**Assignment 4: RDD**

**Santiago Andrés León Rincón**

**Sunday, June 14th,2020**

**Directions**: Please turn in your answers on separate paper, typed, and **beautifully written** with **beautiful tables** and **beautiful figures**.**[[1]](#footnote-1)**

**Github repo and summary (worth 2 points)**

1. Download Hansen\_dwi.dta from github at the following address.

use https://github.com/scunning1975/causal-inference-class/raw/master/hansen\_dwi, clear

Create a new github repo named “RDD”. Inside the RDD directory, put all the subdirectories we’ve discussed in class. Post the link to the repo so I can see it’s done as discussed in your assignment. Save the Hansen\_dwi.dta file into your new /data subdirectory. Note: The outcome variable is “recidivism” or “recid” which is measuring whether the person showed back up in the data within 4 months.

**Summary**

Hansen’s (2015) article asks if whether the harsher punishments and sanctions on driving under the influence (DUI) are an effective way to reduce recidivism. For this purpose, the author uses administrative records on DIU stops in the state of Washington (WA) from 1995 to 2011 which total up to 512,964 observations. It is important to consider that above a 0.08 blood alcohol content (BAC) is considered a DUI while a BAC over 0.15 is considered an aggravated DUI. Therefore, the author uses BAC as the running variable which enables the development of a quasi-experiment.

The method that he used to estimate the effect of the punishments on recidivism was a regression discontinuity design (RDD) thanks to the different cutoffs. Given the above, Hansen uses a local linear regression discontinuity with a kernel a rectangular kernel as following:

Before he applies this, he needed to prove the assumption of the continuity of the underlying conditional regression and distribution functions. For this he uses a McCarary (2008) method which suggested little evidence of endogenous sorting to one side of either of the thresholds. In addition, for robustness in the previous conclusion, he uses Frandsen (2013) method which estimated the same lack of evidence of endogenous sorting.

After all of this, Hansen found out that having a BAC above the DUI cutoffs reduces by up to 2 percentage points the recidivism. While having BAC exceeding the aggravated DUI threshold diminishes recidivism by an extra percentage point.

As a conclusion, the paper says that the sanctions experienced by drunk drivers are effective in reducing recidivism in short and long term.

**Replication (worth 6 points)**.[[2]](#footnote-2)

1. In the United States, an officer can arrest a driver if after giving them a blood alcohol content (BAC) test they learn the driver had a BAC of 0.08 or higher. We will only focus on the 0.08 BAC cutoff. We will be ignoring the 0.15 cutoff for all this analysis. Create a dummy equaling 1 if **bac1**>= 0.08 and 0 otherwise in your do file or R file.

Code used:

gen over\_bac=1 if bac1>=0.08

replace over\_bac=0 if over\_bac==.

1. The first thing to do in any RDD is look at the raw data and see if there’s any evidence for manipulation (“sorting on the running variable”). If people were capable of manipulating their blood alcohol content (bac1), describe the test we would use to check for this. Now evaluate whether you see this in these data? Either recreate Figure 1 using the bac1 variable as your measure of blood alcohol content or use your own density test from software. Do you find evidence for sorting on the running variable?

I would use two tests to check if there is any evidence of people manipulating their BAC. The first one is the so-called eyeball test, which means creating a histogram that on its vertical axis refers to the frequency of observations and on the horizontal axis the BAC.

After this I will develop the McCrary (2008) test to confirm the conclusions suggested by the eyeball test. This formal test suggests proving the null hypothesis of the continuity of the density of the covariate that underlies the assignment at the discontinuity point, against the alternative of a jump in the density function at that point (NYU Wagner, n.d.).

**Eyeball test**

**Figure 1: Density distribution BAC**

**Imagen que contiene mapa, texto

Descripción generada automáticamente**

The vertical red lines represent the legal thresholds at 0.08. The bin width is 0.001, same as Hansen(2015).

The graph does not show any hints of manipulation.

