Replication1

In the beginning, put a link to my RDD repository: https://github.com/LZHHH4869/RDD.

The second part is to summarize Hansen's paper by several parts. Drunk driving has been a significant factor in hundreds of thousands of traffic deaths since 1975. The punishment was determined by the strict blood alcohol content (BAC) and previous criminal convictions, as the author analyzed. The author used thresholds to determine the severity of penalties and to test whether the more severe penalties and sanctions experienced by offenders under the BAC threshold were effective in reducing the driving under the influence (DUI) by testing the effects of tougher penalties and sanctions on DUI.

The author used administrative records of 512,964 DUI tests in the state of Washington from 1995 to 2011. He used a threshold of 0.08 to determine DUI and a threshold of 0.15 to determine 'aggravated DUI'. In addition, the article limited the attention to people above the legal drinking age.

For the research design, at cutoff points of DUI and aggravated DUI, regression discontinuity design (RDD) was allowed to test the effects of punishment imposed at the BAC threshold on recidivism. For the regression model, the author used the local linear regression discontinuity design to estimate BAC above DUI or aggravated DUI threshold for the recidivism, allowing the slope at the discontinuity to change. The main results based on local linear regression discontinuity of a rectangular nuclear design.

Estimates from RDD showed a significant reduction in recidivism rates due to thresholds of 0.08 (DUI) and 0.15 (aggravated DUI). Also, BAC levels above the legal limit of 0.08 and 0.15 were associated with lower recidivism rates by grouping drivers who had ever taken a breathalyzer test. Then, by the estimated effect of BAC above the aggravated DUI threshold on recidivism, it was found that BAC in that area was associated with varying degrees of reduced recidivism among all potential offenders or among those who have not been tested and those who has been tested. Then, by analyzing the effect of BAC above the DUI threshold on the more detailed definition of recidivism, the results show that BAC above the DUI threshold leads to less or more concentrated drunk driving. Furthermore, the author further studies several mechanisms, including deterrence, incapacitation and rehabilitation. Studies showed that BAC thresholds played a role in the effect of drunk driving primarily through deterrence operations, although some effects through incapacitation and rehabilitation operations cannot be completely excluded. In general, these results showed that harsher penalties and sanctions associated with BAC restrictions would reduce future drunk driving.

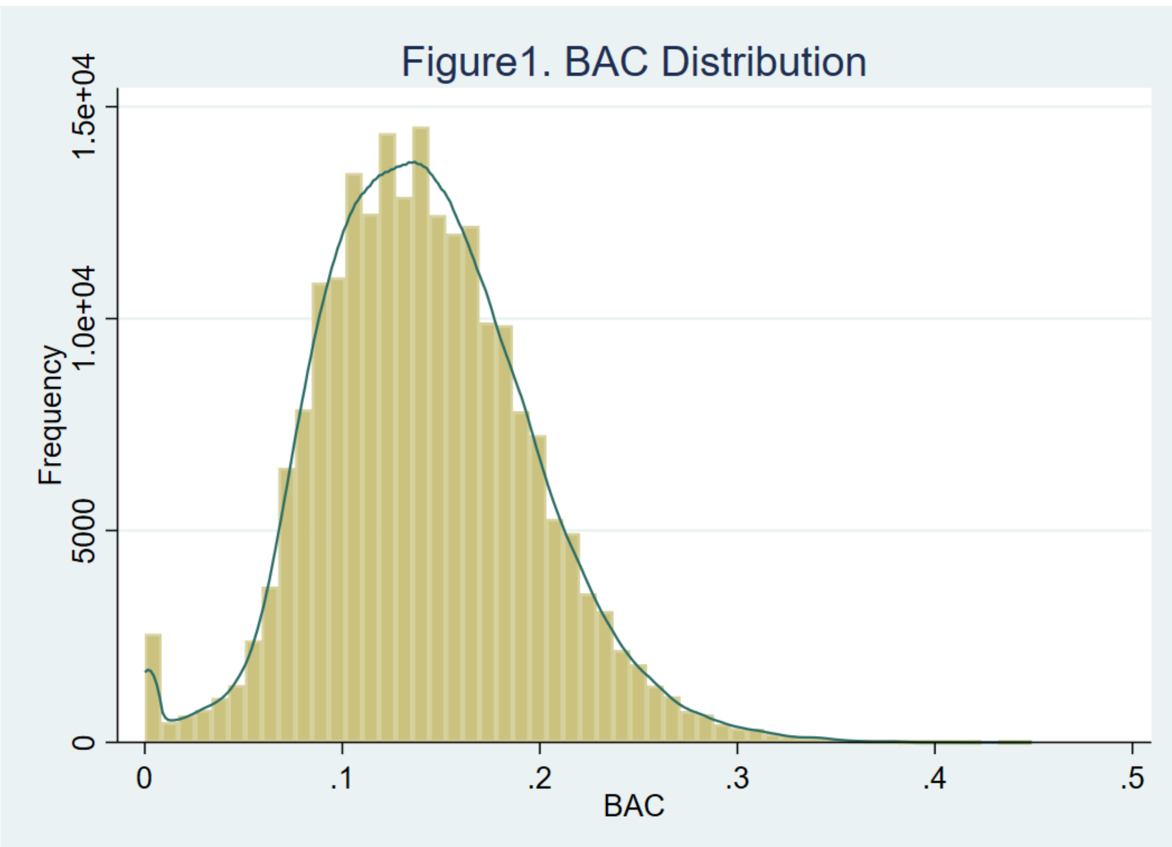
```
. cd C:\Users\Yzhizhang\Desktop\xx
C:\Users\Yzhizhang\Desktop\xx

. use "E:\DOU\UT Au\Spring2021\Casual\作业\Replication1\hansen_dwi.dta"
```

