

Github repo and summary

1. Create a new github repository

[Xing's GitHub repositories](#)

2. Briefly summarize the paper.

The research focus on the effect of harsher punishments and sanctions on driving under the influence. Punishments are determined by strict rules on blood content and prevent offenses. The data comes from administrative records on 512,964 DUI BAC tests in the state of Washington from 1995 to 2011 and data from 1999-2007 to analyze the causal effect of having a BAC above either the 0.08 or 0.15 threshold on recidivism within four years of the original BAC test. In terms of “identification strategy”, the author assumes that it is locally random for a driver to have a BAC either just below or just above the BAC thresholds. Finally, the author concludes that regression discontinuity derived estimates suggest that having a BAC above the DUI threshold reduces recidivism by up to 2 percentage points (17 percent); having a BAC over the aggravated DUI threshold reduces recidivism by an additional percentage point (9 percent). Besides, the additional sanctions experienced by drunk drivers at BAC thresholds are effective in reducing repeat drunk driving.

Reproducing somewhat Hansen's results

3. Create a dummy

```
gen d = 0  
replace d = 1 if bac1 >= 0.08
```

4. Recreate Figure 1

```
histogram bac1, frequency width(0.0001) xline(0.08, lc(red)) xtitle("BAC")
```

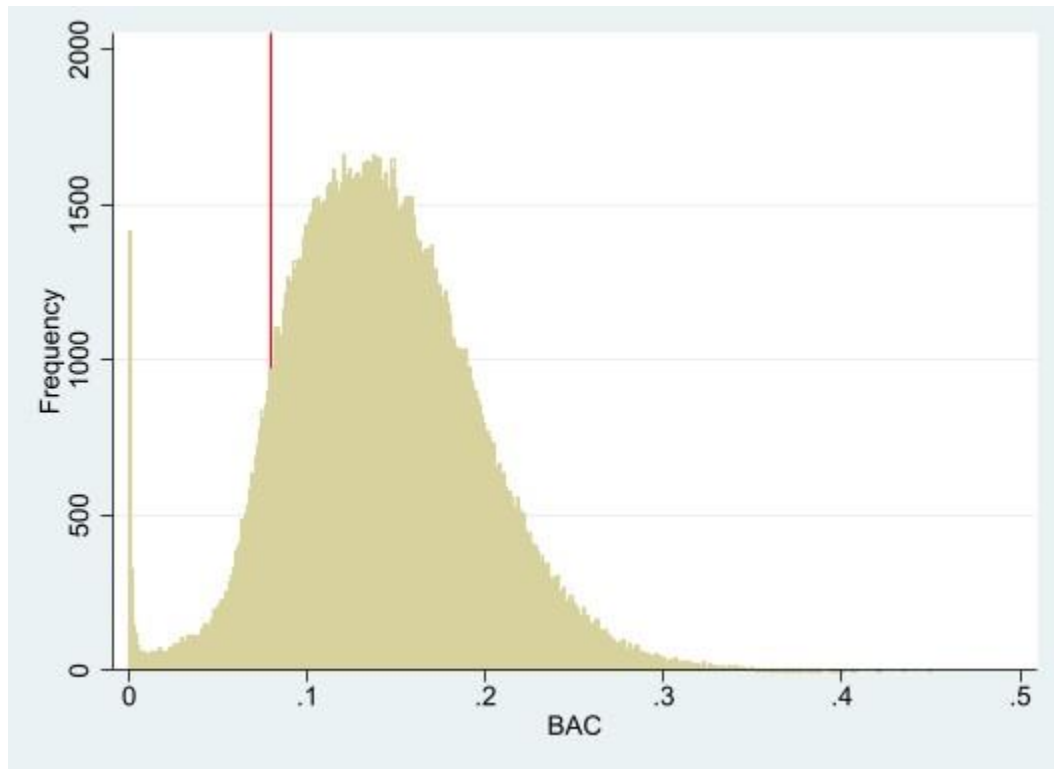


Figure 1

I didn't find evidence for sorting on the running variable since the distribution of BAC shows little evidence of endogenous sorting to one side of either of the thresholds studied.

