RDD Replication

Ashesh Shrestha

28 Feb 2021

*Question1

https://github.com/AsheshShrestha7/RDD

*Question 2

The main objective of this paper is to test the effect of sanctions and punishment for driving under influence (DUI). It seeks to investigate whether harsher punishment for driving under influence reduces recidivism or not. This paper uses the administrative record of 512,964 stops from the state of Washington where blood alcohol content (BAC) measured above 0.08 is DUI and BAC measured above 0.15 is aggravated DUI. A DUI entails fines, jail time and license suspensions, all of which are higher for aggravated DUI. Moreover, the statutory future penalty increases for each DUI received regardless of whether the previous offense was an ordinary DUI or aggravated DUI. Thus, the use of BAC threshold to determine punishment severity has allowed Hansen to use regression discontinuity design (RDD) to test whether the harsher punishments and sanctions offender experience at the BAC thresholds are effective in reducing drunk driving or not.

The key assumption of RDD is the assumption of smoothness. The underlying conditional regression should be continuous across the threshold. In other words, the drivers should not be able to manipulate which side of the threshold their BAC would fall which means, there should not be sorting on running variable. To test if there are signs of manipulation, Hansen has conducted density test and also checked for co-variate balance. With the p-values of 0.59 and 0.38 respectively at the 0.08 and 0.15 thresholds, there was no evidence of manipulation. Furthermore, while conducting a local linear polynomial regression with various covariates including key driver demographics like age, gender and race, as dependent variables, Hansen failed to reject the null that these characteristics are unrelated to the BAC cutoffs for DUI and aggravated DUI. Thus, he concluded that co-variates are balanced at cut-off and claimed this as a further evidence against sorting.

Based on regression discontinuity design he concludes that having a BAC level above the threshold of 0.08 decreases recidivism by 2 percentage point in four year follow-up period. Similarly, having a BAC above the aggravated DUI threshold reduces recidivism by 1.1 percentage point which translates to 8.9 percentage point, given the baseline recidivism rate.

*Question 3

```
. use hansen_dwi, clear

. generate DUI = 1 if (bac1>= 0.08)
```

```
(23,010 missing values generated)

. replace DUI= 0 if (bac1<0.08)
(23,010 real changes made)
```

*Question 4

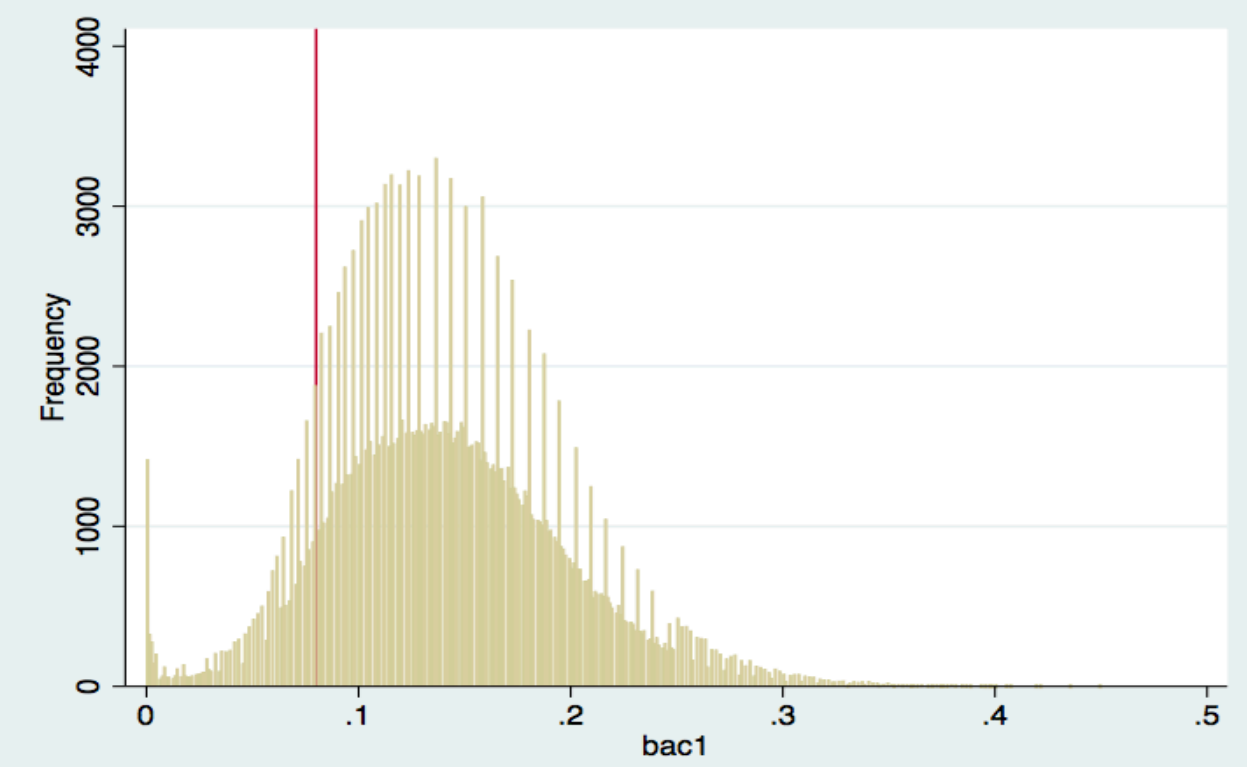
One of the tests we can conduct in order to check if people have manipulated their blood alcohol contest is McCrary density test. The following outlines the steps for implementing the test:

1. First create a histogram of the density of the running variable which is blood alcohol content (bac1) using a particular bin size such that no bins overlaps the cut point (BAC of 0.08).
2. Then, run two local linear regressions, one to the right and one the left of the cut-point.
3. Lastly, test whether log difference in height of the regression line just to the right and to the left of the cut-off is statistically different from zero.

