**Assignment 4: Due Saturday, June 13th 2020[[1]](#footnote-1)**

**Student: Lewis Enrique Polo Espinosa**

**Directions**: Please turn in your answers on separate paper, typed, and **beautifully written** with **beautiful tables** and **beautiful figures**.

**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. 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.

<https://github.com/Lewis0126/RDD#rdd>

<https://github.com/Lewis0126/RDD.git>

1. **In the writing subdirectory, place your assignment. For the first part, read Hansen’s paper in the articles directory of the main class github entitled “Hansen AER”. Briefly summarize this paper. What is his question? What data does he use? What is his research design? What does he find?**

The article evaluates the fact if punishments and sanctions are effective in reducing drunk driving in order give recommendations of enforcement and punishment that might improve social welfare. To answer that question, the author employs administrative records on 512,964 DUI stops from the state of Washington from 1995 to 2011. Then, under a context of quasi-experimental approach and using regression discontinuity estimates he exploits a discrete threshold (0.08) to determine both the current and potential future punishments for drunk drivers, after getting a causal effect of having a blood alcohol content(BAC) over 0.08. The main result suggests that punishments and sanctions associated with BAC limits reduce future drunk driving, providing validity information about the effectiveness of current BAC thresholds in reducing drunk driving.

**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 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.**

**Table 1. Distribution of cutoff variable**



1. **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), describe the test we would use to check for this. Now evaluate whether you see this 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?**

According to Table 2, it evidences data manipulation on the running variables employing the MacCrary density test, which is a statistical test on the difference around the cutoff and it shows this difference is significant at the 5% level suggesting there is evidence for manipulation. Likewise, Figure 1 plots the number of observations into bins which allows to visualize this manipulation around the threshold and this graph supports the test where the running variable was manipulated.

**Table 2. McCrary density test**



**Figure 1. Plot of data manipulation**

Imagen que contiene luz

Descripción generada automáticamente

1. **The second thing we need to do is check for covariate balance. Recreate Table 2 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.**

According to Table 3, this indicates that the covariates are balanced at the cutoff. Nonetheless, if we use the complete sample the covariates are not going to be balanced because of the effect of the treatment.

**Table 3. Replication Table 2**



1. **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.**

According to Figure 2 (A and B) and Hansen’s paper, we can observe that demographic characteristics such as age, male, white and accident are stable across the thresholds (0.08 and 0.15) which gives additional credibility that the regression discontinuity design can deliver unbiased estimates variables. In Figure 2.a we fit linear with confidence intervals and in Figure 2.b fit quadratic with confidence intervals.

**Figure 2. Panel A-D BAC and Characteristics**

**A. Fit Linear**

**(A) Accident (B) Male**



**(C) Age (D) White**



**B. Fit Quadratic**

**(A) Accident (B) Male**

****

**(C) Age (D) White**

****

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.:**
2. **Column 1: control for the bac1 linearly**
3. **Column 2: interact bac1 with cutoff linearly**
4. **Column 3: interact bac1 with cutoff linearly and as a quadratic**
5. **For all analysis, use heteroskedastic robust standard errors.**

**Table 4. Estimates for the Effect of BAC on Recidivism**



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**

**Figure 3. BAC and Recidivism (Fit linear)**



* 1. **Fit quadratic fit using only observations with less than 0.15 bac on the bac1**

**Figure 4. BAC and Recidivism (Fit quadratic)**



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. Much of this advice applies to Stata commands, but you can check the R files for lmb.r to see ways of doing the same in R. [↑](#footnote-ref-2)