Causal Inference-Replication of Hansen (2015)

Wen-Hsin (Molly) Chang

Date: 2021.2.15

1. Create new Github repo named " RDD"

   https://github.com/wsmolly/RDD

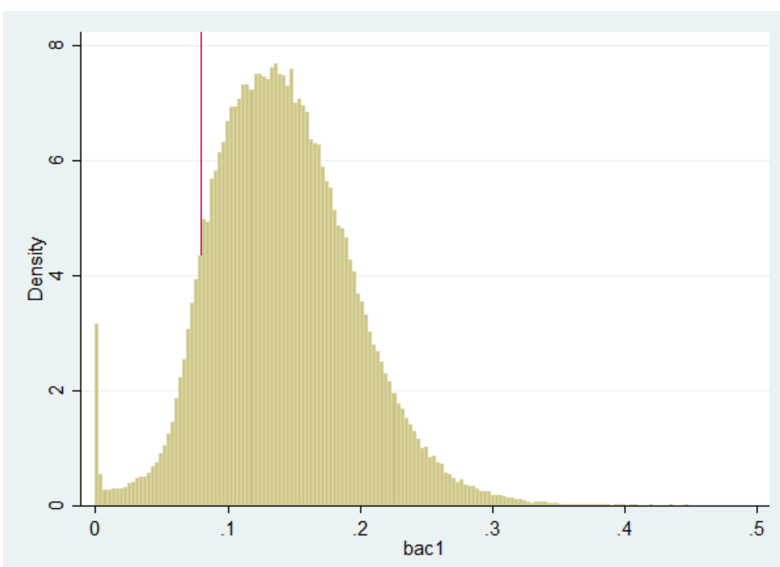
2. Summary of the paper

   The research question of this paper is whether the punishment severity has on commission of future crime. The author uses the data from the administrative records of drinking under influence (DUI) stops from the states of Washington from 1995 to 2011. The research design mainly follows the regression discontinuity approach, and the author relies upon the reasoning that drivers and officers are unable to manipulate the blood alcohol content (BAC). In conclusion, the author suggests that harsher punishment associated with BAC limits reduce future drunk driving.

3. Create a dummy equaling 1 if bac>=0.08 and 0 otherwise.

```
. use https://github.com/scunning1975/causal-inference-class/raw/master/
  hansen_dwi, clear
. gen DUI=1 if bac1>=0.08
. replace DUI=0 if DUI==.
```

4. Replication of Figure 1: sorting the running variable



There seems to be no evidence for sorting on the running variable. Specifically, the distribution of bac1 is smooth around the 0.08 threshold. The result gives credence to Hansen's regression discontinuity design.