5. Table 2 Panel A

reg d white male aged acc, robust

Linear regression	Number of obs	=	214,558
	F(4, 214553)	=	93.40
	Prob > F	=	0.0000
	R-squared	=	0.0016
	Root MSE	=	.30918

d	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
white	.0171179	.0020259	8.45	0.000	.0131472	.0210886
male	.0036559	.0016481	2.22	0.027	.0004256	.0068862
aged	-.0004864	.0000611	-7.97	0.000	-.0006061	-.0003667
acc	.0277418	.0017297	16.04	0.000	.0243517	.031132
_cons	.8880377	.0030581	290.39	0.000	.8820439	.8940314

Panel A focuses on the estimated effect of BAC above the DUI threshold. They are balanced and we fail to reject the null that the predetermined characteristics are unrelated to the BAC cutoffs for DUI.

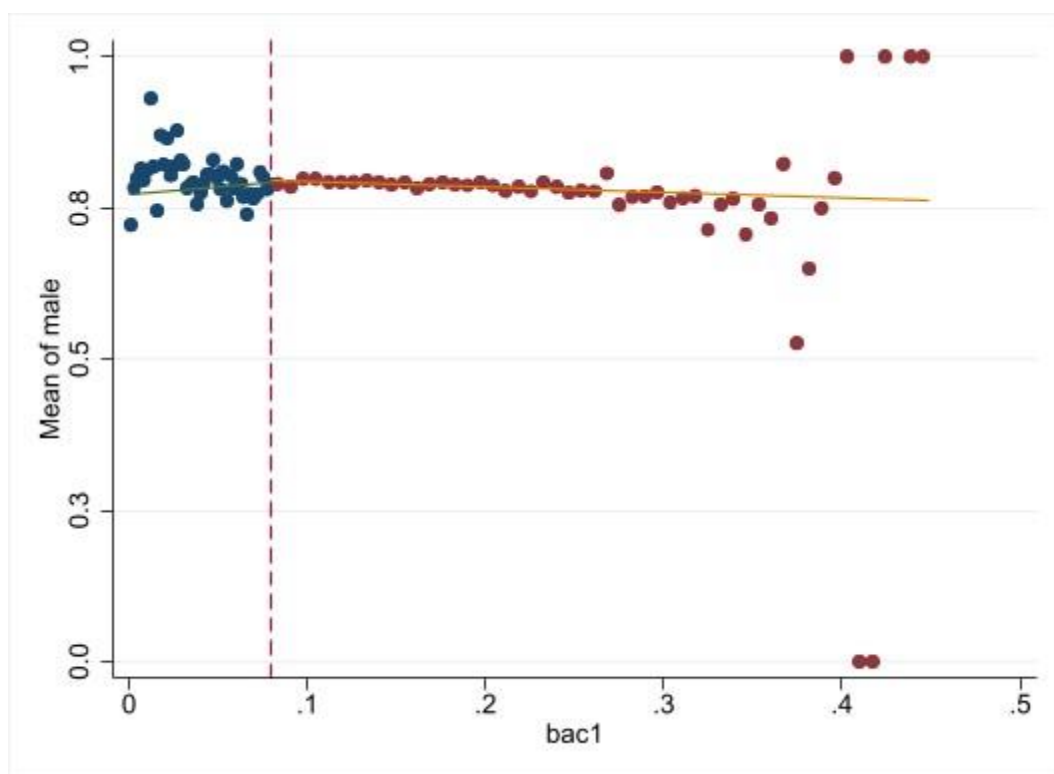
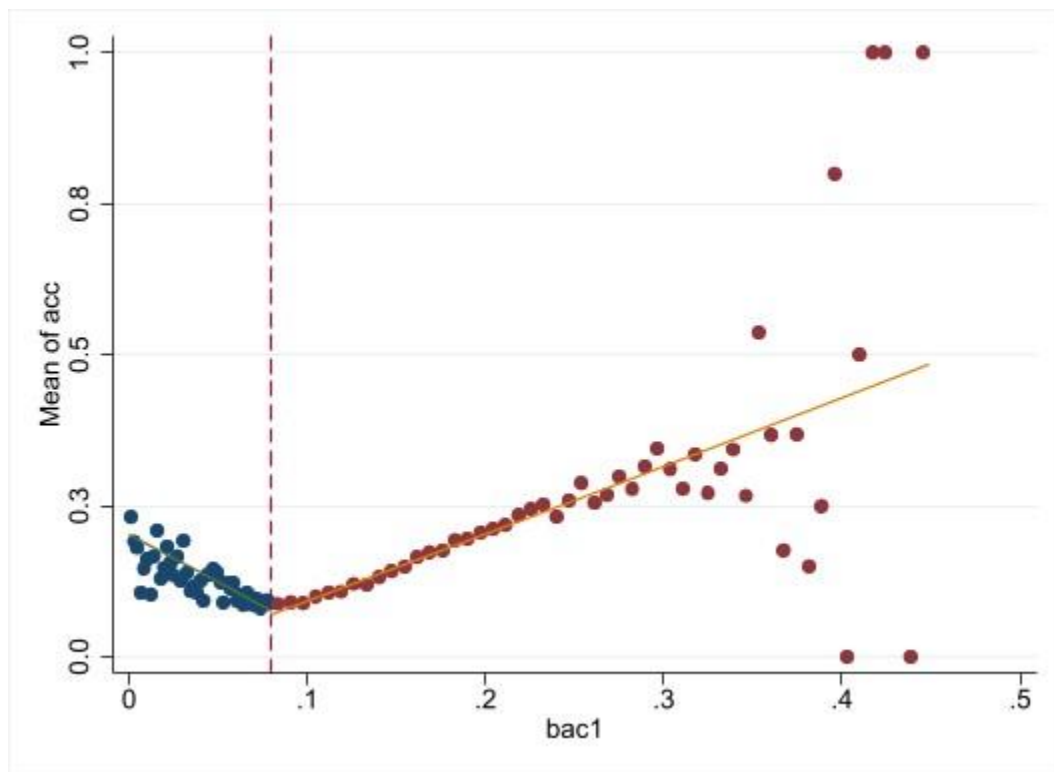
6. Figure 2 panel A-D