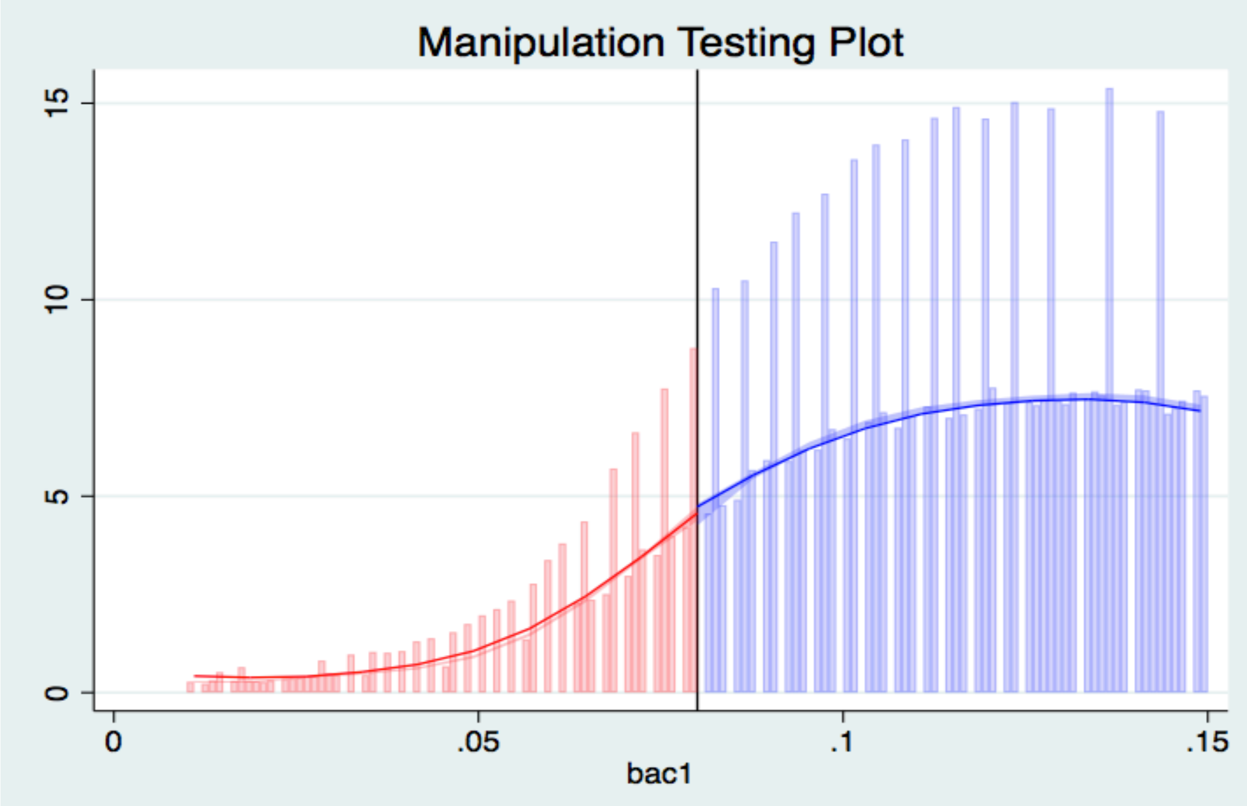
```
. hist bac1, freq width(0.001) xline(0.08)
(bin=450, start=0, width=.001)

. graph export histogram.png
(file /Users/user/Documents/untitled folder/histogram.png written in PNG format)

. rddensity bac1, c(0.08) hist_width(0.001) plot
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

```
. graph export densityplot.png
(file /Users/user/Documents/untitled folder/densityplot.png written in PNG
format)
```



4000
3000
2000
1000
0
Frequency
0
.1
.2
.3
.4
.5
bac1



Manipulation Testing Plot
15
10
5
0
0
.05
.1
.15
bac1

The distribution of bac1 offers little evidence of manipulation of blood alcohol content. The p-value of 0.88 signifies that log difference in height of the regression line just to the right and to the left of the threshold of 0.08 is not statistically difference from zero.

Hansen's result from the density test produces similar result to what I have found. He also finds no evidence of sorting on the running variable with p-value of the test 0.59 at the cut-off of 0.08.

*Question 5

```
. gen bac1_DUI = bac1*DUI

. eststo: quietly rdrobust white bac1, c(0.08) h(0.05) kernel(uniform)
(est1 stored)

. eststo: quietly rdrobust male bac1, c(0.08) h(0.05) kernel(uniform)
(est2 stored)

. eststo: quietly rdrobust aged bac1, c(0.08) h(0.05) kernel(uniform)
(est3 stored)

. eststo: quietly rdrobust acc bac1, c(0.08) h(0.05) kernel(uniform)
(est4 stored)

