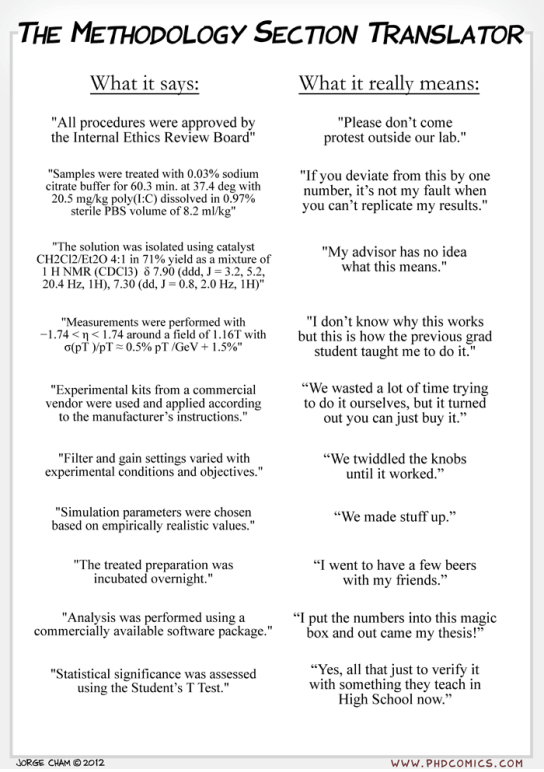
Quantitative Methods Across the Social Sciences
Communication Studies 840-001
Josh Pasek, Ph.D.
Fall, 2016
jpasek@umich.edu

Course Meetings:	Office Hours:
W – 4:00-7:00 PM	M – 1:30-2:30 PM
270 Tappan	5413 North Quad

What's the question?

Methodological innovation has become perhaps the fastest growing area in contemporary social science. Across all of the social sciences, new techniques are being developed that allow researchers to more effectively and efficiently answer their questions. But the expanding scope of methodological techniques has also posed distinct challenges. Researchers across different disciplines are increasingly unfamiliar with the strategies used in neighboring fields. Similarly, the plethora of novel statistical approaches has made it increasingly difficult to determine which approach is ideal for a particular question. And the diversity of methodological choices has made it increasingly difficult to evaluate research and to discriminate findings that represent methodological confounds from those providing ironclad evidence.

To the extent that researchers wish to understand, use, and communicate results from various techniques, the ever-changing methodological scope has presented a conundrum: The methodological toolbox of the social sciences has become far too large for common mastery. This course attempts to circumvent the problem. Instead of trying to train researchers to be proficient at each of dozens of



is to develop a broad roadmap. In a whirlwind we attempt to demarcate the generalized ntational conventions underlying the statistical

Researchers should emerge from this class with a better sense of what a variety of quantitative methods actually do, how they are used in practice, the assumptions that are being made, and what kinds of questions they can answer.

Course Content

Weekly classes will each focus on a general area of statistical methodology that is being used in the social sciences. The course will begin with the assumption that students have been exposed to statistical and methodological courses that have explained the principals behind basic probability and ordinary least squares regression. We will start from that point and try to unpack as many methods as are feasible in the term.

The course will principally use a shared MBox folder for all content. I recommend syncing this folder with your computer.

Course Requirements

In any given week, students will be responsible for (1) reading the articles in the "required" folder for that

week, (2) selecting an additional article that uses that method to evaluate (either from the "optional folder" or by searching the literature for their favorite journal), (3) writing a one-page assessment of how the week's methods are being used in that article, and (4) presenting the article to the class. Once each student has selected an article to summarize and present for the following week, he or she should **move that article from the "optional" folder to the "selected" folder** (so nobody else chooses it).

A number of academic articles are available for each week's optional readings. These can be found in the course MBox folder for that week under "optional". The articles themselves are also summarized in the "Reference List" inside of that folder. Please note that there are a large number of articles for many weeks, which were identified through keyword searches. This means that some articles may not use the methods identified (perhaps mentioning them as alternative approaches) and that others may include the methods identified as well as other, more complicated approaches. When choosing an article, I highly recommend skimming the article first to make sure that it does use the method mentioned and, preferably, does not include other methods that you have not heard of. This should improve your ability to interpret the article.

For both the one-page assessments and the presentations, students should focus on five aspects of each article: (1) the types of data that are being used, (2) how the method is applied, (3) the conclusions that are being drawn from the method, (4) how the method and results are presented, and (5) any assumptions

that are being made about the nature of the method or the data to which it is applied.

As a final project for the class, students will be expected to use any one of the methods discussed in class to analyze data from their own research. These final projects will be graded on how well they apply the method as well as their description of the benefits and limitations of the method for use with their data. These projects should take the form of the methods section from a typical research article.

Examples for all methods in the class will be provided using R.

Grading

Class Participation / Presentations	40%
One-page write-ups	30%
Final Project	30%

NOTE TOPICS FOR THE FIRST FEW WEEKS ARE LISTED BELOW

IN WEEK 1, WE WILL DISCUSS WHAT TOPICS STUDENTS ARE MOST INTERESTED IN INCLUDING AND INCORPORATE THESE

Course Schedule:

Sept 7 – Week 1 – Introduction

Today we will discuss the organization of the class, requirements for the class, and will outline the semester. We will also go over strategies for quickly and effectively reading empirical research articles. This class will also include a brief introduction to the statistical software package R. Please download R and try some of the examples from the Muenchen chapter before class.