**McCrary Test**

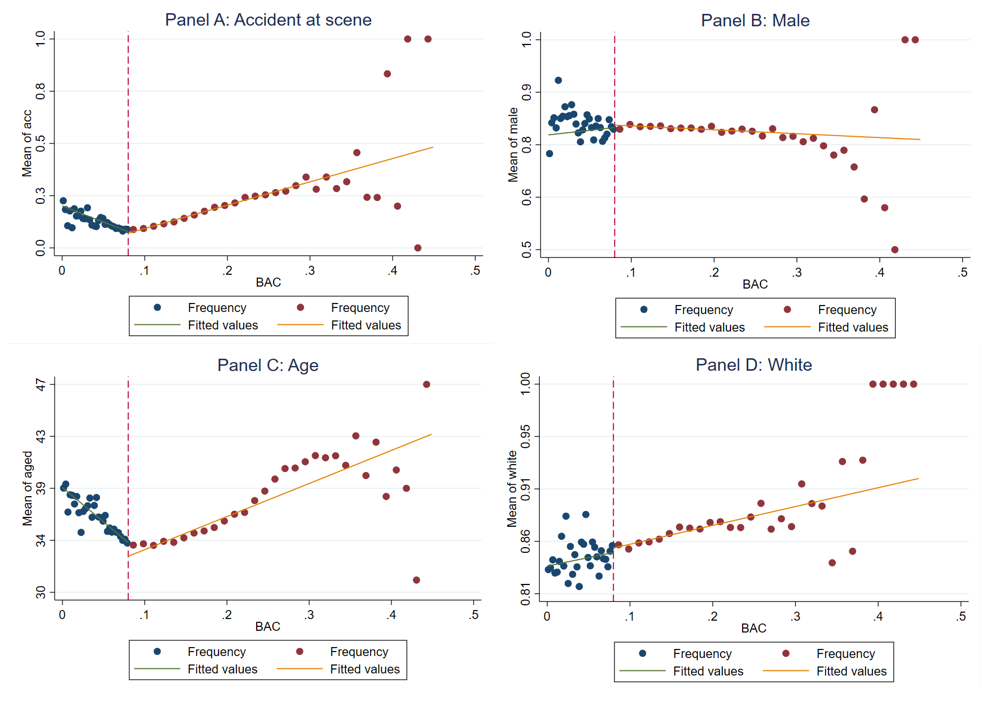
**NEED TO BE CORRECTED**

The test shows a P-Value of 0.472296, which means the null hypothesis is not rejected, therefore it suggests there is no manipulation in the data.

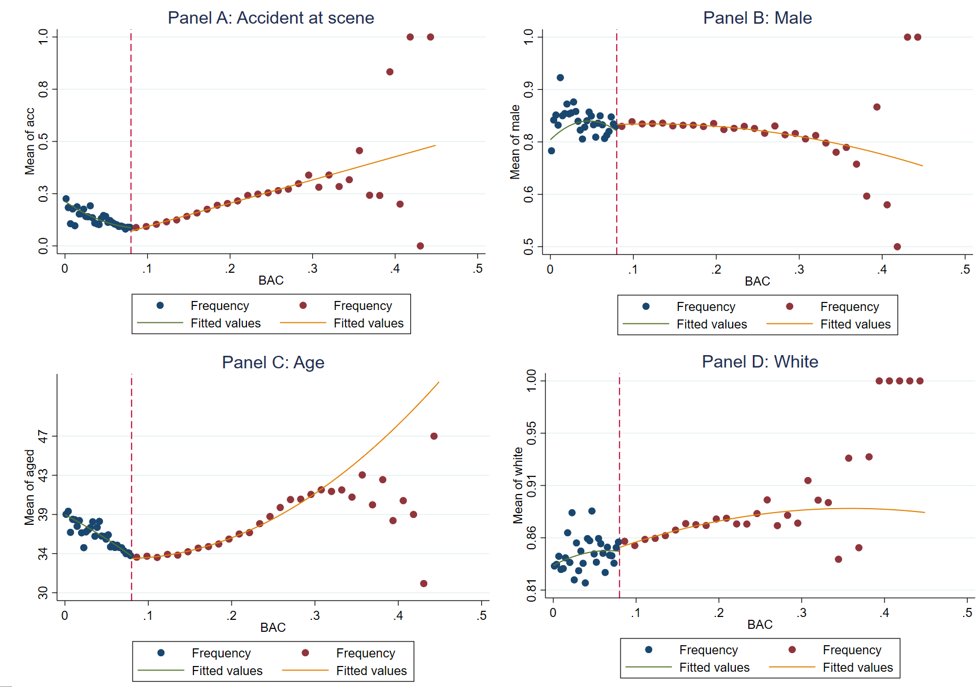
1. The second thing we need to do is check for covariate balance. Recreate Table 2 Panel A but only white male, age and accident (acc) as dependent variables. Use your equation 1) for this. Are the covariates balanced at the cutoff? It’s okay if they are not exactly the same as Hansen’s.
2. Recreate 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.