5. Replication of Table2 :checking for covariate balance

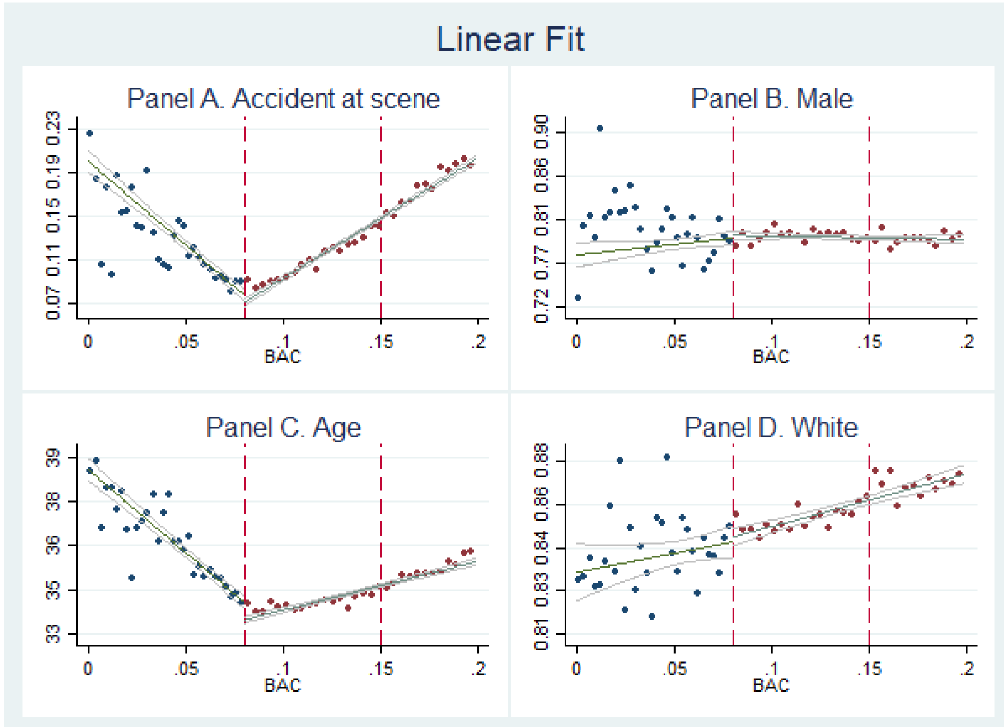
	(1)	(2)	(3)	(4)
VARIABLES	Male	White	Age	Accident
bac1	-0.210	0.0788	-69.16***	-1.096***
	(0.234)	(0.224)	(6.862)	(0.176)
DUI	-0.0184	0.00445	-6.224***	-0.154***
	(0.0198)	(0.0173)	(0.585)	(0.0158)
DUI_bac1	0.307	0.0156	76.05***	1.888***
	(0.263)	(0.238)	(7.647)	(0.203)
Constant	0.801***	0.840***	39.45***	0.171***
	(0.0151)	(0.0150)	(0.466)	(0.0118)
Observations	89,967	89,967	89,967	89,967
R-squared	0.000	0.000	0.002	0.002

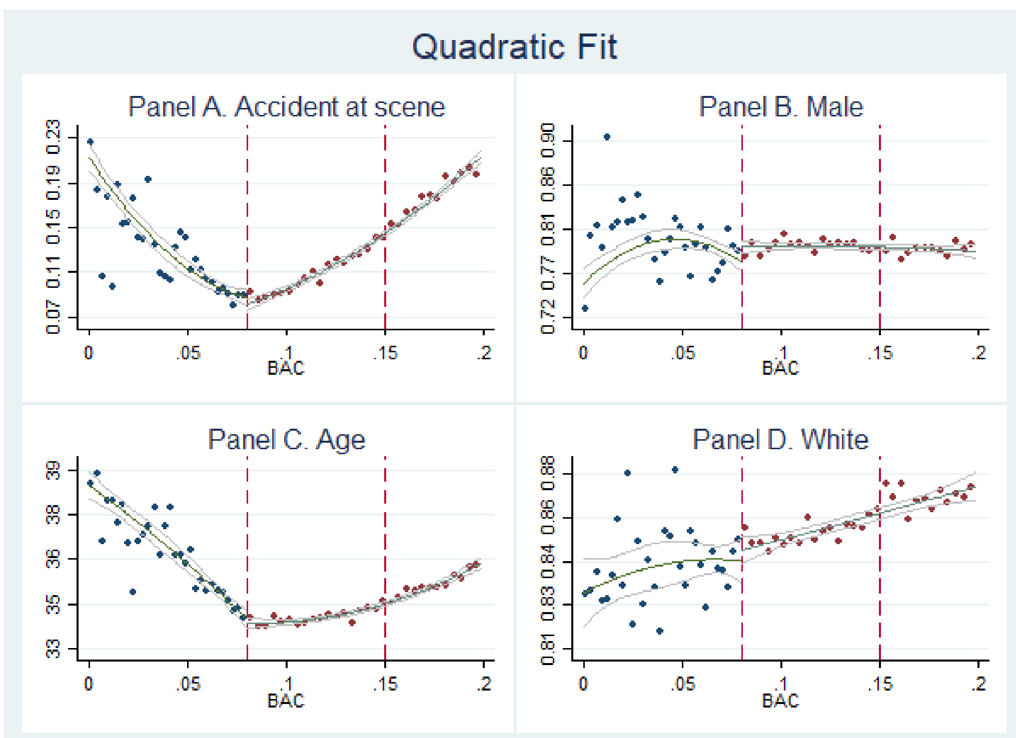
Robust standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

