Economics 142
Problem Set 6
Due Thursday October 20

The data set welfare.csv contains 5,480 observations for people in an unnamed welfare experiment that created incentives for full time work. The experiment had a very strong incentive for people in the treatment group to find a full time job and leave welfare within a year of entering the experiment (when people were randomly assigned to a treatment or control group). The data set has the following variables:

treatment = 1 if in treatment group, 0 for the control group
imm = 1 if immigrant
hsgrad = 1 if high school graduate (or higher) education
agelt25 =1 if age is 25 or less
age35p = 1 if age is 35 or more
working_at_baseline = 1 if person was working when assigned to treatment/control group
anykidsu6 = 1 if any children under age 6
nevermarried = 1 if never married (all cases are single parents)

In addition there are measures of two outcomes at various times after assignment to the program:

- ft15, ft20, ft24, ft48 = set of dummies if working full time at months 15, 20, 24, and 48 after random assignment
- welfare15...welfare48 = set of dummies if receiving welfare at months 15, 20, 24, and 48 after random assignment

1. a) Estimate first stage models for the probability of working FT in months 15, 20, 24, 48, using treatment as the instrumental variable.

b) Estimate reduced form models for the probability of being on welfare in months 15, 20, 24, 48, using treatment as the instrumental variable.

c) Suppose we are interested in the causal effect of working full time on the probability of receiving welfare. Our causal model for period t is:

$$y_i(t) = \beta_0(t) + \beta_1(t)FT_i(t) + u_i(t), \tag{1}$$

where $y_i(t)$ is an indicator equal to 1 if parent i is on welfare in period $t = 15, 20, 24, 48$.

- Estimate this model by OLS for each period
- Estimate a first stage model (by OLS) for each period:

$$FT_i(t) = \pi_0(t) + \pi_1(t)T_i + \eta_i(t),$$

where T_i=1 if the parent is in the treatment group.

- Estimate a reduced form model (by OLS) for each period:

$$y_i(t) = \gamma_0(t) + \gamma_1(t)T_i + \eta_i(t).$$

- Estimate the causal model (1) by 2SLS for each period using T_i as an instument for FT status in period t. Verify that the 2SLS estimate for each period is the ratio of the reduced form and first stage coefficients.

2. Treating $FT_i(t)$ as the "D_i" variable, find the fractions of always takers, never takers, and compliers in each month.

3. Using the techniques described in Lecture 16, find the mean characteristics of the compliers in month 15: i.e., means of the variables imm, hsgrad, agelt25, age35p, working_at_baseline, anykidsu6, and nevermarried. Compare these to the characteristics of always takers and never takers in the same month.