. 1999. “Voter Choice in Multi-Party Democracies: A Test of Competing Theories and Models.” *American Journal of Political Science* 43 (4): 1231–1247 (if you are interested in model comparison).

- Gelman, Andrew et al. 2019. “R-squared for Bayesian Regression Models.” *The American Statistician* 0 (0): 1–7.

HW 9 assigned: Model checking for linear regression (optional; due Friday, July 19).

Day 17: Thursday, July 18
Multilevel models and MRP continued
Bayesian Analysis of Spatial Data

If interested, please read the following for background and applications of spatial modeling using Bayesian inference:

- Lunn et al.: Section 11.3.
 - Sparks, Corey S. 2011. “Violent crime in San Antonio, Texas: An application of spatial epidemiological methods.” *Spatial and Spatio-temporal Epidemiology* 2 (4): 301–309.
-

Day 18: Friday, July 19: Using Bayesian Modeling in Your Applied Work

Please read:

- Gelman, Andrew. 2008. “Objections to Bayesian Statistics.” *Bayesian Analysis* 3 (3): 445–450.
- Gabry, Jonah et al. 2019. “Visualization in Bayesian workflow.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 182 (2): 389–402.
- Humphreys, Macartan and Jacobs, Alan M. 2015. “Mixing Methods: A Bayesian Approach.” *American Political Science Review* 109, no. 4 (04): 653–673.

Also have a look at:

- Stan example models “*ARM Models Sorted by Chapter*” at <https://github.com/stan-dev/example-models/wiki/ARM-Models-Sorted-by-Chapter>.