Please estimate the effects of the MLDA on alcohol consumption and on mortality. The figures and tables you create will form the core of the second paper. The dataset "NHIS.csv" is a sample drawn from the National Health Interview Sample Adult Files 1997-2007. It includes the following variables: HS_diploma, Hispanic, white, black, uninsured, employed, married, workin_lw, going_school, male,days_21 (Days to 21st birthday), drinks_alcohol (Reports they drink alcohol), AGE_yrs, perc_days_drink (Percent of days on which they report drinking). Each row on the dataset is a record for an individual. The file mortality data.csv contains mortality rates per 100,000 by age overall and broken down by primary cause of death.

1. Create a table of regression estimates of the level of each of the demographic variables for people just under 21 and how they change at 21 (Hint: Same regression as for part 5 with demographic variable as outcome). This is the table that documents that the people just under 21 are very similar to those just over 21 at least on their observable characteristics similar to the balance table in an experiment.

2. Make a figure with the age profile of deaths due to all causes.

3. Make a figures of the age profile for each cause of death. Where possible include more than one age profile in a figure.

4. Make a regression table with estimates of the increase in deaths overall and in each sub category of deaths. Put one regression per column.

5. Does the MLDA reduce death rates? Are the reductions in causes of death that are likely to be affected by alcohol consumption?

6. Estimate the effect of drinking on overall deaths and each cause. Use a two sample instrumental variables approach (reduced form/first stage).
7. Estimate the standard errors for each IV estimate using the delta method (formula provided in class).

8. Do you believe the assumptions under which the MLDA is a valid instrument for alcohol consumption are met? Assess them one at a time.