Political Methodology Comprehensive Examination, September 2015
Department of Political Science, George Washington University

Instructions: *Read all questions before answering any of them. Answer 6 questions in total. You can answer question 8 or question 9, but do NOT answer both question 8 and 9. After the exam turn in an empirical paper demonstrating your ability to use statistical models OR schedule an oral exam. Good luck!*

1. Suppose you are interested in testing how X affects Y, but there are a number of potential controls you can add to your regression. You don't have enough data to control for all of them at once, forcing you to choose some subset of controls. How might this choice affect statistical inference? How should this choice be made? Should it be based on theory or the data? What are some alternatives to making a single choice and instead using information from multiple models?

2. Suppose you are presenting a paper at a conference that regresses Y on independent variables X, Z, and a set of controls. You show that X not only affects Y, but has a larger effect than Z, which has garnered more attention in the literature. A Z-scholar in the audience interrupts and says, "Wait a second! Z has a higher significance level than X! Doesn't that contradict your argument?" How would you respond? What are some reasons that a variable might be "more significant" without having a larger substantive effect? Are there any lessons empirical scholars can draw from this?

3. One of the most common criticisms of causal inference and experimental research is "external validity." Explain what this means and contrast it with "internal validity." Why is external validity a special concern for causal and experimental work? What are some of the ways researchers can address external validity concerns?

4. You go to a job talk where the candidate is making the argument that individuals' issue positions and gender attitudes have differently structured their preferences for the Republican party over time because the time periods have offered contexts with greater or fewer women candidates on the ballot, which conditions the effects of attitudes and positions on partisan preferences. Regardless of context, however, the candidate argues, Republicans are more preferred if they are the incumbent party.

(a) Use the following variables to write out a single regression equation that might capture the argument being made in the talk, adding any subscripts that might help clarify the equation:

Republican Preference

Gender Attitudes

Issue Positions

%Female Candidates

Incumbent Party

(b) Rather than run a single regression, the job candidate says that he estimated a multilevel model. The system of equations that specify his model are:

$$Republican\ Preference_{it} = \beta_{0t} + \beta_{1t}Issue\ positions + \beta_{2t}Gender\ attitudes + \epsilon_{it} \qquad (1)$$

$$\beta_{0t} = \gamma_{00} + \gamma_{01}Incumbent\ party + u_{0t} \qquad (2)$$

$$\beta_{1t} = \gamma_{10} + \gamma_{11}\%candidates\ female + u_{1t} \qquad (3)$$

$$\beta_{2t} = \gamma_{20} + \gamma_{21}\%candidates\ female + u_{2t} \tag{4}$$

Substitute the values assigned to β_{0t}, β_{1t}, and β_{2t} by equations 2 through 4 into equation 1, multiply out, and collect the stochastic terms together. What problems would you have if you ran OLS to estimate this equation? (i.e., explain what assumptions would be violated and what the consequences for the OLS estimator would be.)

(c) Comment on the similarities and differences between the equation you wrote in part b and the model you wrote out in part a. What does this tell you about why the job candidate made the modeling choice(s) he did?

5. After showing a scatterplot exhibiting a positive linear correlation between a pretest X and a posttest Y ($r \approx .65$), a presenter comments that the lowest decile observations on the X variable increased more on average than the highest decile observations on the X variable. He proceeds to give an elaborate explanation as to why this is substantively important. Can you propose a simpler statistical explanation for his "finding"? Explain.

6. In the analysis of time-series data, what are the sources and consequences of both a stationary dynamic process and an integrated (or "random walk") time series. In other words, what does it mean – both substantively and statistically – that we have a stationary process and what does it mean that we have an integrated process? And why does it matter? Discuss the ultimate implications of each process for drawing causal inferences about the effect of X on Y. Use a substantive example from politics to illustrate your arguments.