We can see that the coefficient of DUI is significant for Age and Accident and insignificant for Male and White. Therefore, covariate balance seems to hold under Male and White, while it does not hold for Age and Accident. The result is different from Hansen's original result. I suspect one of the reasons may be the difference between our data set and Hansen's original one.

6. Replication of Figure2 : BAC and characteristic





Although two of the variables violate covariate balance in our replication of table 2, the replication results of Figure 2 above seem quite similar to Hensen's paper. Hensen does not use quadratic fits in his paper, but I suspect the quadratic model can serve a betterpurpose in projection.

7. Table3 Panel A:recidivism outcome (0.03 < bac1 < 0.13)

.

Panel A (bandwidth0.03-0.13) VARIABLES	(1)	(2)	(3)
bac1	0.312***	-0.0367	2.943*
	(0.0748)	(0.187)	(1.638)
DUI	-0.0271***	-0.0576***	0.116
	(0.00403)	(0.0152)	(0.0843)
DUI_bac1		0.420**	-4.276**
		(0.204)	(2.112)
bac1sq			-25.01*
			(13.74)
DUI_bac1sq			33.09**
			(15.11)
white	0.0145***	0.0145***	0.0145***
	(0.00282)	(0.00282)	(0.00282)
male	0.0326***	0.0325***	0.0326***
	(0.00233)	(0.00233)	(0.00233)
aged	-0.000854***	-0.000860***	-0.000860***
	(8.49e-05)	(8.50e-05)	(8.50e-05)
acc	0.00456	0.00434	0.00431
	(0.00345)	(0.00345)	(0.00345)
year	-0.00268***	-0.00267***	-0.00267***
	(0.000411)	(0.000411)	(0.000411)
Constant	5.456***	5.455***	5.378***
	(0.823)	(0.823)	(0.824)

