Problem Set 4: Oregon Health Insurance Experiment

ECON 24450: Inequality and the Social Safety Net (Deshpande)

Due Thursday, May 27th, 9am

Read the journal article posted on Canvas, "The Oregon Health Insurance Experiment: Evidence from the First Year," by Finkelstein et al. (QJE 2013). Provide *brief* answers to the following questions.

1. What is the research question being addressed and why is it an interesting empirical question?
2. The following questions relate to the internal validity of the Oregon Health Insurance experiment (OHIE).
   (a) What is the identifying assumption of the randomized controlled trial? What are possible violations of the identifying assumption?
   (b) What do Finkelstein et al. do to test for random assignment? Refer to specific page numbers, figures, and tables as appropriate.
   (c) Why is it a problem if the treatment and control groups differentially report outcomes? Is differential reporting expected to be a problem in this context? Why or why not?
   (d) Is imperfect take-up of Medicaid in the treatment group a threat to validity? Why or why not? Compare the consequences of imperfect take-up of Medicaid in the treatment group on the OHIE to the situation we saw in class with differential attrition across treatment groups in the RAND health insurance experiment. **access to administrative information**
   (e) What other possible violations of the exclusion restriction do Finkelstein et al. consider? How do they test for these violations?
3. The following questions relate to the paper's estimation strategy.
   (a) Write down the structural equation of interest. Write down the reduced form equation and first stage equation. Why do Finkelstein et al. say that there is ambiguity about what the correct endogeneous variable is?
   (b) Why do Finkelstein et al. include controls in equation (1)?
   (c) What is the first stage estimate for ever on Medicaid? Interpret the estimate. Why is it less than one?
   (d) Describe who the complier population is for this local average treatment effect. How generalizable is this particular LATE is to the effect of Medicaid on other populations (e.g., children and pregnant women)? How generalizable is this LATE to the effect of the Medicaid expansions under the Affordable Care Act?
4. This question asks you to consider the results of this paper, plus the results of the Oregon Health Insurance Experiment more generally.
   (a) Summarize this paper's results on the effect of Medicaid on
      i. Health care utilization
      ii. Financial strain
      iii. Health

(b) The Oregon Health Insurance Experiment has received extensive media coverage. In 2011, an NPR headline read "Medicaid Makes 'Big Difference' in Lives, Study Finds." In 2013, a *Washington Post* headline read "Spending on Medicaid doesn't actually help the poor." These headlines seem contradictory. Given the results of this paper and the other OHIE papers we summarized in class, which of these headlines do you think is correct and why?

5. Data exercise.

(a) Download the data set from Canvas. Import the data set to your preferred statistical analysis program.

(b) Calculate the average 12-month survey response rate (*returned_12m*) for the control group and for the treatment group. Is there evidence of differential survey response rates between the control group and the treatment group?

(c) Consider first stage equation (4) from the paper:

$$Insurance_{ih} = \delta_0 + \delta_1 Lottery_h + X_{ih}\delta_2 + \mu_{ihj}$$

$Lottery_h$ is an indicator variable that equals 1 if the household was selected in the lottery, and 0 otherwise. X_{ih} includes household size dummy variables, survey wave dummy variables, and the interactions between survey wave dummy variables and household size dummy variables. We have provided these fixed effect terms for you in the data (variable names beginning with "ddd"). If you are using Stata, include the term "ddd*" in your regression equation to add the full set of fixed effects.

i. Estimate first stage equation (4) by OLS for each of the following outcome variables:
   A. Ever on Medicaid during the study period (*ohp_all_ever_survey*)
   B. Number of months on Medicaid (*ohp_all_mo_survey*)
   C. On Medicaid at end of study period (*ohp_all_end_survey*)

ii. In a table, report estimate coefficients $\hat{\delta}_1$. Report standard errors below each estimate. Indicate whether each estimated coefficients is significant at a 1%, 5% or 10% level.

(d) For the remainder of the exercise, use "ever on Medicaid during the study period" (*ohp_all_ever_survey*) as your measure of insurance coverage $Insurance_{ih}$.

i. Consider structural equation (3) from the paper:

$$y_{ijh} = \pi_0 + \pi_1 Insurance_{ih} + X_{ih}\pi_2 + \nu_{ijh}$$

Estimate structural equation (3) by 2SLS, using the lottery as an instrument, for each of the following outcome variables:
   A. Currently taking any prescription medications (*rx_any_12m*)
   B. Any primary care visits in last 6 months (*doc_any_12m*)
   C. Any ER visits in last 6 months (*er_any_12m*)
   D. Any hospital visits in last 6 months (*hosp_any_12m*)
   E. Any out of pocket medical expenses, last 6 months (*cost_any_oop_12m*)
   F. Owe money for medical expenses currently (*cost_any_owe_12m*)
   G. Self-reported health fair, good, very good, or excellent (*health_notpoor_12m*)
   H. Health about the same or gotten better over last 6 months (*notbaddays_tot_12m*)

ii. For each of the outcome variables listed in (4)(i), also estimate intent-to-treat equation (1) by OLS:

$$y_{ijh} = \beta_0 + \beta_1 Lottery_h + X_{ih}\beta_2 + V_{ih}\beta_3 + \epsilon_{ijh}$$

iii. In a table, report estimated coefficients $\hat{\pi}_1$ and $\hat{\beta}_1$. Report standard errors below each estimate. Indicate whether each estimated coefficient is significant at a 1%, 5% or 10% level.

iv. In a sentence, explain how to interpret the coefficient on insurance from IV estimation of equation (3). In another sentence, explain how to interpret the coefficient on the lottery dummy variable from intent-to-treat equation (1).