In the paper, we know that 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, so in question3 we will create a dummy equaling 1 if bac1>= 0.08 and 0 otherwise

```
. gen DUI=(bac1>=0.08)
```

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), so in question4, we should do a test to check for this.

```
. histogram bac1, frequency kdensity ytitle(Frequency) xtitle(BAC) title(Figure1.
 BAC Distribution
> )
(bin=53, start=0, width=.0084717)

. graph export bac.png
(file bac.png written in PNG format)
```



We can also use McCrary density test as shown below.

```
. rddensity bac1, c(0.08)
Computing data-driven bandwidth selectors.

Point estimates and standard errors have been adjusted for repeated observations.
(Use option nomasspoints to suppress this adjustment.)

RD Manipulation test using local polynomial density estimation.

      c =       0.080 | Left of c  Right of c          Number of obs =       214558
----------------------+----------------------          Model         =  unrestricted
     Number of obs    |     23010      191548          BW method     =          comb
Eff. Number of obs    |     14727       28946          Kernel        =    triangular
    Order est. (p)    |         2           2          VCE method    =     jackknife
    Order bias (q)    |         3           3
       BW est. (h)    |     0.023       0.023

Running variable: bac1.
---------------------------------------------
               Method |         T        P>|T|
----------------------+----------------------
               Robust |   -0.1387       0.8897
---------------------------------------------

P-values of binomial tests. (H0: prob = .5)
-----------------------------------------------------
 Window Length / 2    |        <c         >=c |  P>|T|
----------------------+-----------------------+------
               0.000  |       909           0 | 0.0000
               0.000  |       909           0 | 0.0000
               0.000  |       909           0 | 0.0000
               0.000  |       909           0 | 0.0000
               0.000  |       909           0 | 0.0000
               0.000  |       909           0 | 0.0000
               0.000  |       909           0 | 0.0000
               0.000  |       909           0 | 0.0000
               0.000  |       909           0 | 0.0000
               0.000  |       909           0 | 0.0000
-----------------------------------------------------
```

Figure 1 is a histogram displaying the number of observations in 0.08 BAC level. This distribution is very similar to Hansen's finding. Similarly, the distribution of BAC shows little evidence of endogenous sorting to one side of the threshold studied. On the other hand, after operating McCrary density test, it shows that the p-value is 0.8897 at the 0.08 threshold, which reveals no evidence of manipulation.

The second thing we need to do is check for covariate balance in question5.

```
. eststo clear
```

```
. quietly eststo: rdrobust male bac1, c(0.08) h(0.05) kernel(uniform)

. quietly eststo: rdrobust white bac1, c(0.08) h(0.05) kernel(uniform)

. quietly eststo: rdrobust aged bac1, c(0.08) h(0.05) kernel(uniform)

. quietly eststo: rdrobust acc bac1, c(0.08) h(0.05) kernel(uniform)

. esttab
```