. esttab, se
```

	(1) white	(2) male	(3) aged	(4) acc
RD_Estimate	0.00570 (0.00501)	0.00618 (0.00570)	-0.140 (0.164)	-0.00335 (0.00406)
N	214558	214558	214558	214558

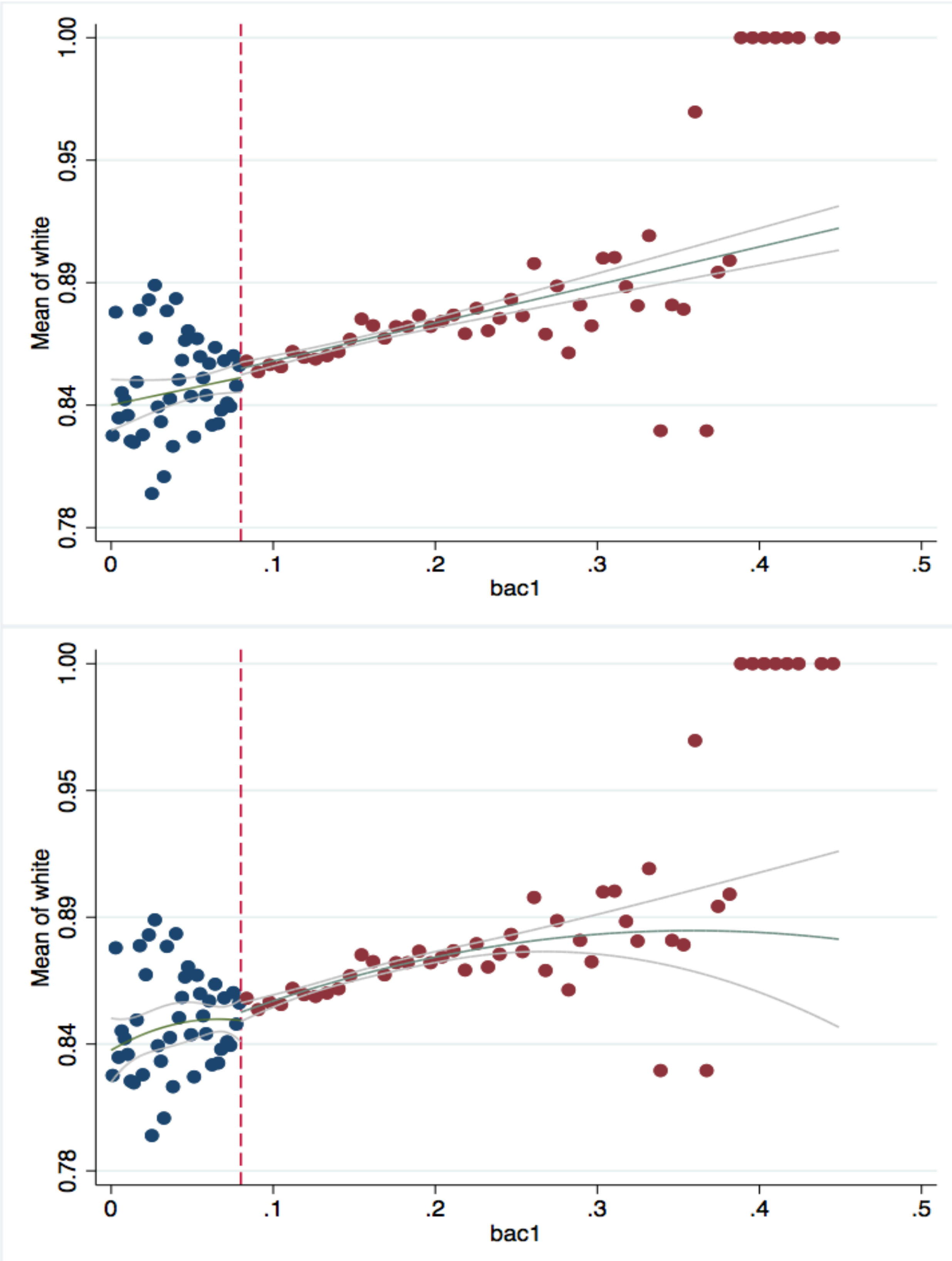
```
Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

. eststo clear
```

Since the RD estimates are insignificant, I fail to reject the null hypothesis that white, male, age and accident are unrelated to blood alcohol content (bac1) at cut-off of 0.08. Thus, I conclude that the covariates are balanced at the cut-off.

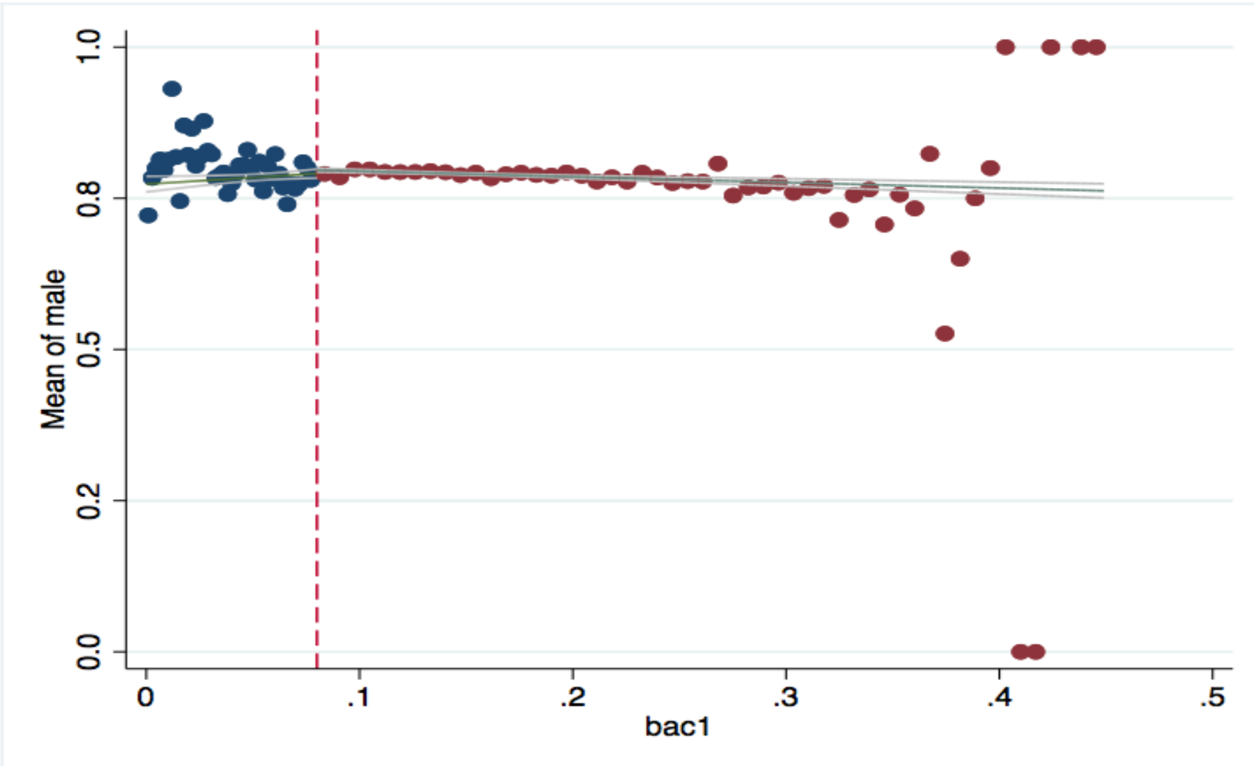
*Question 6