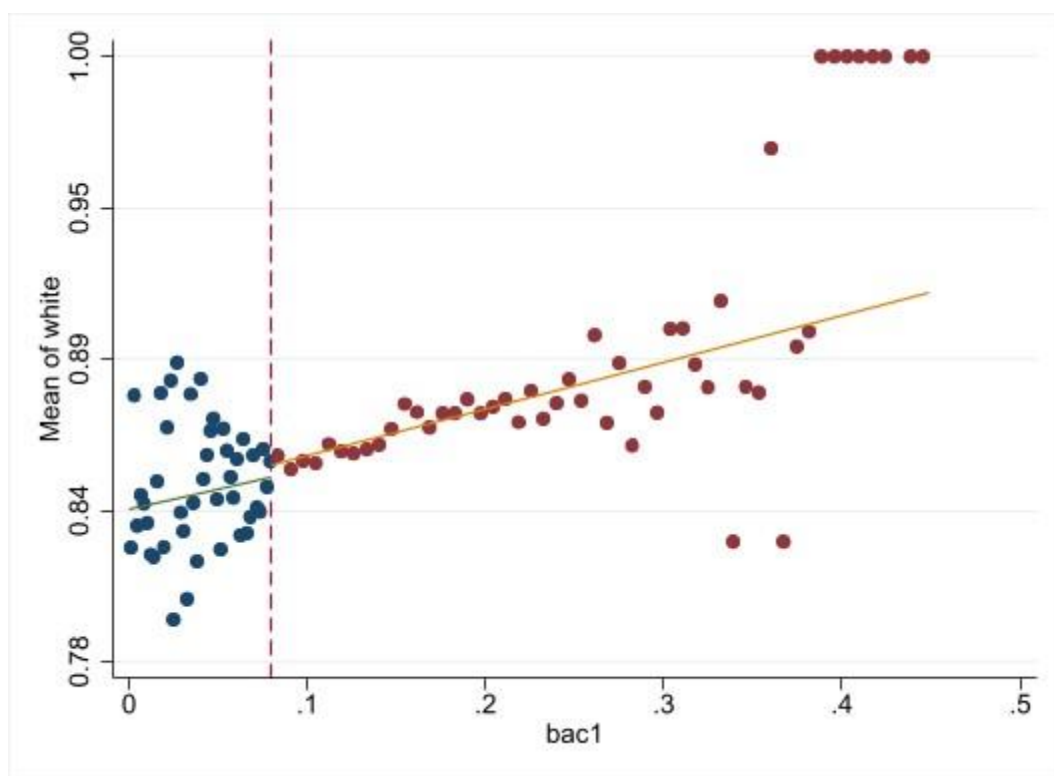
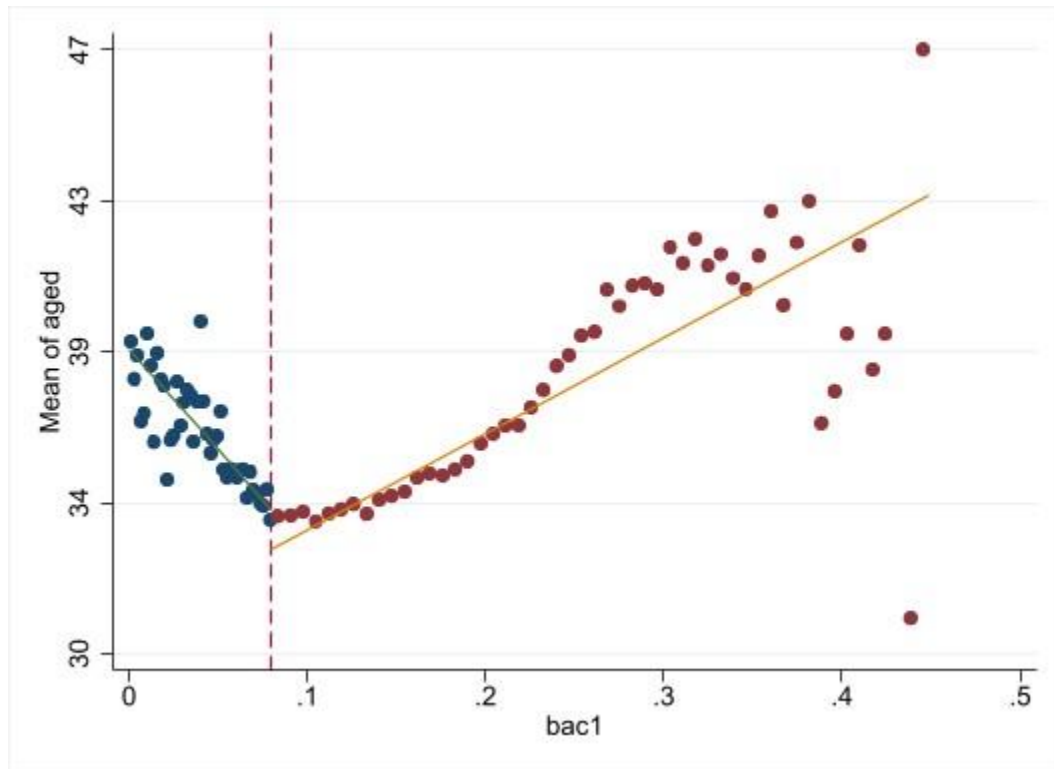
```
cmogram acc bac1, cut(0.08) scatter line(0.08) lfit  
cmogram acc bac1, cut(0.08) scatter line(0.08) qfit
```

```
cmogram male bac1, cut(0.08) scatter line(0.08) lfit  
cmogram male bac1, cut(0.08) scatter line(0.08) qfit
```

```
cmogram aged bac1, cut(0.08) scatter line(0.08) lfit  
cmogram aged bac1, cut(0.08) scatter line(0.08) qfit
```

```
cmogram white bac1, cut(0.08) scatter line(0.08) lfit  
cmogram white bac1, cut(0.08) scatter line(0.08) qfit
```