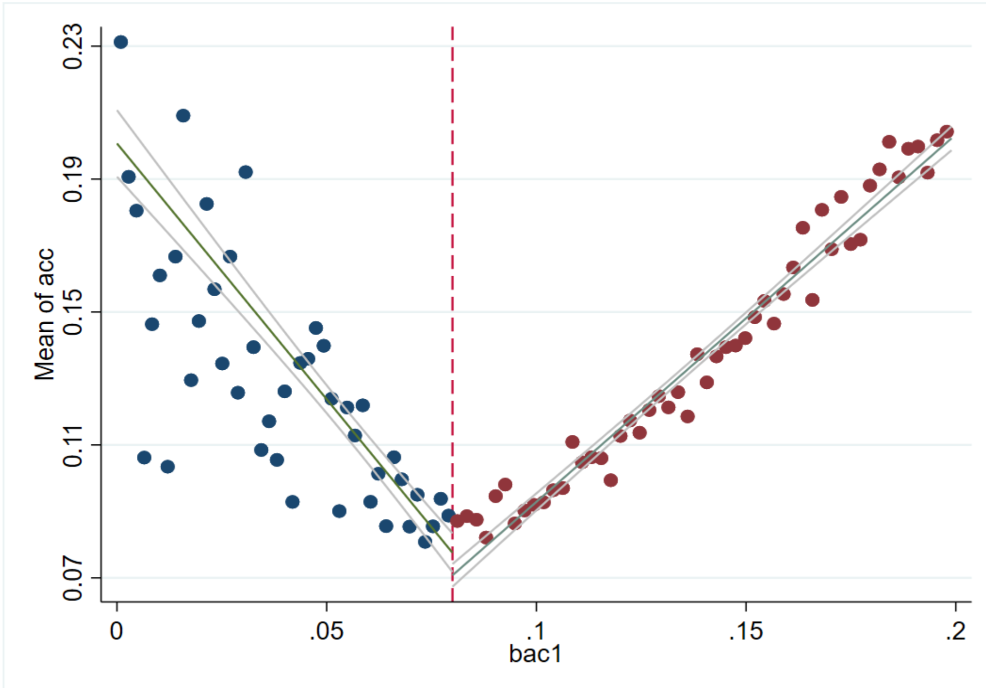
	(1) male	(2) white	(3) aged	(4) acc
RD_Estimate	0.00618 (1.08)	0.00570 (1.14)	-0.140 (-0.85)	-0.00335 (-0.82)
N	214558	214558	214558	214558

```
t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001
```

This table contains regression discontinuity based estimates of the effect of having BAC above the 0.08 thresholds on four characteristics: white, male, age, and accident at scene. The results from these regressions show that all of the estimated coefficients are not statistically significant at 5% level, even at 10% level. This is consistent with the result from Hansen's paper. And this indicates that we cannot reject the null that the predetermined characteristics are unrelated to the BAC cutoffs for DUI.

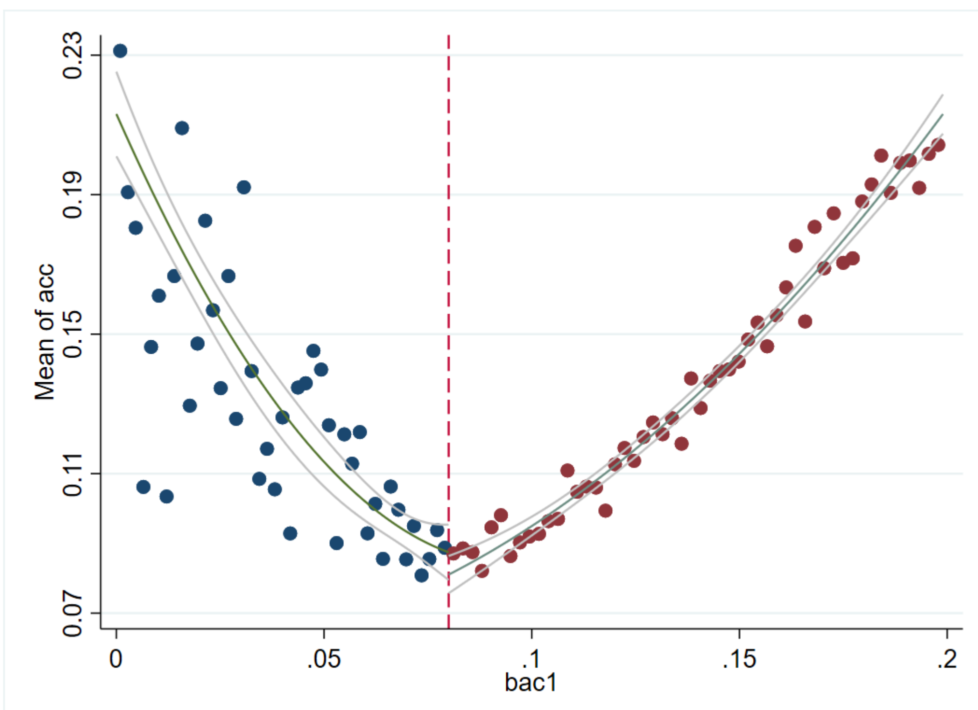
Next in question6, we will create several graphs to present some predetermined characteristics and corresponding fitted regression lines which should remain unchanged across the punishment thresholds if offenders or police are unable to manipulate the running variable. This time we should fit both linear and quadratic with confidence intervals.

```
. quietly cmogram acc bac1 if bac1<=0.2, cut(0.08) scatter line(0.08) lfitci

. graph export acc1.png
(file acc1.png written in PNG format)
```