```
. quietly cmogram white bac1, cut(0.08) scatter line(0.08) lfitci

. graph export cmogram1.png
(file /Users/user/Documents/untitled folder/cmogram1.png written in PNG format)

. quietly cmogram white bac1, cut(0.08) scatter line(0.08) qfitci

. graph export cmogram2.png
```

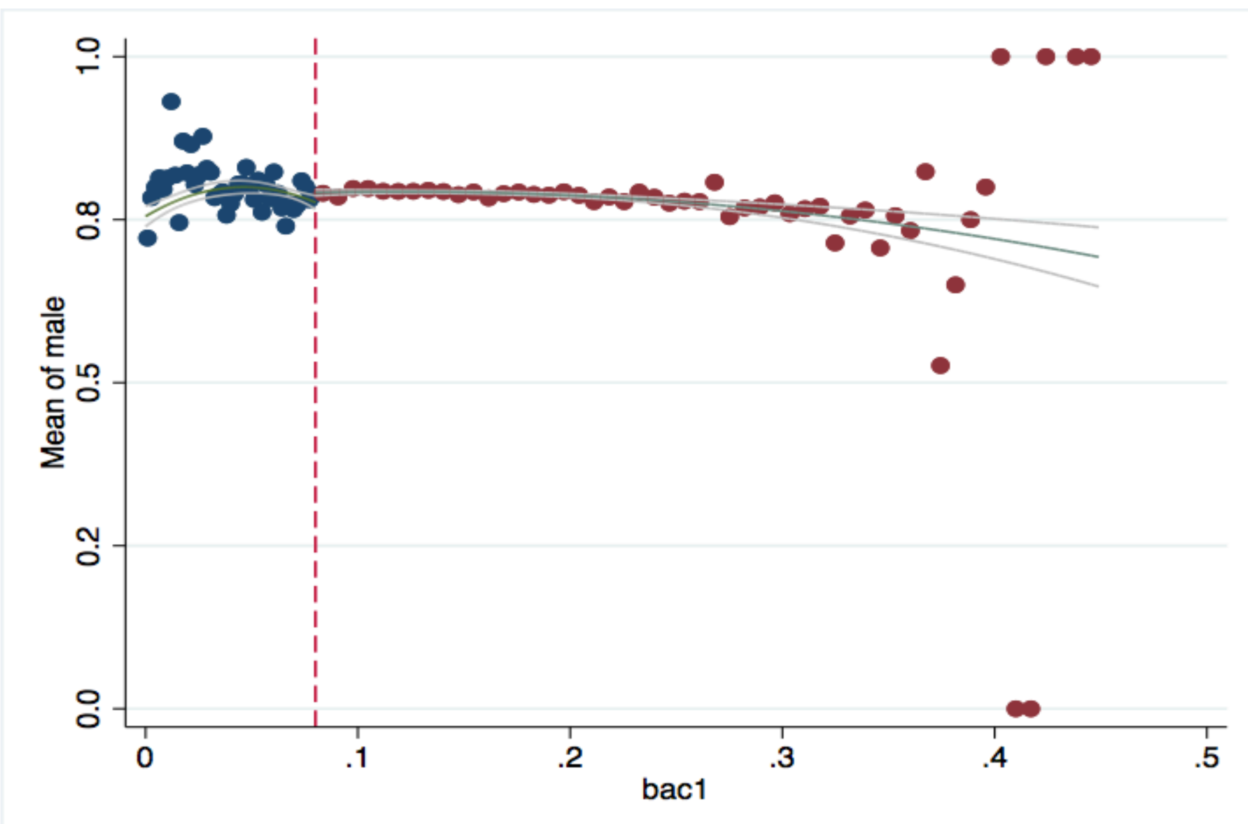
```
(file /Users/user/Documents/untitled folder/cmogram2.png written in PNG format)

. quietly cmogram male bac1, cut(0.08) scatter line(0.08) lfitci

. graph export cmogram3.png
(file /Users/user/Documents/untitled folder/cmogram3.png written in PNG format)

. quietly cmogram male bac1, cut(0.08) scatter line(0.08) qfitci

. graph export cmogram4.png
(file /Users/user/Documents/untitled folder/cmogram4.png written in PNG format)

. quietly cmogram aged bac1, cut(0.08) scatter line(0.08) lfitci

. graph export cmogram5.png
(file /Users/user/Documents/untitled folder/cmogram5.png written in PNG format)

. quietly cmogram aged bac1, cut(0.08) scatter line(0.08) qfitci

. graph export cmogram6.png
(file /Users/user/Documents/untitled folder/cmogram6.png written in PNG format)