Demographic factors such as age, race (defined by white versus non-white), and gender are stable across the DUI punishment thresholds. The stability of predetermined characteristics gives additional credibility that the regression discontinuity design can deliver unbiased estimates in this scenario as it suggests

neither the impaired driver nor the police officer is able to manipulate testing on either side of the 0.08 thresholds.

7. Table 3

```
gen bac1_c = bac1 - 0.08
gen bac1_c2 = bac1_c^2
```

```
** panel A
gen bac1_c = bac1 - 0.08
gen bac1_c2 = bac1_c^2
```

```
* column 1
xi: reg recidivism bac1 acc male aged white if bac1 >= 0.03 & bac1 <= 0.13, robust
estimates store linearly, title(model 1)
```

```
* column 2
xi: reg recidivism i.d * bac1_c acc male aged white if bac1 >= 0.03 & bac1 <= 0.13, robust
estimates store lcoeff, title(model 2)
```

```
* column 3
xi: reg recidivism d##(c.bac1_c c.bac1_c2) acc male aged white if bac1 >= 0.03 & bac1 <= 0.13, robust
estimates store lcoeffquad, title(model 3)
```

```
estout linearly lcoeff lcoeffquad
```

```
** panel B
```

```
* column 1
xi: reg recidivism bac1 acc male aged white if bac1 >= 0.055 & bac1 <= 0.105, robust
estimates store linearly, title(model 1)
```

```
* column 2
xi: reg recidivism i.d * bac1_c acc male aged white if bac1 >= 0.055 & bac1 <= 0.105, robust
estimates store lcoeff, title(model 2)
```

```
* column 3
xi: reg recidivism d##(c.bac1_c c.bac1_c2) acc male aged white if bac1 >= 0.055 & bac1 <= 0.105, robust
estimates store lcoeffquad, title(model 3)
```

```
estout linearly lcoeff lcoeffquad
```

Linear regression

Number of obs = 89,967
 F(5, 89961) = 61.93
 Prob > F = 0.0000
 R-squared = 0.0031
 Root MSE = .30857

recidivism	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
bac1	-.0754879	.0482512	-1.56	0.118	-.1700597	.0190839
acc	.0047692	.0034524	1.38	0.167	-.0019975	.0115358
male	.0331526	.0023292	14.23	0.000	.0285874	.0377178
aged	-.0008381	.0000849	-9.87	0.000	-.0010045	-.0006716
white	.0161118	.0028028	5.75	0.000	.0106183	.0216052
_cons	.1025354	.0062892	16.30	0.000	.0902086	.1148622

bac1	-.0754879	0	
acc	.0047692	4.16e-19	.0041839
male	.0331526	1.06e-19	.0332123
aged	-.0008381	-1.70e-19	-.0008538
white	.0161118	2.18e-18	.0162247
_Id_1		-1.08e-17	
Date		9.45e-20	
Alcohol1		2.43e-18	
Alcohol2		2.33e-18	
low_score		-4.58e-18	
recidivism		1	
year		-3.44e-17	
bac2		0	
d		0	
bac1_c		0	-1.052646
bac1_c2		-2.17e-15	-24.71688
bac1_c1		0	
_est_lcoff~d		0	
_est_lcoff		0	
_est_linea~y		0	
_Id_1		0	
bac1_c		0	
acc		0	
male		0	
aged		0	
white		0	
0.d			0
1.d			-.0143237
0.d#c.bac1_c			0
1.d#c.bac1_c			1.026786
0.d#c.bac1~2			0
1.d#c.bac1~2			32.73074
_cons	.1025354	6.75e-14	.1001477

Linear regression

Number of obs = 89,967
 F(9, 89957) = 40.23
 Prob > F = 0.0000
 R-squared = 0.0037
 Root MSE = .30848