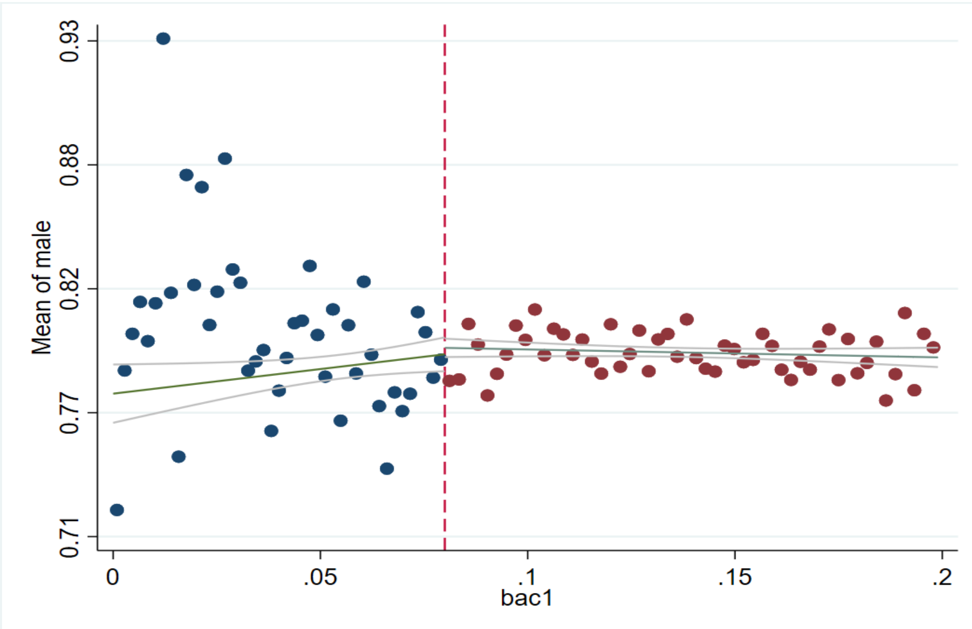
Panel A1. Accident at scene

```
. quietly cmogram acc bac1 if bac1<=0.2, cut(0.08) scatter line(0.08) qfitci

. graph export acc2.png
(file acc2.png written in PNG format)
```



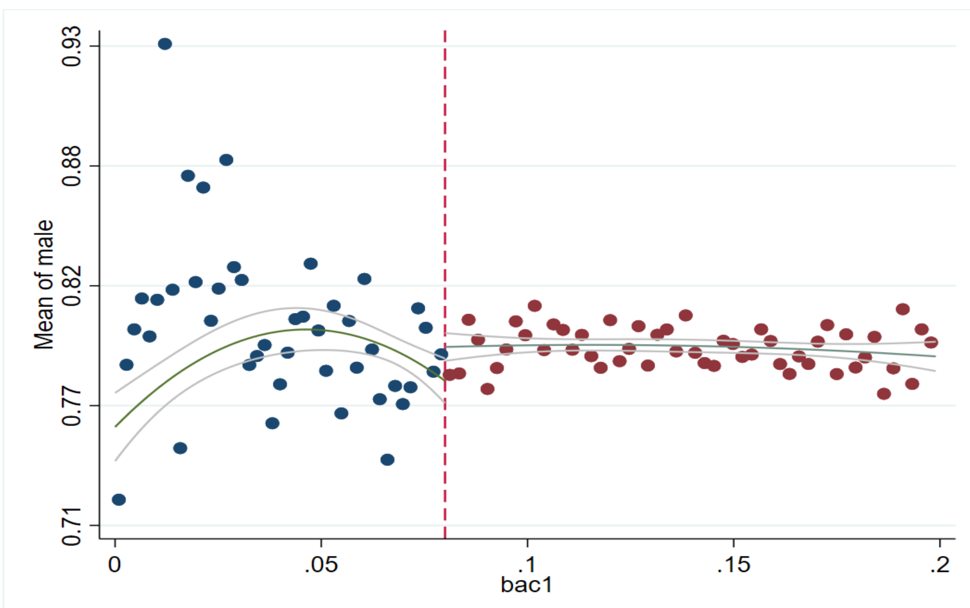
Panel A2. Accident at scene

```
. quietly cmogram male bac1 if bac1<=0.2, cut(0.08) scatter line(0.08) lfitci

. graph export male1.png
(file male1.png written in PNG format)
```