. quietly cmogram acc bac1, cut(0.08) scatter line(0.08) lfitci

. graph export cmogram7.png
(file /Users/user/Documents/untitled folder/cmogram7.png written in PNG format)

. quietly cmogram acc bac1, cut(0.08) scatter line(0.08) qfitci

. graph export cmogram8.png
(file /Users/user/Documents/untitled folder/cmogram8.png written in PNG format)
```



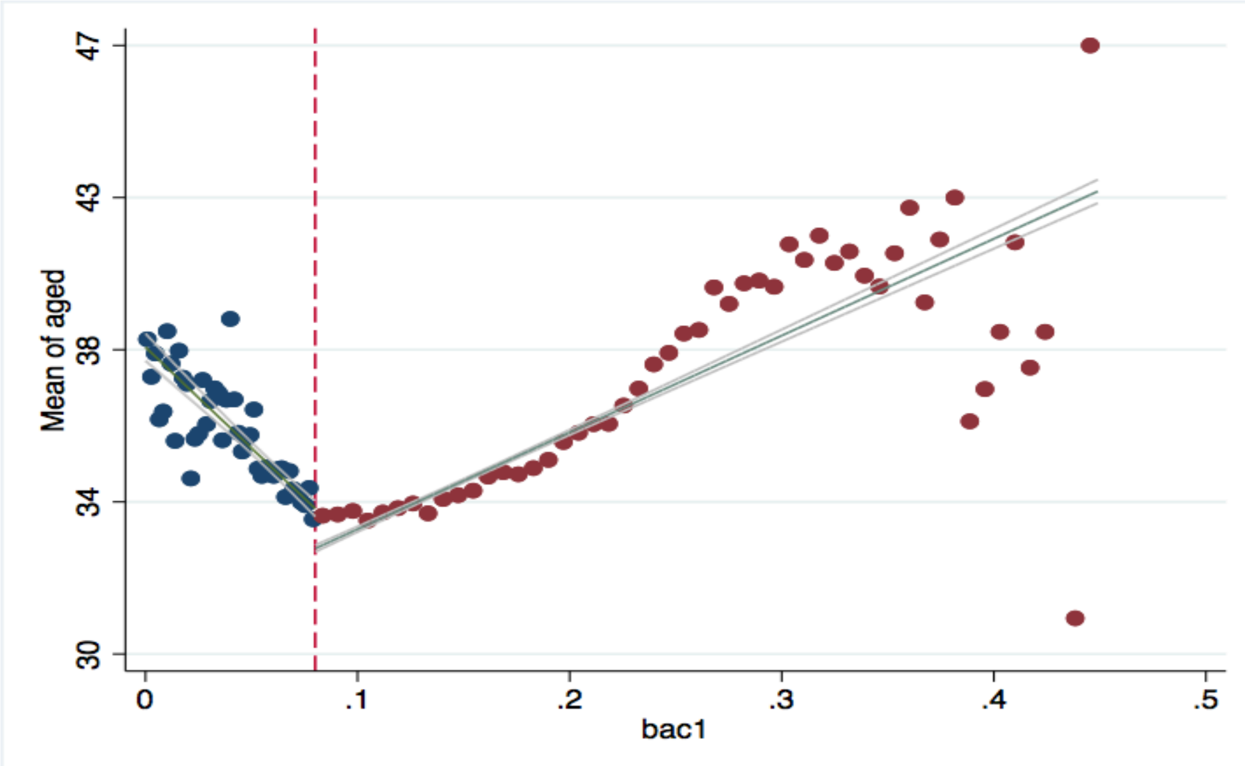
Mean of white
1.00
0.95
0.89
0.84
0.78
0
.1
.2
.3
.4
.5
bac1
Mean of white
1.00
0.95
0.89
0.84
0.78
0
.1
.2
.3
.4
.5
bac1



