**Assignment 4: Due Sunday, June 14th 2020**

**Oriana Alvarez Vos**

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

1. **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**.

<https://github.com/OrianaAlvarez2020/RDD.git>

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

The paper offers quasi-experimental evidence on the effectiveness of blood alcohol content (BAC) thresholds that it is a primary policy used in curbing drunk driving. He used administrative records on 512.964 DUI BAC tests in the state of Washington (WA) from 1995 to 2011.

The author exploit threshold that determine both the current as well as potential future punishment for drunk drivers. BAC measured above 0.08 is considered a DUI while a BAC above 0.15 is considered an aggravated DUI. He used a regression discontinuity as a research design to test the effect of the punishment imposed at BAC threshold on recidivism. The paper found evidence that having a BAC above either the 0.08 DUI threshold or the 0.15 aggravated DUI is associated with reduced repeat drunk driving both in the short and long term. Having a BAC over 0.08 correspond with a 2 percent point decline in repeat drunk driving over the next four years. Likewise, having a BAC over 0.15 enhanced punishment limit is associated with an additional 1 percentage point decline in repeat drunk driving. The author also found that 10 percent increase in sanctions and punishments is associated with 2.3 percent decline in drunk driving and that the increase in type of sanction at each threshold and the timing of the reductions of drunk driving are most consistent with deterrence serving as a primary channel in the reduction in recidivism.

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

**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.**

gen d=1

replace d=0 if bac1<0.08

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 could manipulate 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?**

The test will be MacCrary Density Test to see whether there is bunching of units at the cut off. The null hypothesis the density should be continuous at the cutoff point and under the alternative, the density should increase at the kink.



There is no evidence that there was manipulation in the running variable at the cut off according with the rddensity test.

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 the same as Hansen’s.**



There is a lack of significance in the regressions coefficients which imply that the continuity conditions in the individual’s characteristics hold. In other words, the observable individuals’ characteristics X which also determine the outcome variable Y, did not jump in the cut off points. This means that the individuals just to the left of the cut-off points are very similar to individuals just to the right of the cut-off.

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.**

One of the key arguments to use this command is to decide the kind of polynomial to fit left and right of the cutoff. What it is possible to see in these two graphs there are strong trends in the running variable (BAC) that is necessary to use the quadratic fit linear as it has been shown in the Hansen´s paper. When I use the linear fit of the data left and right of the cutoff on Hansen´s dataset there is an influence of outliers far from the actual cutoff in the estimate of causal effect at the cutoff for example in Panel B(male). Hansen´s report a quadratic fit because it is the polynomial that best fit his data.



**Figure 2: BAC and Characteristics (linear fit)**



**Figure 2: BAC and Characteristics (quadratic fit)**

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. **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



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



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)