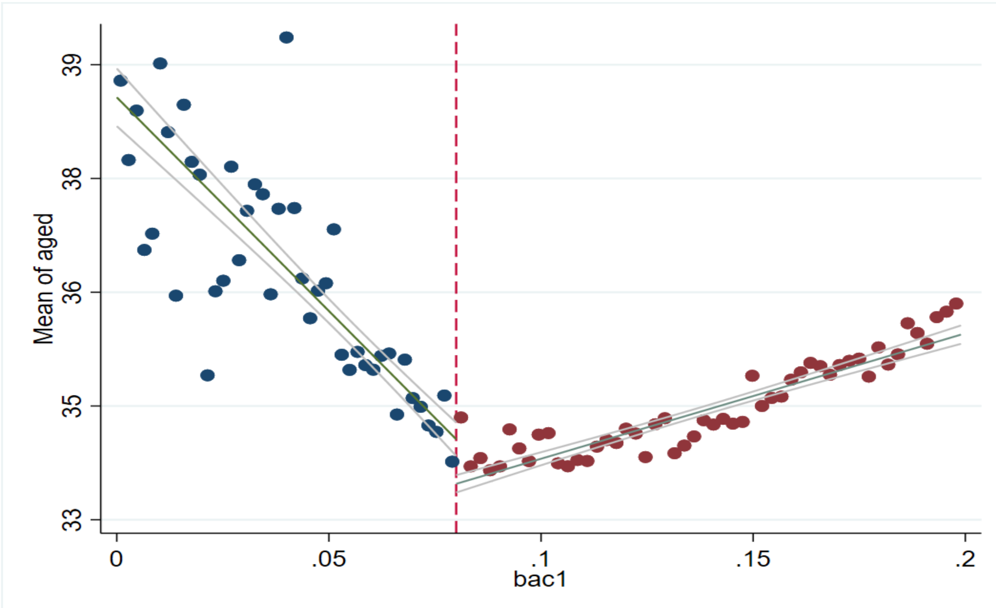
Panel B1. Accident at scene

```
. quietly cmogram male bac1 if bac1<=0.2, cut(0.08) scatter line(0.08) qfitci

. graph export male2.png
(file male2.png written in PNG format)
```



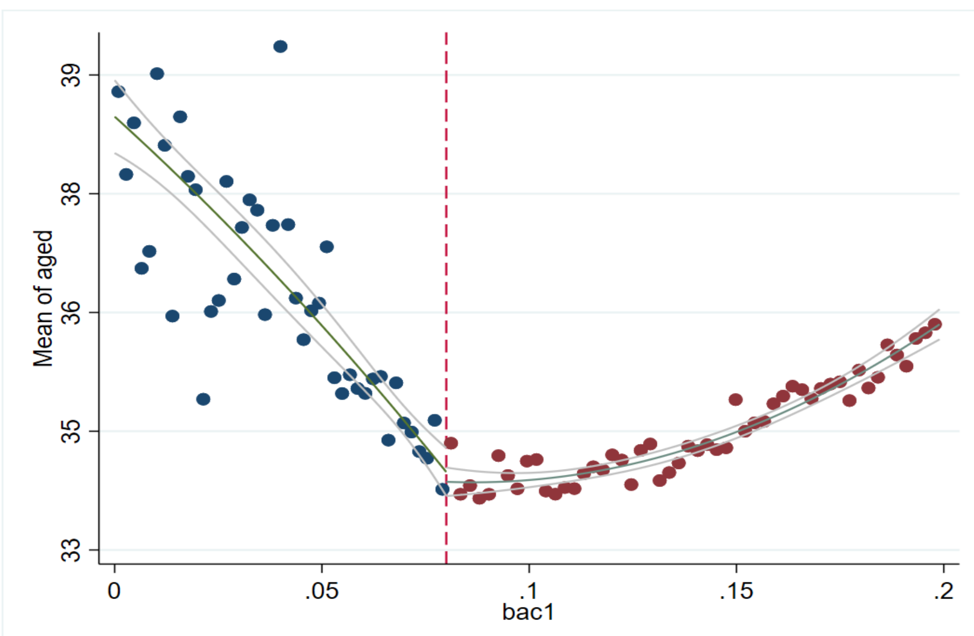
Panel B2. Accident at scene

```
. quietly cmogram aged bac1 if bac1<=0.2, cut(0.08) scatter line(0.08) lfitci

. graph export age1.png
(file age1.png written in PNG format)
```