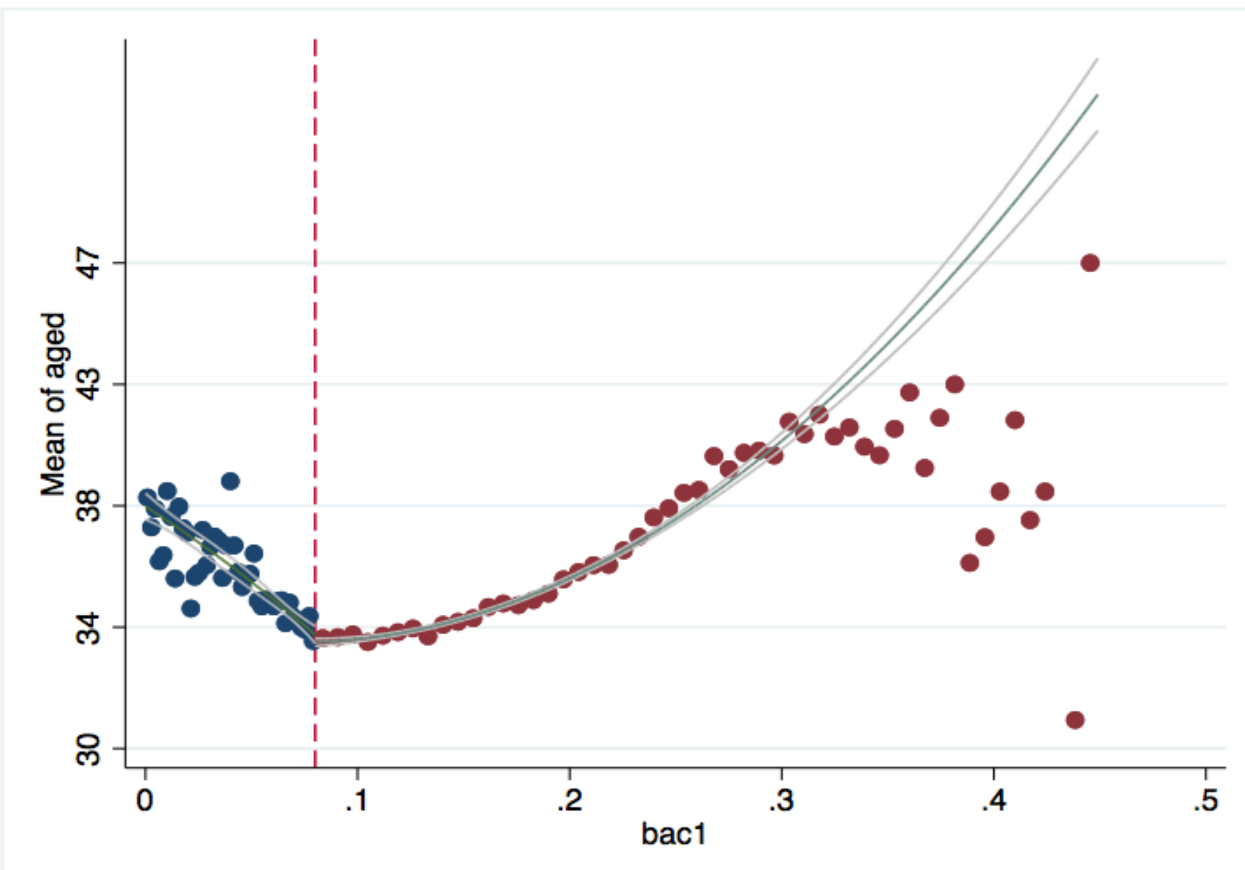
1.0
0.8
0.5
0.2
0.0
Mean of male
0
.1
.2
.3
.4
.5
bac1



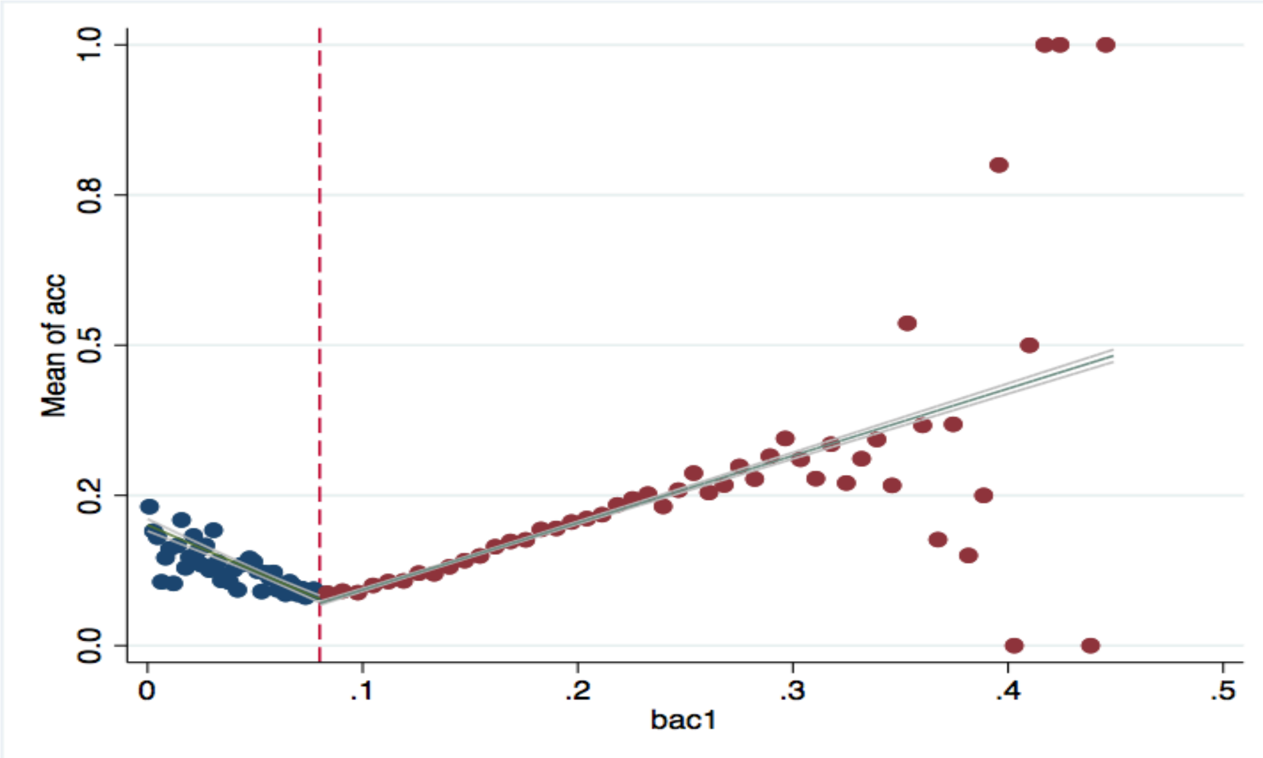
1.0
0.8
0.5
0.2
0.0
Mean of male
0
.1
.2
.3
.4
.5
bac1



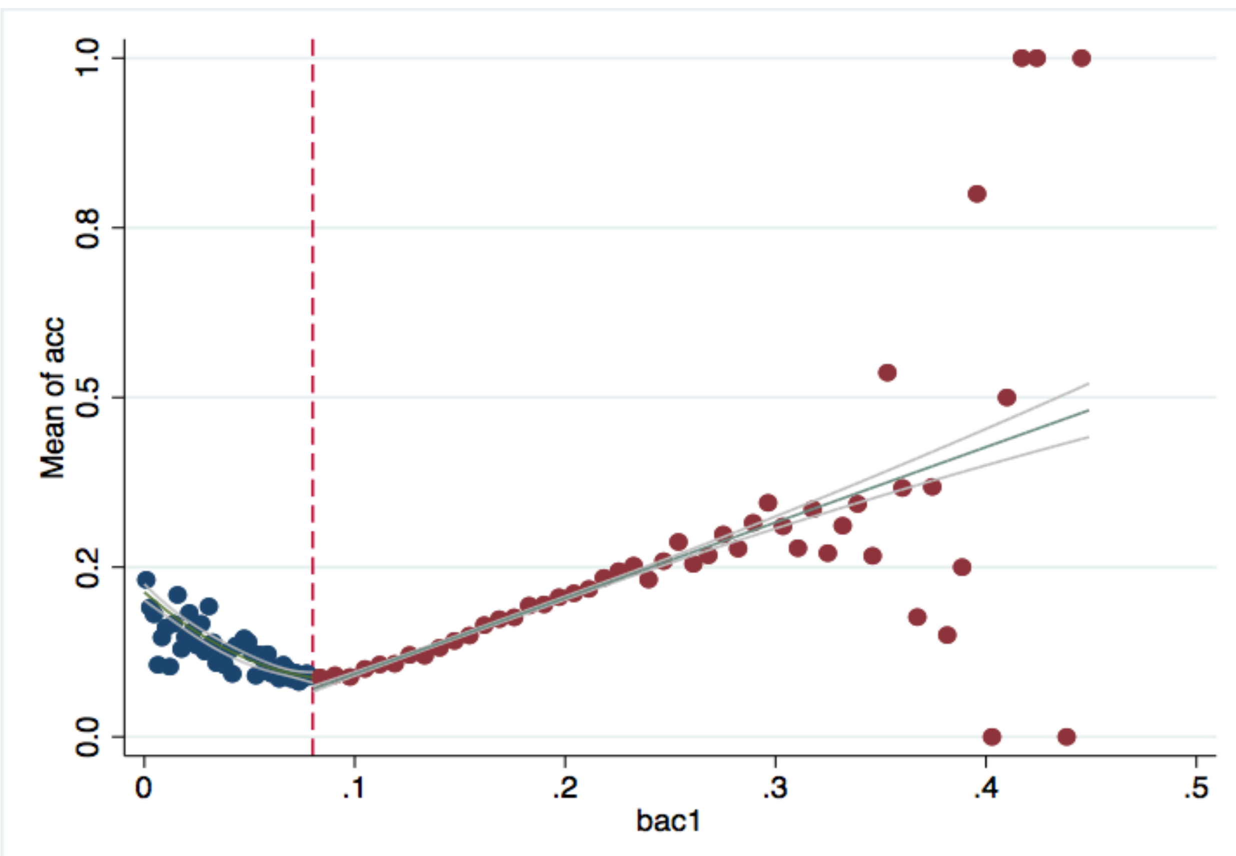
47
43
38
34
30
Mean of aged
0
.1
.2
.3
.4
.5
bac1



47
43
38
34
30
Mean of aged
0
.1
.2
.3
.4
.5
bac1



Mean of acc
1.0
0.8
0.5
0.2
0.0
0
.1
.2
.3
.4
.5
bac1



Mean of acc
1.0
0.8
0.5
0.2
0.0
0
.1
.2
.3
.4
.5
bac1

In the figures, we can see that white, male, age and accident are stable across the DUI punishment threshold of 0.08. There is no sudden significant jump in the regression line in either side of the cutoff. The confidence interval of the regression lines overlap right at the cutoff in at least one of the linear and quadratic models. In Hansen's paper, the linear regression lines of BAC on conditional mean of white, male, age and accident remain unchanged across the threshold signifying that there is no manipulation in either side of the 0.08 threshold. Thus, my results are similar to what Hansen has found in his paper.

*Question 7

