Political Methodology Comprehensive Examination, September 2013
Department of Political Science, George Washington University

Instructions: You have six hours to write the exam. Answer all questions in part I, one question in part II, and after the exam turn in an empirical paper demonstrating your ability to use statistical models OR schedule an oral exam.

Part I: Answer all 5 questions; read all questions before answering any of them.

1. A popular empirical approach in the social sciences to exploit a "natural experiment." What does it mean for something to be a natural experiment? Give a brief example. How does one verify that a source of variation counts as a natural experiment? What are the strengths and weaknesses of this approach?

2. After presenting a paper showing a significant result using OLS, your discussant objects that both your dependent variable and main independent variable are likely measured with error, and therefore we should be dubious about your significant finding. Is this conclusion valid? What are some of the problems that result from each variable being measured with error?

3. OLS and other parametric methods identify relationships among several data points, but single observations can have a large influence. Under what conditions are outliers likely to be a problem for your parameter estimates? What are some of the techniques one can use to identify whether specific observations have an outsized influence on your estimated coefficients?

4. Political scientists often confront the following situation. The outcome of interest to be explained by a statistical model is ordinal, with a small number of values (3, 4, or 5, say). A standard model to estimate in such situations is either the ordered probit or ordered logit model. Why do analysts argue that such models are preferable to OLS? Are there any situations where OLS would be preferable? What is the proper way to interpret the coefficient estimates in an ordered probit or ordered logit model?

5. At the annual meeting of the American Political Science Association a few weeks ago, a prominent political scientist argued in favor of what he called "results-blind reviewing." If implemented, reviewers would review manuscripts on the basis of the research design (tested hypotheses, data collection strategy, etc.) but the reviewers would not see the results of the data analysis. This idea would, in principle, reduce the amount of atheoretical data mining that is possible with observational data or with experimental data and subgroup analysis and make reviewers and authors less fixated on whether results are "significant at the .05 level" . Do you think results-blind reviewing is a good idea? What problems might it solve? What problems might it create?

Part II: Answer ONE of these two questions

6a. Missing data are often a problem in political science data. What are some solutions for dealing with missing data, and what are their strengths and weaknesses? Do you think the effectiveness of missing data solutions vary by substantive area? If not, why not? If so, in what areas will the solutions be more or less effective?

6b. Analyzing clustered data (hierarchical data, panel and TSCS data) presents some key modeling obstacles. If you had to choose the two most important issues/considerations associated with clustered data to communicate to practitioners, what would they be? Why are they so important?