Observations	89,967	89,967	89,967
R-squared	0.004	0.004	0.004

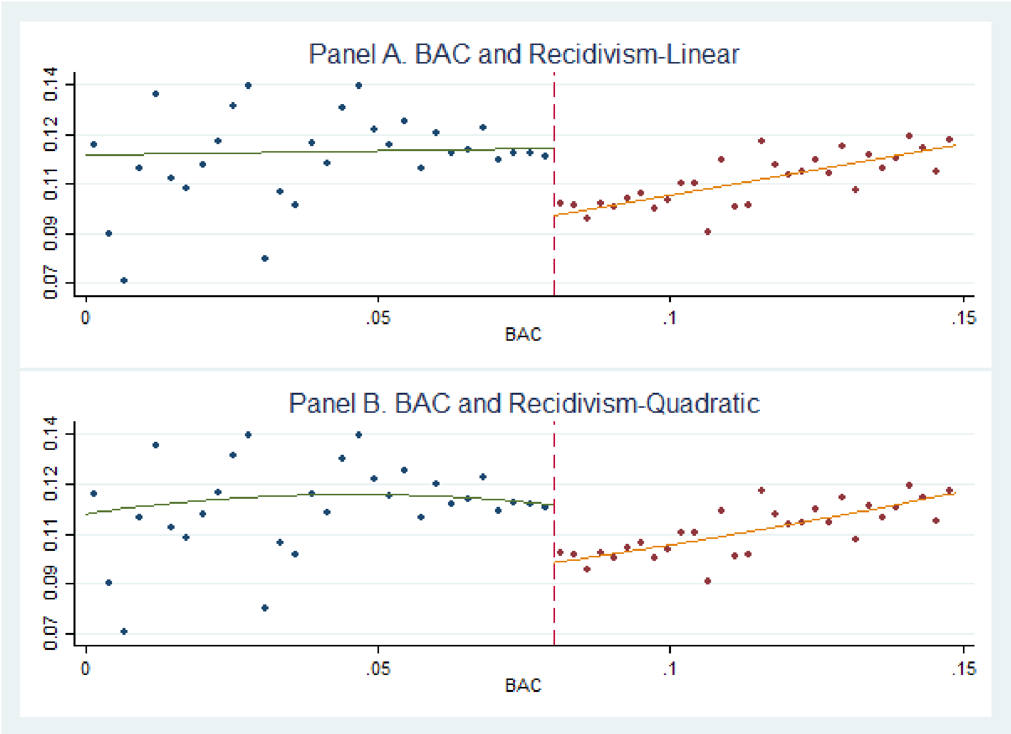
Panel B (bandwidth0.055-0.105) VARIABLES	(1)	(2)	(3)
bac1	0.177	-0.194	6.124
	(0.201)	(0.382)	(8.113)
DUI	-0.0217***	-0.0628*	0.379
	(0.00558)	(0.0350)	(0.422)
DUI_bac1		0.530	-10.66
		(0.449)	(10.60)
bac1sq			-45.74
			(58.70)
DUI_bac1sq			71.86
			(69.16)
white	0.0156***	0.0156***	0.0156***
	(0.00383)	(0.00383)	(0.00383)
male	0.0349***	0.0349***	0.0349***
	(0.00317)	(0.00317)	(0.00317)
aged	-0.000761***	-0.000763***	-0.000764***
	(0.000115)	(0.000115)	(0.000115)
acc	0.00416	0.00408	0.00409
	(0.00496)	(0.00496)	(0.00496)
year	-0.00299***	-0.00299***	-0.00299***
	(0.000567)	(0.000567)	(0.000567)
Constant	6.087***	6.103***	5.890***
	(1.136)	(1.137)	(1.172)

Observations	46,957	46,957	46,957
R-squared	0.005	0.005	0.005

In Panel A, the coefficient of DUI is significantly negative in column 1 and 2, indicating that there is a decrease in recidivism due to the punishment of exceeding the 0.08 BAC threshold. However, it is interesting that the coefficient of DUI becomes insignificant when including quadratic interaction. A further examination will be needed to determine whether it is appropriate to model the punishment effect using quadratic approximation. Overall, the results indicate the effectiveness of 0.08 BAC punishment on recidivism at least in the linear case.

As for Panel B, we can see that the coefficient of DUI becomes marginally significant in column (2), where we control for the interaction between bac1 and DUI, and insignificant in column (3), where we control for the quadratic interaction between bac1 and DUI. According to Table 3, we may infer that the regression discontinuity design is sensible to the bandwidth, but column (1) closely replicates Hensen's original result.

8. Replication of Figure3: BAC and Recidivism



Our replication of Figure 3 shows a salient plunge in recidivism either using linear fit or quadratic fit. The result is qualitatively similar to that of Hensen's, and indicates that the BAC 0.08 punishment threshold is effective in reducing recidivism. However, it is also worth noting that when offenders are way above the BAC 0.08 punishment threshold, they become more likely to recidivate

9. Conclusion

This replication exercise is very helpful to understand the intuition and the mechanism behind the RDD design. Moreover, I benefit from getting more familiar with STATA commands. Consistent with Hensen's hypothesis, I expect that DUI punishment is effective in reducing recidivism. However, the replication of the main table(Table 3) provides preliminary evidence showing that Hensen's original conclusion may be sensible to model specification (linear vs quadratic), and the underlying assumption of covariate balance may fail for some of the variables in our dataset. Overall, my results mostly support Hansen's findings in linear model specification.