7. One of the points of modeling is to draw broad generalizations about political phenomena. Does the specification of interaction terms (and conditional effects more generally) in an empirical model detract from the goal of generalization? Regardless of your answer to this first question, make a case for how specifying interaction terms can still lead to meaningful and powerful generalizations. In your answer, clear up any misconceptions about the specification, interpretation, and presentation of interaction terms that might help convince a skeptic.

Questions 8 and 9 have the common introduction:

Hot in the (social science) news recently was the report of the large scale Reproducibility Project in psychology. The project took published work (called "originals") then replicated them (called "replications") and compared their effects.

8. Figure 1 below is figure 1 from the replication project report. Panel A compares the p-values for the original studies vs. the replications and Panel B compares the effect sizes. How could the patterns in Panels A and B occur if all actors (authors, reviewers, and editors) are behaving ethically and following the strictures of the scientific method? How could the patterns in Panels A and B occur if some actors (authors, reviewers, and editors) are not behaving ethically and not following the strictures of the scientific method?

9. Figure 2 below is figure 3 from the replication project report. On the x-axis are the effect sizes of the original studies; on the y-axis are the effect sizes of the replications. Red indicates non-significant effects (i.e., p NOT $< .05$) and blue indicates significant effects (i.e., $p < .05$). The conclusion of the project was that 36 out of 100 efforts "failed to replicate" and much lamentation trumpeted across the land. Just given the results in Figure 3, tell us two stories. First, tell us a story about why these results should make us pessimistic about the state of social science. Second, tell us a story about why these results don't seem as bad as portrayed in the popular press and should make us at least mildly optimistic about the state of social science.
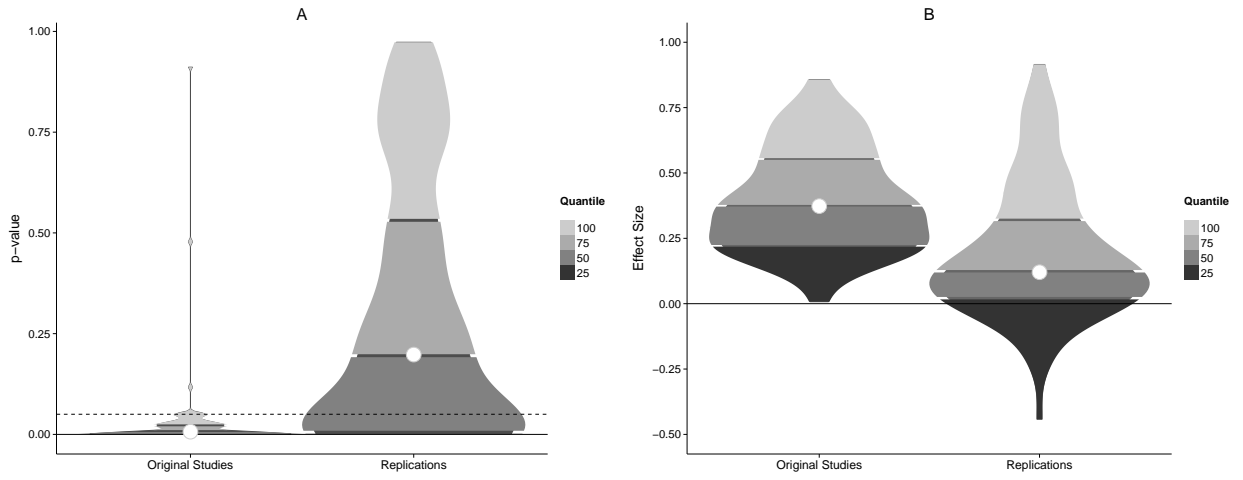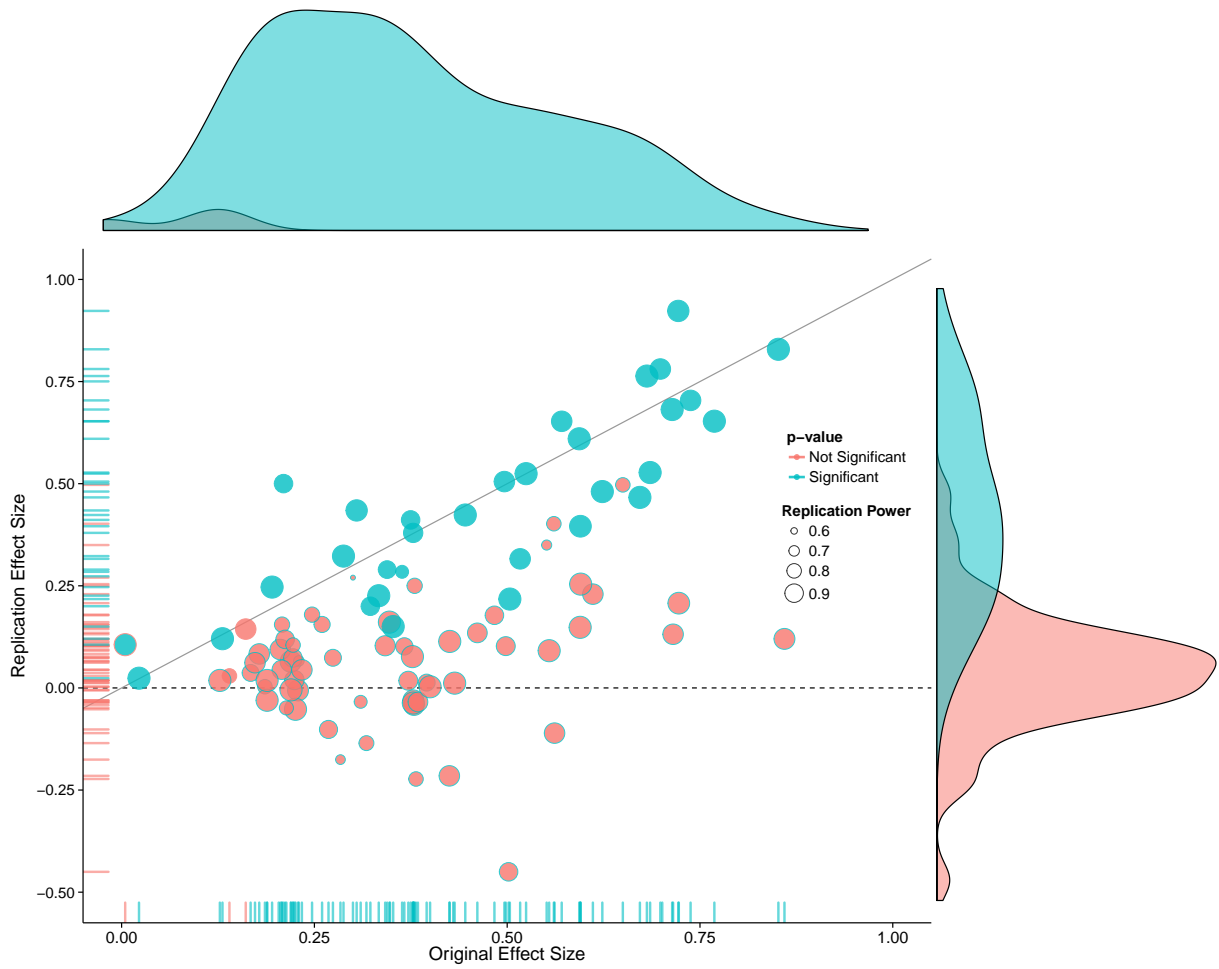
Figure 1: Figure 1 from Replication project

Figure 2: Figure 3 from Replication project

3

10. When the outcome of interest is an unordered discrete variable, we are confronted with a series of modeling decisions. Do we allow the cross alternative errors to be correlated, e.g.? Given the large set of potential statistical models to fit to an unordered DV, articulate a set of principles that can be applied when confronted with an unordered DV.