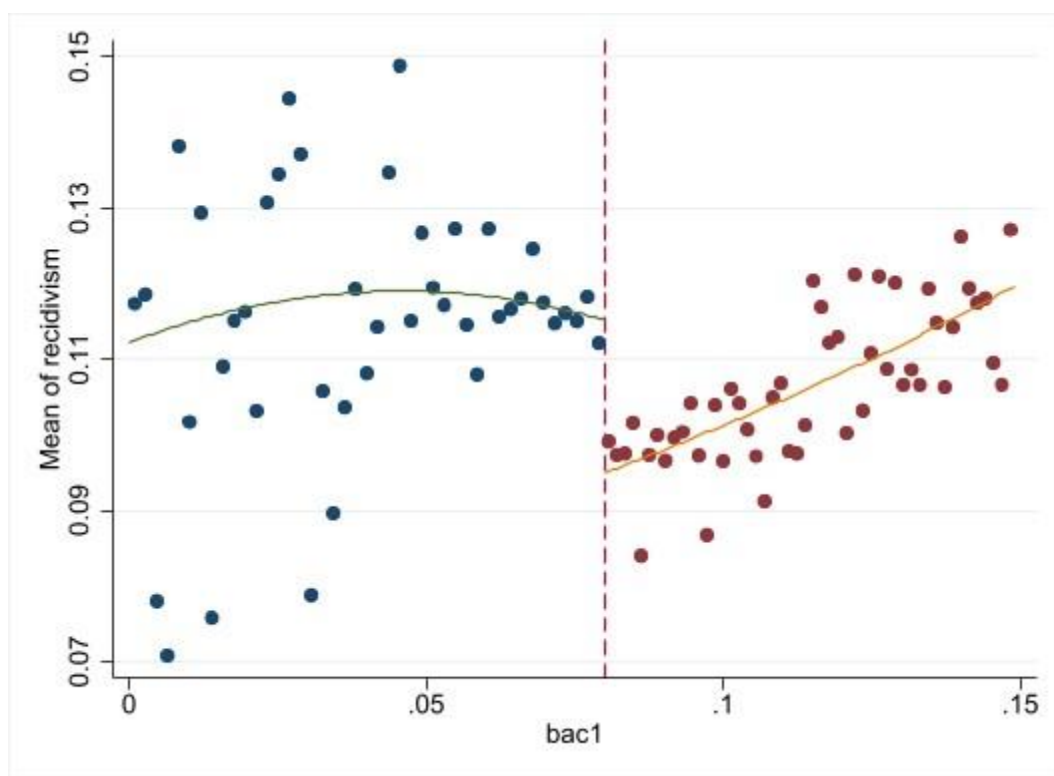
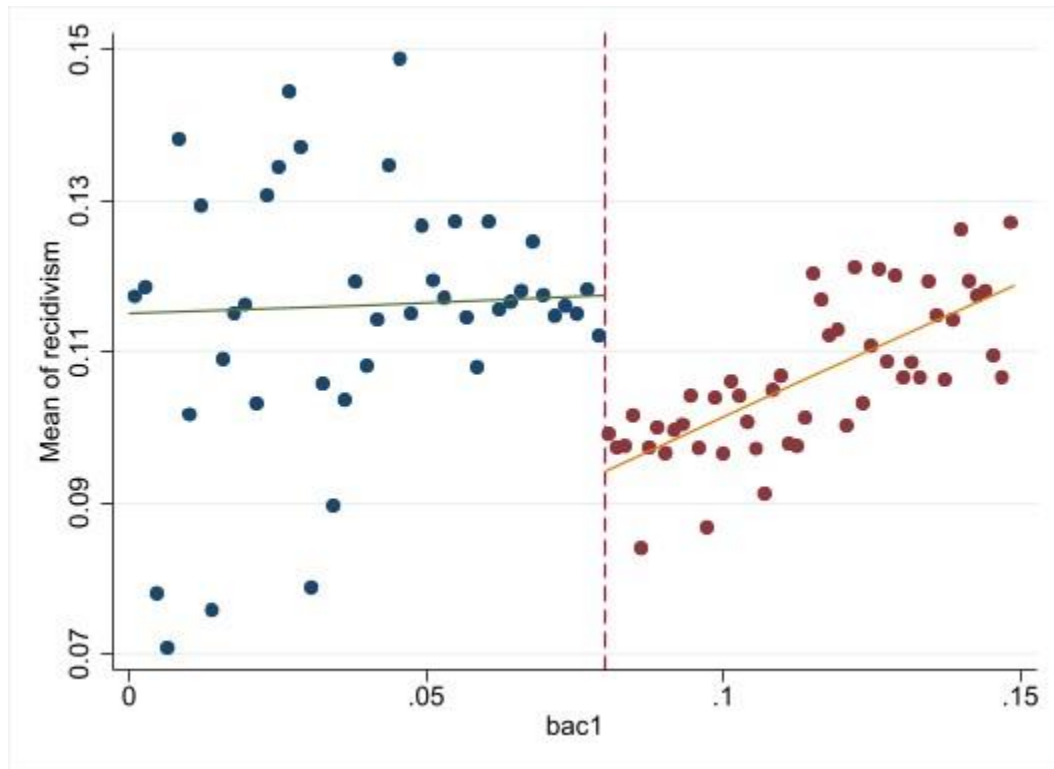
recidivism	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
1.d	-.0143237	.006222	-2.30	0.021	-.0265188	-.0021286
bac1_c	-1.052646	.6012363	-1.75	0.080	-2.231064	.1257711
bac1_c2	-24.71688	13.73897	-1.80	0.072	-51.64513	2.211376
d#c.bac1_c						
1	1.026786	.6890208	1.49	0.136	-.3236884	2.37726
d#c.bac1_c2						
1	32.73074	15.10452	2.17	0.030	3.126029	62.33545
acc	.0041839	.0034514	1.21	0.225	-.0025807	.0109485
male	.0332123	.0023288	14.26	0.000	.028648	.0377767
aged	-.0008538	.000085	-10.05	0.000	-.0010203	-.0006872
white	.0162247	.0028024	5.79	0.000	.0107321	.0217173
_cons	.1001477	.0062711	15.97	0.000	.0878564	.1124389