Required Articles:

Muenchen, R. A. (2007). R for SAS and SPSS Users. Springer: New York. Available from: http://www.et.bs.ehu.es/~etptupaf/pub/R/RforSAS&SPSSusers.pdf

R-Core (2012). The R Project for Statistical Computing. www.rproject.org

Sept 14 – Week 2 – Limited Dependent Variable Regressions

This week we will discuss strategies for running regressions when the outcome variables are not continuous measures. These will include examples of binomial, ordinal, multinomial, and seemingly unrelated regressions. Binomial regressions are used for dichotomous outcome variables, ordinal regressions are used for categorical variables with ordered categories, multinomial regressions are used for multiple-category outcomes that are not ordered, and seemingly unrelated regressions are used when estimating multiple continuous outcomes that are interrelated.

Search terms for articles: "logistic regression", "logit", "probit", "ordinal logit", "multinomial logit", "multinomial probit", "seemingly unrelated regression"

Required Articles:

Alvarez, R. M. & Nagler, J. (1998). When Politics and Models Collide: Estimating Models of Multiparty Elections. *American Journal of Political Science* 42: 55-96.

Bender, R. & Grouven, U. (1997). Ordinal Logistic Regresion in Medical Research. *Journal of the Royal College of Physicians of London, 31*(5), 546-551.

Peng, C.-Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An Introduction to Logistic Regression Analysis and Reporting. *Journal of Educational Research, 96*(1), 3-14.

Sept 21 – Week 3 – Factor Analysis, Principal Components, and Item Response Theory

This week we will examine tools for identifying shared variance across multiple measures. These tools allow us to take multiple measures of some underlying construct and generate a single measure for it. We will discuss how different approaches to combining measures differently incorporate measurement errors and shared variance across items.

Search terms for articles: "factor analysis", "principal components", "Item Response Theory", "IRT"

Required Articles:

Fabrigar, J. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *Psychological Methods, 4*(3), 272-299.

Park, H. S., Dailey, R., & Lemus, D. (2002). The Use of Exploratory Factor Analysis and Principal Components Analysis in Communication Research. *Human Communication Research*, *28*(4), 562-577.

Reise, S. P., Ainsworth, A. T., & Haviland, M. G. (2005). Item Response Theory Fundamentals, Applications, and Promise in Psychological Research. *Current Directions in Psychological Science*, *14*(2), 95–101.

Sept 28 – Week 4 – Structural Equation Modeling and Instrumental Variables

Structural equation modeling and instrumental variable analysis are tools that leverage shared variance across measures to estimate chains of causal relations.

Search terms for articles: "structural equation modeling", "SEM", "Instrumental variable"

Required Articles:

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak. *Journal of the American Statistical Association, 90*(430), 443-450.

Hox, J. J., Bechger, T. M. (1998). An Introduction to Structural Equation Modeling. *Family Science Review, 11*, 354-373.

MacCallum, R. C., & Austin, J. T. (2000). Applications of Structural Equation Modeling in Psychological Research. *Annual Review of Psychology, 51*(1), 201-226.

Oct 5 – Week 5 – Permutation Tests and Bootstrapping

Permutation tests and bootstrapping are examples of resampling statistics, where researchers presume that the distributional forms of their data are related to what they observed in their data collection. These methods are ideal for small-N data, non-normal distributions, and correcting for difficult-to-predict systematicity in data.

Search terms for articles: "bootstrap", "permutation test", "randomization test", "Monte Carlo simulation", "Jackknife". [NOTE: If you search for Monte Carlo Simulation, make sure that your article does NOT include "Markov Chain" in it, as this is part of the empirical Bayes approach]

Required Articles:

Henderson, A. R. (2005). The bootstrap: A technique for data-driven statistics. Using computer-intensive analyses to explore experimental data. *Clinica Chimica Acta, 359*(1), 1-26.

Preacher, K. J., & Selig, J. P. (2012). Advantages of Monte Carlo Confidence Intervals for Indirect Effects. Communication Methods and Measures, 6(2), 77–98. http://doi.org/10.1080/19312458.2012.679848

Oct 12 – Week 6 – [CLASS WILL BE RESCHEDULED DUE TO YOM KIPPUR] Fixed and Random Effects *and* **Hierarchical and Multilevel Models**

Oct 19 – Week 7 – Dealing with Missing Data
Oct 26 – Week 8 – Empirical Bayes
Nov 2 – Week 9 – Network Analysis
Nov 9 – Week 10 – Working with GIS and Spatial Data
Nov 16 – Week 11 – Survey Weighting and Matching Techniques
Nov 23 – Week 12 – NO CLASS – Thanksgiving
Nov 30 – Week 13 – Time Series, Regression Discontinuity, and Growth Curves
Dec 7 – Week 14 – Word Counts, Sentiment, and Topic Modeling (Textual Data)
Dec 14 – Week 15 – Presentations on Applications