```
. gen bac1_squared = bac1^2

. gen bac1_squared_DUI = bac1_squared*DUI

. eststo: quietly rdrobust recid bac1, c(0.08) h(0.05) covs(bac1) all
(est1 stored)

. eststo: quietly rdrobust recid bac1, c(0.08) h(0.05) covs(bac1_DUI) all
(est2 stored)

. eststo: quietly rdrobust recid bac1, c(0.08) h(0.05) covs(bac1_DUI
bac1_squared_DUI) all
(est3 stored)

. esttab, se
```

Panel A- Using bandwidth 0.03 to 0.13

	(1) recid	(2) recid	(3) recid
Conventional	-0.0199***	-0.0359***	-0.0677***
	(0.00471)	(0.00471)	(0.00471)
Bias-corre~d	-0.0164***	-0.0324***	-0.0665***
	(0.00471)	(0.00471)	(0.00471)
Robust	-0.0164*	-0.0324***	-0.0665***
	(0.00662)	(0.00662)	(0.00662)
N	214558	214558	214558

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001
Column 1: controlling for bac1 linearly
Column 2 : interacting bac1 with cutoff linearly
Column 3: interacting bac1 with cutoff linearly and as a quadratic

```
. eststo clear
```

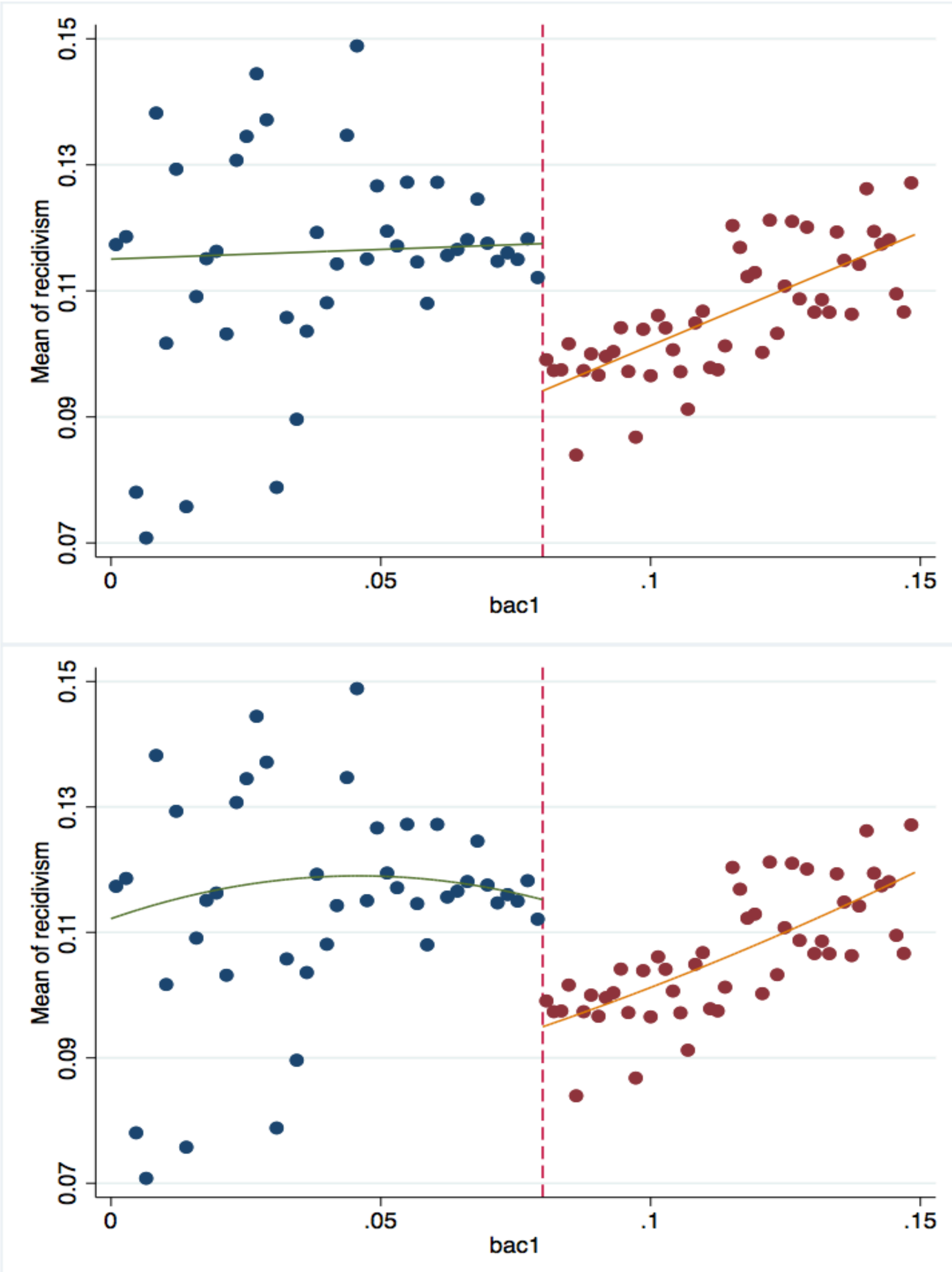
```
. eststo: quietly rdrobust recid bac1, c(0.08) h(0.025) covs(bac1) all
(est1 stored)

. eststo: quietly rdrobust recid bac1, c(0.08) h(0.025) covs(bac1_DUI) all
(est2 stored)

. eststo: quietly rdrobust recid bac1, c(0.08) h(0.025) covs(bac1_DUI
bac1_squared_DUI) all
(est3 stored)

. esttab, se

Using bandwidth 0.055 to 0.105
```

	(1) recid	(2) recid	(3) recid
Conventional	-0.0181**	0.0619***	-0.639***
	(0.00626)	(0.00626)	(0.00626)
Bias-corre~d	-0.0159*	0.0641***	-0.639***
	(0.00626)	(0.00626)	(0.00626)
Robust	-0.0159	0.0641***	-0.639***
	(0.00906)	(0.00906)	(0.00906)
N	214558	214558	214558

```
Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Column 1: controlling for bac1 linearly
Column 2: interacting bac1 with cutoff linearly
Column 3: interacting bac1 with cutoff linearly and as a quadratic

. eststo clear
```

*Question 8

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

. graph export cmogram9.png
(file /Users/user/Documents/untitled folder/cmogram9.png written in PNG format)

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

. graph export cmogram10.png
(file /Users/user/Documents/untitled folder/cmogram10.png written in PNG format)
```


Mean of recidivism
0.07
0.09
0.11
0.13
0.15
0
.05
.1
.15
bac1
Mean of recidivism
0.07
0.09
0.11
0.13
0.15
0
.05
.1
.15
bac1

*Question 9

This exercise in general has helped me to understand the concepts surrounding the Regression Discontinuity design. I have better understood the assumption of smoothness and the importance of this assumption in order to establish causal relationship using discontinuity design. Now, I understand and I am able to test if there is any evidence of manipulation at the cut-off. However, I still have some confusion regarding operation of RDD which I hope will get clearer with further readings and practice exercises.

The main hypothesis tested was whether harsher punishments and sanctions on driving under the influence, that is having BAC level higher than 0.08, reduces recidivism or not. As seen in the table for question number 7, in all three columns of both the panels(with bandwidth of 0.05 and 0.025), the coefficient of local polynomial discontinuity estimates are negative and statistically significant. Hence, I can conclude that having BAC level above 0.08 reduces recidivism.

I think Hansen has thoroughly followed the procedure of regression discontinuity design to come to the conclusion. Moreover, from the replication exercise, I also have come to similar conclusion. Hence, I am confident in Hansen's conclusion.