Appendix. STATA Code:

```
use https://github.com/scunning1975/causal-inference-class/raw/master/hansen_dwi, clear

*Q3.Create dummy for bac>0/08
gen DUI=1 if bac1>=0.08
replace DUI=0 if DUI==.

*Q4.Figure1:BAC hitogram
cd "C:\Users\user\Desktop\Austin course\Causal Inference\Replicate1\Figures"
histogram bac1, bin(150) xline(0.08)
graph export "C:\Users\user\Desktop\Austin course\Causal
Inference\Replicate1\Figures\Figure1.png", as(png) replace
cd "C:\Users\user\Desktop\Austin course\Causal Inference\Replicate1\Do"

*Q5.Table2:Corvariate balance
ssc install outreg2
gen DUI_bac1 = bac1*DUI
reg male bac1 DUI DUI_bac1 if bac1>0.08-0.05& bac1<0.08+0.05,cluster(bac1)
outreg2 using ../Tables/Table2.doc, replace ctitle(Male) label

reg white bac1 DUI DUI_bac1 if bac1>0.08-0.05& bac1<0.08+0.05,cluster(bac1)
outreg2 using ../Tables/Table2.doc, append ctitle(white) label

reg aged bac1 DUI DUI_bac1 if bac1>0.08-0.05& bac1<0.08+0.05,cluster(bac1)
outreg2 using ../Tables/Table2.doc, append ctitle(aged) label

reg acc bac1 DUI DUI_bac1 if bac1>0.08-0.05& bac1<0.08+0.05,cluster(bac1)
outreg2 using ../Tables/Table2.doc, append ctitle(acc) label

*Q6.Figure2

*lineat(#) draws a vertical line at the specified point along the x axis.
*cutpoint(#) splits the graph at the specified point along the x axis. No bin spanning the cutoff.
*histopts(options) controls the definition of bins with histogram options.When graphs are split by
cutpoint, these options govern bindefinitions for each side of the graph
ssc install cmogram
cmogram acc bac1 if bac1 < 0.2, lineat(0.08 0.15) cutpoint(0.08)scatter lfitci title(Panel A. Accident
at scene) histopts(bin(30)) graphopts(xtitle("BAC") ytitle(""))
graph save ../figures/Fig2.ACC.gph

cmogram male bac1 if bac1 < 0.2, lineat(0.08 0.15) cutpoint(0.08)scatter lfitci title(Panel B. Male)
histopts(bin(30)) graphopts(xtitle("BAC") ytitle(""))
graph save ../figures/Fig2.male.gph
```

```
cmogram aged bac1 if bac1 < 0.2, lineat(0.08 0.15) cutpoint(0.08)scatter lfitci title(Panel C. Age)
histopts(bin(30)) graphopts(xtitle("BAC") ytitle(""))
graph save ../figures/Fig2.age.gph

cmogram white bac1 if bac1 < 0.2, lineat(0.08 0.15) cutpoint(0.08)scatter lfitci title(Panel D. White)
histopts(bin(30)) graphopts(xtitle("BAC") ytitle(""))
graph save ../figures/Fig2.white.gph

graph
combine ../figures/Fig2.ACC.gph ../figures/Fig2.male.gph ../figures/Fig2.age.gph ../figures/Fig2.w
hite.gph, title(Linear Fit) saving(../figures/Fig2.Linear.gph)
graph export "C:\Users\user\Desktop\Austin course\Causal
Inference\Replicate1\Figures\Figure2_linear.png", as(png) replace

*quadratic
cmogram acc bac1 if bac1 < 0.2, lineat(0.08 0.15) cutpoint(0.08)scatter qfitci title(Panel A.
Accident at scene) histopts(bin(30)) graphopts(xtitle("BAC") ytitle(""))
graph save ../figures/Fig2.ACC_q.gph

cmogram male bac1 if bac1 < 0.2, lineat(0.08 0.15) cutpoint(0.08)scatter qfitci title(Panel B. Male)
histopts(bin(30)) graphopts(xtitle("BAC") ytitle(""))
graph save ../figures/Fig2.male_q.gph