**Figure 2: BAC and Control variables**

**Linear Fit**

****

**Quadratic fit**

****

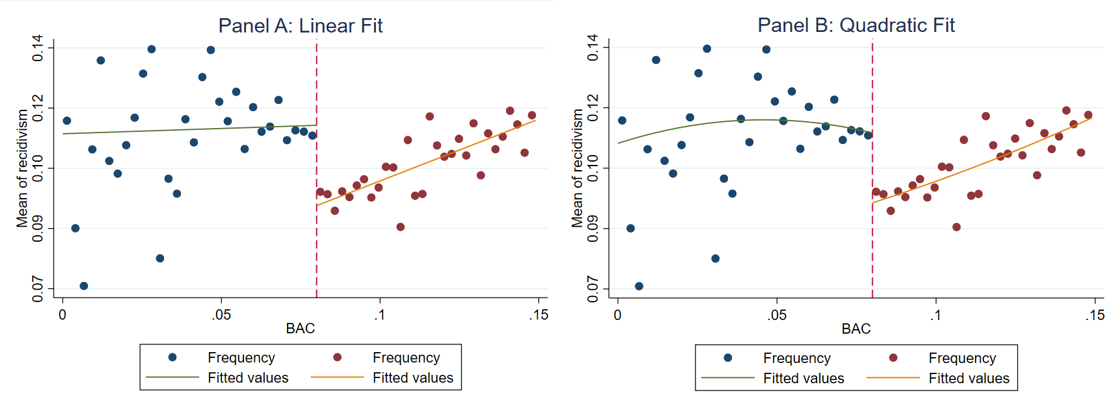
ITS MISSING THE DISCUSSION

1. Estimate equation (1) with recidivism (recid) as the outcome. This corresponds to Table 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, Panels A and B. Note that these are local linear regressions and Panel A uses as its bandwidth 0.03 to 0.13. But Panel B has a narrower bandwidth of 0.055 to 0.105. Your table should have three columns and two A and B panels associated with the different bandwidths.:
   1. Column 1: control for the bac1 linearly
   2. Column 2: interact bac1 with cutoff linearly
   3. Column 3: interact bac1 with cutoff linearly and as a quadratic
   4. For all analysis, use heteroskedastic robust standard errors.

IT IS ONLY MISSING TO CONFIRM WITH SCOTT IF REGRESSIONS ARE CORRECTLY SPECIFIED AND ALSO MERGING THE TWO BIG PANELS

1. Recreate the top panel of Figure 3 according to the following rule:
   1. Fit linear fit using only observations with less than 0.15 bac on the bac1
   2. Fit quadratic fit using only observations with less than 0.15 bac on the bac1

**BAC and Recidivism**



**References**

Regression Discontinuity: Advanced Topics NYU Wagner Rajeev Dehejia

1. Again, my preference is that you attempt to create automated tables and automated figures as much as you can. I’ve placed a simple estout program called ols.do in the estout subdirectory. You just need to edit. [↑](#footnote-ref-1)
2. [↑](#footnote-ref-2)