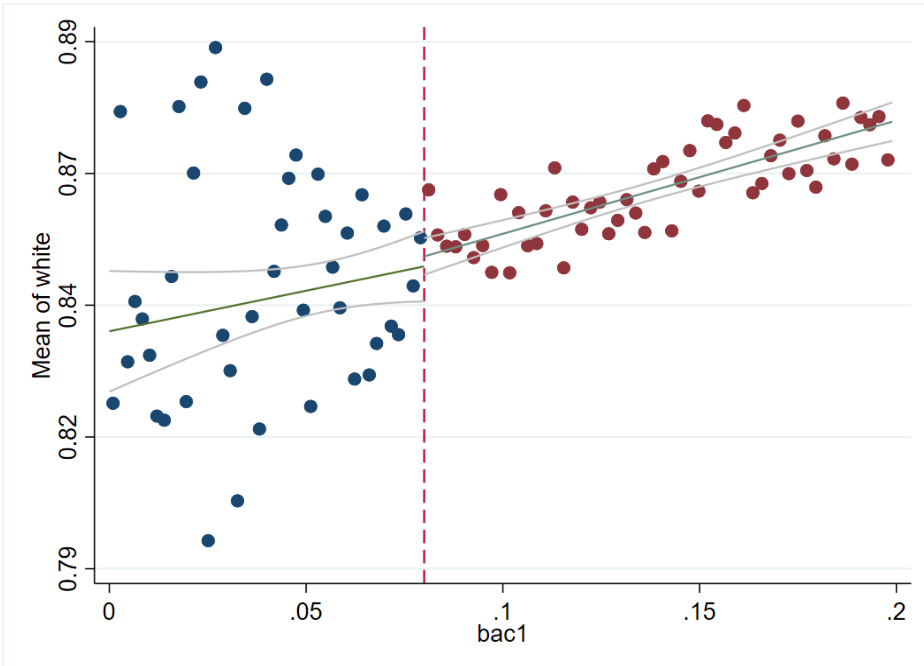
Panel C1. Accident at scene

```
. quietly cmogram aged bac1 if bac1<=0.2, cut(0.08) scatter line(0.08) qfitci

. graph export age2.png
(file age2.png written in PNG format)
```



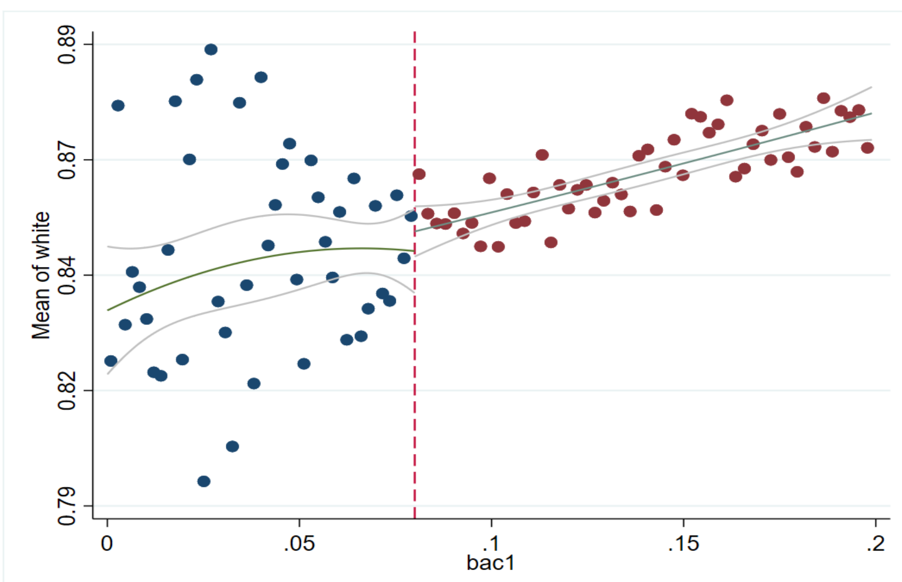
Panel C2. Accident at scene

```
. quietly cmogram white bac1 if bac1<=0.2, cut(0.08) scatter line(0.08) lfitci

. graph export white1.png
(file white1.png written in PNG format)
```



Panel D1. Accident at scene

```
. quietly cmogram white bac1 if bac1<=0.2, cut(0.08) scatter line(0.08) qfitci

. graph export white2.png
(file white2.png written in PNG format)
```



Panel D2. Accident at scene

In general, the graphs made by myself are similar to those in Hansen's finding. To be specific, in panel A, the beginning of the trend in the BAC<0.08 is higher than that in the article. And in the panel B, when BAC<0.08, the trend tends to be increasing smoothly, which is a bit different from the smoothly decreasing trend in the article's panel B. Furthermore, in the panels fitted quadratic with confidence intervals, before 0.08 threshold there exists a trend like a quadratic function form (not just increasing), which is a little different from Hansen's finding. The trends in other three panels are similar to those fitted linearly with confidence intervals and those in Hansen's article.

