

Assignment 1: Hansen Alcohol DWI AER – Due: 2/27/20 in class

Directions: Please turn in your answers on separate paper, typed, and **beautifully written** with **beautiful tables** automated using -estout- or -outreg2- and **beautiful figures**.

Background and set up

1. Write a brief (1-3 paragraph) summary of Hansen's project. What is his question and the theory behind it? What data does he use and what his research design? What does he find?
2. Create a new directory, labeled appropriately, with the following subdirectories: /Article, /Do, /Data, /Misc, /Figures, /Tables, /Writing. Place your data /data.
3. Create a new do file for this project and store it in the /Do subdirectory. Your do file should have a commented out header stating the name of the file, a description of the file, the file's authorship, and the date of last time you worked on the do file.
4. Include in the do file the following lines:
 - a. clear
 - b. and "cd ..." which will change working directory to /do.
5. Download Hansen_dwi.dta from canvas and store in the /data subdirectory.
6. The outcome variable is "recidivism" or "recid" which is measuring whether the person showed back up in the data within 4 months.

Replication

7. We are going to only focus on the first cutoff – a blood alcohol content of 0.08 – for this replication. So create a dummy equaling 1 if bac1 \geq 0.08 and 0 otherwise.
8. The first thing to do in any RDD is look at the raw data and see if there's any evidence for manipulation. If people were capable of manipulating their blood alcohol content (bac1), what would that look like? And do we see that in these data? Recreate Figure 1 using the bac1 variable as your measure of blood alcohol content. Do you find evidence for sorting on the running variable?
9. The second thing we need to do is check for covariate balance. You will need to estimate equation (1) with white, male, age and accident (acc) as dependent variables. Are the covariate balanced at the cutoff?
10. Recreate Figure 2 panel A-D. You can use the -cmogram- command in Stata to do this. Fit both linear and quadratic with confidence intervals. Discuss what you find and compare it with Hansen's paper.
11. Estimate equation (1) with recidivism (recid) as the outcome. This corresponds to 3 column 1, but since I am missing some of his variables, your sample size will be the entire dataset of 214,558. Nevertheless, replicate Table 3, column 1. Your table should have three columns:
 - a. Column 1: control for the bac1 linearly
 - b. Column 2: interact bac1 with cutoff linearly
 - c. Column 3: interact bac1 with cutoff linearly and as a quadratic
 - d. For all analysis, use heteroskedastic robust standard errors.
12. Recreate the top panel of Figure 3 according to the following rule:
 - a. Fit linear fit using only observations with less than 0.15 bac on the bac1
 - b. Fit quadratic fit using only observations with less than 0.15 bac on the bac1