

Political Methodology Comprehensive Examination
May 2010
Department of Political Science, George Washington University

Part I

1. Following are the means and standard deviations of normally distributed national scores on four standardized tests:

Test	μ	σ
A	50	10
B	50	20
C	100	15
D	100	10

- (i) Relative to which test is a score of 70 the highest?
 - (ii) Relative to which test is a score of 85 the lowest?
 - (iii) What percentage of students will score higher than 55 on test A?
 - (iv) Out of 100 students, how many will score between 45 and 55 on test B?
 - (v) George Washington University claims that its students on average do better on test A than the national average. The assertion is based on a sample of 26 students, who had a mean score of 54.51 with $s=7.0$. Test the validity of the University's assertion.
2. You have data on international conflicts from 1946 to 2009 to examine the causes of war. The data are arranged with observations for each year and country dyad. As you prepare to present your OLS results at a conference, one of your colleagues suggests that your observations might not be independent. Please explain why this is a concern, highlighting the assumption(s) that this would implicate, diagnostic tools for determining whether this creates a problem, and solutions to remedy the problems.
3. Standard practice in political science in recent years has been to account for heteroskedasticity by the use of so-called "robust" standard errors. Are robust (Huber/White/sandwich) standard errors appropriate for all situations? If not, under what conditions would alternative remedies be preferable?
4. The Heckman selection correction model has been widely used in statistical studies of American politics and International Relations to account for the possible selection biases induced by non-random assignment to the treatment of interest. What are the strengths and weaknesses of the Heckman model? Compare the Heckman approach to at least one other approach to dealing with non-random assignment to the treatment. [Note: ignoring the issue does not constitute an approach for the purposes of this exam.]

5. 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?

Answer one of the two following questions:

6a. Some would argue that the three most important considerations in the modeling of longitudinal data are: unobserved heterogeneity, pooling, and dynamics. First, discuss what each concept means. Second, discuss why each is an important consideration when modeling longitudinal data structures. In other words, what are the consequences of NOT accounting for each? Third, discuss methods that explicitly account for each consideration. What statistical modeling specifications are available, and which specifications are most appropriate under different conditions?

6b. Say you are interested in explaining voter turnout in the 2008 presidential election (voted or did not vote). You have data consisting of n voters (level-1 units) nested within j media markets (level-2 units). Partisan strength, education, and race (black versus non-black) are level-1 (individual-level) independent variables. An independent variable measured at the media market level (level 2) is the percentage of presidential campaign advertisements that were "negative ads."

(i) Formally write out a varying intercept model. Make sure your notation is precise. Explain what this model specification accounts for compared to a plain-vanilla pooled logit/probit model. In other words, what specifically makes it different from a regular logit/probit?

(ii). Formally write out a varying intercept and varying slopes model. Make sure your notation is precise. Explain what this model specification accounts for compared to a pooled logit/probit and a varying intercept model. Be sure to explain exactly what you are modeling here with this specification.

Part II

Either submit an empirical research paper along with the exam or schedule an oral exam after the written exam.