In question7, we will estimate equation (1) from the paper with recidivism (recid) as the outcome. these are local linear regressions and Panel A uses as its bandwidth 0.03 to 0.13, while Panel B has a narrower bandwidth of 0.055 to 0.105. This reports the estimated effect of having BAC over the 0.08 threshold for all drivers. Panel A:

```
. eststo clear

. quietly eststo: reg recidivism bac1 male white acc aged if bac1>=0.03 & bac1<=0.
13, robust

. quietly eststo: rdrobust recidivism bac1, c(0.0064) h(0.03 0.13) kernel(uniform)
 p(1) covs(male
> white acc aged)

. quietly eststo: rdrobust recidivism bac1, c(0.0064) h(0.03 0.13) kernel(uniform)
 q(2) covs(male
> white acc aged)

. esttab, keep(bac1 RD_Estimate)
```

	(1) recidivism	(2) recidivism	(3) recidivism
bac1	-0.0755 (-1.56)		
RD_Estimate		0.0240 (1.08)	0.0240 (1.08)
N	89967	214558	214558

```
t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001
```

Panel B:

```
. eststo clear
```

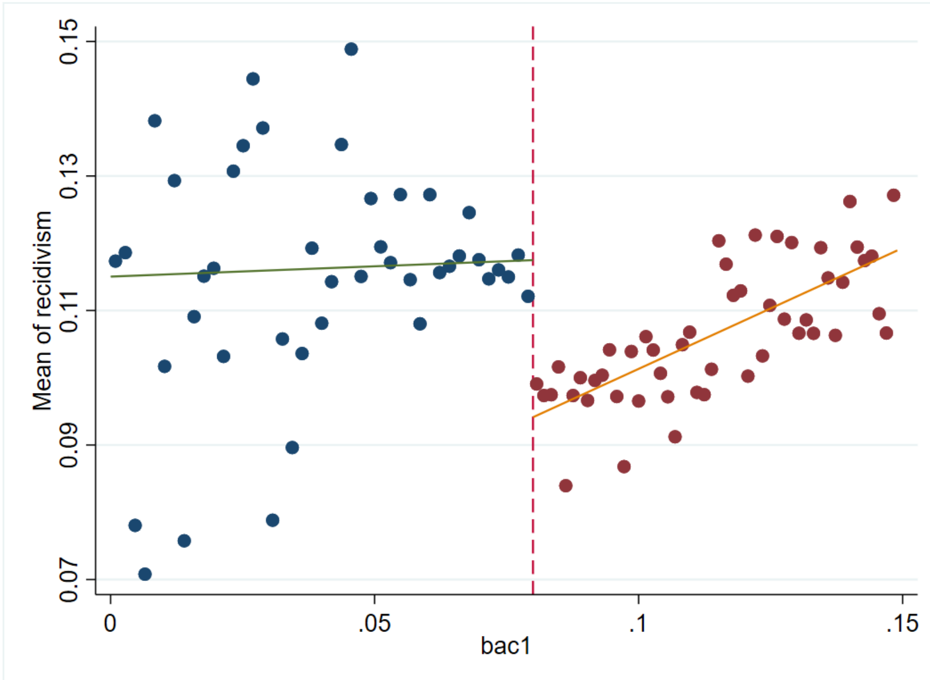
```
. quietly eststo: reg recidivism bac1 male white acc aged if bac1>=0.055 & bac1<=
0.105, robust

. quietly eststo: rdrobust recidivism bac1, c(0.0064) h(0.055 0.105) kernel(unifo
rm) p(1) covs(mal
> e white acc aged)

. quietly eststo: rdrobust recidivism bac1, c(0.0064) h(0.055 0.105) kernel(unifo
rm) q(2) covs(mal
> e white acc aged)

. esttab, keep(bac1 RD_Estimate)
```

	(1) recidivism	(2) recidivism	(3) recidivism
bac1	-0.476***		
	(-4.26)		
RD_Estimate		0.0428	0.0428
		(1.92)	(1.92)
N	46957	214558	214558

```
t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001
```

