Questions to answer in preparing for February 25th class

**The Diaz and Handa paper (JHR, 2006)**

* 1. Focus of the paper
     1. What research questions does the authors ask?
     2. What conclusions do the authors reach about the answers to their research questions?
  2. Data
     1. How does the *PROGRESA* program (now called *Oportunidades*) work? What aspects of the implementation process are important for this paper?
     2. What data set do the authors use to create comparison groups for their study?
     3. What characteristics of this data set turn out to be important, given the authors’ purpose?
     4. What are the outcome measures from the PROGRESA experiment that the authors use in their study?
     5. The authors use a “restricted sample” as well as an original sample from the data set they use to construct a comparison group. What is the difference between the two samples? What is the authors’ explanation for using the second, restricted sample? In what respects are the results of their analyses sensitive to the choice of the comparison sample? What lessons do you draw from these results?
     6. Consider the row in Table 1 called “No Bathroom.” Explain to which group the relevant figure in each column in that row applies. Explain the pattern that you see in the numbers in this row. Do they make sense?
  3. Methodology
     1. What assumptions must hold for estimates of the effects of PROGRESA participation in the ENIGH sample data to be defensibly causal in nature?
     2. How are participants assigned to PROGRESA in the ENCEL sample?
     3. On what observable characteristics does assignment to PROGRESA treatment in the ENIGH sample rely?
     4. Putting aside for a moment the details of the alternative matching methods, describe what the authors do in the paper. (Hint: See how the outcome is defined in the logit models presented on page 343.)
     5. Explain how the methodology that the authors use differs slightly from the methodology described on pages 321-324 of Chapter 12 of *Methods Matter*.
     6. Explain what the balance requirement is. What do the authors suggest doing when they fail to satisfy the balanced requirement?
     7. How do the authors examine the extent to which the biases they detect stem from differences in the questions asked in the different data sets? (Explain their analysis strategy for addressing this question.)
     8. How does “nearest neighbor” matching work? What does it mean to do this “with replacement? How is “caliper matching” different from “nearest neighbor” matching?
  4. Lessons:
     1. Explain why the data that Diaz and Handa examine provides a particularly favorable application of propensity scores?
     2. What lessons do you take away from this paper about the value of propensity score matching (PSM) for evaluating social interventions?

1. **Chapter 12 o*f Methods Matter***
2. What are the relative strengths and weaknesses of stratification versus direct control for covariates using regression analysis?
3. What does the term, region of common support, mean? What role does this play in the stratification and regression analysis approaches to dealing with sources of bias from observed variables?
4. Describe the three different methods of using propensity scores that Murnane and Willett highlight to estimate treatment effects.
5. When using Inverse Probability Weighting, which treatment units are assigned additional weight? Which control units are assigned additional weight?