cmogram aged bac1 if bac1 < 0.2, lineat(0.08 0.15) cutpoint(0.08)scatter qfitci title(Panel C. Age)
histopts(bin(30)) graphopts(xtitle("BAC") ytitle(""))
graph save ../figures/Fig2.age_q.gph

cmogram white bac1 if bac1 < 0.2, lineat(0.08 0.15) cutpoint(0.08)scatter qfitci title(Panel D.
White) histopts(bin(30)) graphopts(xtitle("BAC") ytitle(""))
graph save ../figures/Fig2.white_q.gph

graph
combine ../figures/Fig2.ACC_q.gph ../figures/Fig2.male_q.gph ../figures/Fig2.age_q.gph ../figures/
Fig2.white_q.gph, title(Quadratic Fit) saving(../figures/Fig2.Quadratic.gph)
graph use "C:\Users\user\Desktop\Austin course\Causal
Inference\Replicate1\Figures\Fig2.Quadratic.gph"

*Q7.Table3 Panel A:recidivism outcome (0.03 < bac1 < 0.13)

reg recid bac1 DUI  white male age acc year if bac1>0.03& bac1<0.13,robust
outreg2 using ../Tables/Table3A.doc, replace title("Table 3 Panel A: Regression Discontinuity
Estimates for the Effect of Exceeding the 0.08 BAC Threshold on Recidivism") ctitle(1) label

reg recid bac1 DUI DUI_bac1 white male age acc  year if bac1>0.03& bac1<0.13,robust
outreg2 using ../Tables/Table3A.doc, append title("Table 3 Panel A: Regression Discontinuity
Estimates for the Effect of Exceeding the 0.08 BAC Threshold on Recidivism") ctitle(2) label
```

```
gen bac1sq_DUI= bac1*bac1*DUI
reg recid  bac1 DUI DUI_bac1 bac1sq bac1sq_DUI white male age acc year if bac1>0.03& bac1<0.13,robust
outreg2 using ../Tables/Table3A.doc, append title("Table 3 Panel A: Regression Discontinuity Estimates for the Effect of Exceeding the 0.08 BAC Threshold on Recidivism") ctitle(3) label


*Q7.Table3 Panel B:recidivism outcome (0.055 < bac1 < 0.105)

reg recid bac1 DUI white male age acc year if bac1>0.055& bac1<0.105,robust
outreg2 using ../Tables/Table3B.doc, replace title("Table 3 Panel B: Regression Discontinuity Estimates for the Effect of Exceeding the 0.08 BAC Threshold on Recidivism") ctitle(1) label

reg recid bac1 DUI DUI_bac1 white male age acc year if bac1>0.055& bac1<0.105,robust
outreg2 using ../Tables/Table3B.doc, append title("Table 3 Panel B: Regression Discontinuity Estimates for the Effect of Exceeding the 0.08 BAC Threshold on Recidivism") ctitle(2) label

gen bac1sq=bac1*bac1
reg recid bac1 bac1sq DUI DUI_bac1 bac1sq bac1sq_DUI white male age acc year if bac1>0.055& bac1<0.105,robust
outreg2 using ../Tables/Table3B.doc, append title("Table 3 Panel B: Regression Discontinuity Estimates for the Effect of Exceeding the 0.08 BAC Threshold on Recidivism") ctitle(3) label


*Q8.Figure3: BAC and Recidivism
cmogram recid bac1 if bac1 < 0.15, lineat(0.08) cutpoint(.08)scatter lfit title(Panel A. BAC and Recidivism-Linear ) histopts(bin(30)) graphopts(xtitle("BAC") ytitle(""))
graph save ../figures/Fig3.Linear.gph

cmogram recid bac1 if bac1 < 0.15, lineat(0.08) cutpoint(.08)scatter qfit title(Panel B. BAC and Recidivism-Quadratic ) histopts(bin(30)) graphopts(xtitle("BAC") ytitle(""))
graph save ../figures/Fig3.Quadratic.gph

graph combine ../figures/Fig3.Linear.gph ../figures/Fig3.Quadratic.gph, rows(2) cols(1) saving(../figures/Fig3.gph)
```