recidivism	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
_Id_1	-1.08e-17	4.49e-19	-24.13	0.000	-1.17e-17	-9.95e-18
Date	9.45e-20	9.02e-22	104.80	0.000	9.27e-20	9.63e-20
Alcohol1	2.43e-18	3.69e-20	65.89	0.000	2.36e-18	2.50e-18
Alcohol2	2.33e-18	4.10e-20	56.74	0.000	2.25e-18	2.41e-18
low_score	-4.58e-18	6.60e-20	-69.40	0.000	-4.71e-18	-4.45e-18
male	1.06e-19	2.12e-19	0.50	0.616	-3.09e-19	5.22e-19
white	2.18e-18	2.52e-19	8.66	0.000	1.69e-18	2.68e-18
recidivism	1	7.69e-19	1.3e+18	0.000	1	1
acc	4.16e-19	3.03e-19	1.37	0.170	-1.79e-19	1.01e-18
aged	-1.70e-19	7.72e-21	-22.03	0.000	-1.85e-19	-1.55e-19
year	-3.44e-17	3.31e-19	-103.91	0.000	-3.51e-17	-3.38e-17
bac1	0	(omitted)				
bac2	0	(omitted)				
d	0	(omitted)				
bac1_c	0	(omitted)				
bac1_c2	-2.17e-15	1.89e-16	-11.50	0.000	-2.54e-15	-1.80e-15
bac1_c1	0	(omitted)				
_est_linearly	0	(omitted)				
_est_lcoffquad	0	(omitted)				
_est_lcoff	0	(omitted)				
_Id_1	0	(omitted)				
bac1_c	0	(omitted)				
acc	0	(omitted)				
male	0	(omitted)				
aged	0	(omitted)				
white	0	(omitted)				
_cons	6.75e-14	6.50e-16	103.85	0.000	6.62e-14	6.87e-14

8. Recreate the top panel of Figure 3

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

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



9. Discuss what you learned from this exercise

Back to the research itself, I believed the author's conclusion makes sense but I couldn't get the same result due to the data difference or just my code. For example, I found little evidence correspond with the conclusion for question 7.

Thanks to the assignment, I was trying many new stuff for the first time, such as inserting a picture with Markdown, replicating the result from famous economists.

Practice makes perfect. I have to say it really depresses me when I was unable to run the code successfully. However, the happiness after doing this work makes me more confident and I better understand how RDD work thorough actual economic problem.

I realized that why our professor want us to connect knowledge of theory with estimation, to be a more empirical one.

c