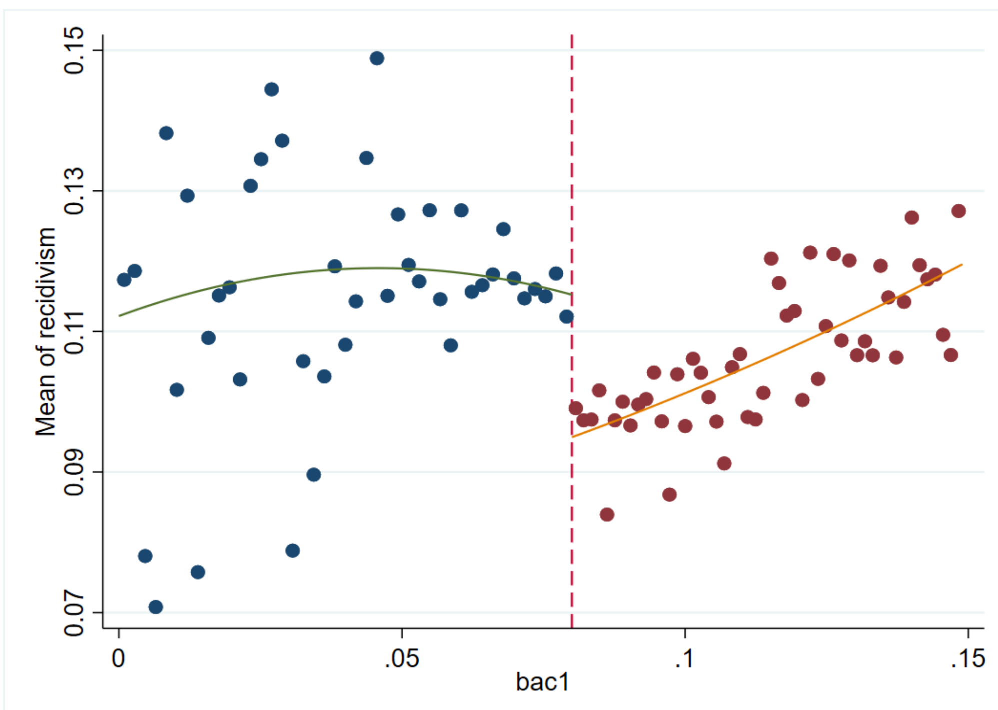
In question8, we will recreate the top panel of paper's Figure 3, which plots means of recidivism rates and predicted recidivism rates based on simple regression models for all offenders and highlights the stark changes in recidivism which occur at the 0.08 thresholds.

```
. quietly cmogram recidivism bac1 if bac1<0.15, cut(0.08) scatter line(0.08) lfit

. graph export bac11.png
(file bac11.png written in PNG format)
```



Panel A3. All offenders(linearly)

```
. quietly cmogram recidivism bac1 if bac1<0.15, cut(0.08) scatter line(0.08) qfit

. graph export bac12.png
(file bac12.png written in PNG format)
```



Panel A4. All offenders(quadratic)

The general trend is similar to each other. To be specific, there is a difference before the 0.08 cutoff. The slope of regression line has a smoothly increasing trend (almost parallel to the x-axis) while the slope in the article's graph has a smoothly decreasing trend. When BAC exceeds 0.08, both of the trends are increasing. The most significant similarity between our regression and author's regression is that there is a notable drop in recidivism at the 0.08 threshold. This indicates that the increase in punishments and sanctions at the threshold is effective in reducing future drunk driving, and having a BAC over the 0.08 legal limit is associated with lower recidivism rates.

In the last section, a conclusion should be needed. In our replication, we just focused on one of the thresholds, 0.08, to analyze problems. Among those, a hypothesis that the original BAC bins determined by the implicit rounding from the breathalyzer has been tested and the BAC distribution showed little evidence of endogenous sorting to one side of either of the thresholds studied. And later the p-values in McCrary test also verified that. From these a series of replications above, most of the results are close to those in Hansen's findings. For example, the fitted regression lines by using regression discontinuity design have very similar trend like the lines in Hansen's paper. And the regression results from regression discontinuity estimates for the effect of exceeding the 0.08 BAC threshold on recidivism are also similar to those in the paper. Thus, I think Hansen's findings are to some extent convincing, since he provided a classic project by using RDD method. There are also several differences from Hansen's results, I think this may be caused by the data which is not perfectly same as Hansen used, and also the operations in Stata by my replication cannot be same as Hansen did.