RDD

chengkan_tao ~ ~

```
devtools::install_github(c('rstudio/rmarkdown', 'yihui/tinytex'))
tinytex::install_tinytex()
```

abstract

The author tries to find the effect of punishments and deterrence on driving under the influence (DUI) through administrative records on 512,964 DUI BAC tests in the state of Washington from 1995 to 2011 by discontinuity approach. Blood alcohol content and previous offenses in the data set are used to estimate the effect. This paper finds having a BAC above either DUI or aggravated DUI threshold reduces repeat drunk driving. Having a BAC over the 0.08 DUI threshold is associated with a about 2 percent point decline in recidivism over the next four years. And, having a BAC over the 0.15 DUI will leads to about 3 percentage point decline in recidivism.

read data off-line

Q3

```
BAC = hansen_dwi %>%
  mutate_vars(.cols = bac1, .func = function(bac1) ifelse(bac1<0.08,0,1))
BAC = data.frame(BAC,hansen_dwi$bac1)
```

```
BAC = rename(BAC,c(bac1="DUI"))
BAC = rename(BAC,c(hansen_dwi.bac1="bac1"))
```

Q4

```
rdr = rdrobust(y = hansen_dwi$recidivism,x=hansen_dwi$bac1, c = 0.08)
```

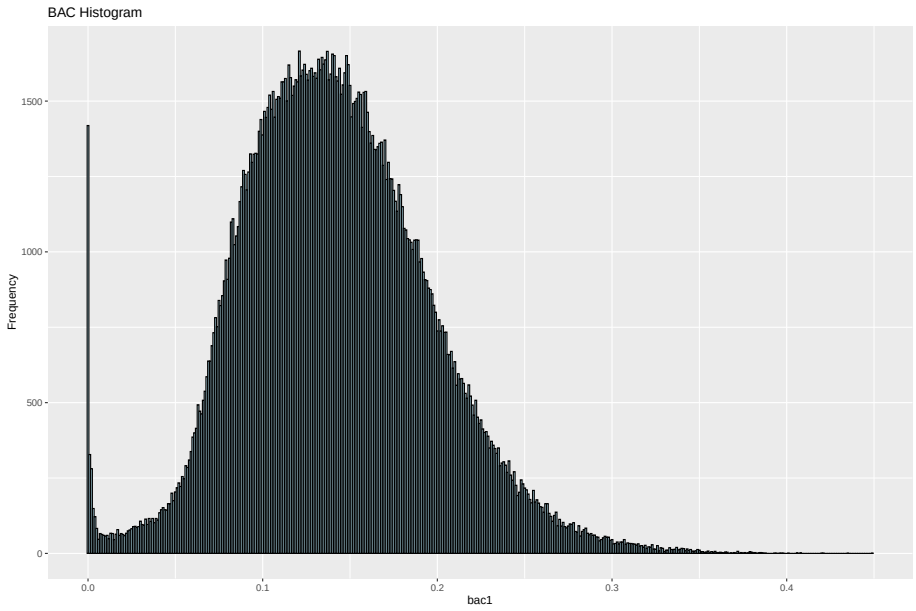
```
## [1] "Mass points detected in the running variable."
```

```
summary(rdr)
```

```
## Call: rdrobust
##
## Number of Obs.                214558
## BW type                        mserd
## Kernel                    Triangular
## VCE method                        NN
##
## Number of Obs.                23010      191548
## Eff. Number of Obs.           16574       41013
## Order est. (p)                    1           1
## Order bias  (q)                   2           2
## BW est. (h)                   0.031       0.031
## BW bias (b)                   0.050       0.050
## rho (h/b)                     0.633       0.633
## Unique Obs.                      81         318
##
## ===============================================================================
##         Method     Coef. Std. Err.         z     P>|z|      [ 95% C.I. ]
## ===============================================================================
##   Conventional    -0.018     0.006    -3.220     0.001    [-0.029 , -0.007]
```

```
##          Robust          -          -    -2.503     0.012    [-0.030 , -0.004]
## ================================================================================
```

```
ggplot(data = hansen_dwi,mapping = aes(bac1)) +
  geom_histogram(binwidth = 0.001, fill = "lightblue", colour = "black") +
  labs(title = 'BAC Histogram', y = 'Frequency', x = 'bac1')
```



To see if there's any evidence for manipulation, we would use density test. Through the picture, I cannot find a obvious discontinuity.

Q5

```
lm_q5_male <- lm_robust(male ~ DUI+bac1+bac1*DUI, data = BAC)
coef(lm_q5_male)
```

```
## (Intercept)         DUI        bac1    DUI:bac1
##  0.77344125  0.03071895  0.21824807 -0.31135817
```

```
lm_q5_white <- lm_robust(white ~ DUI+bac1+bac1*DUI, data = BAC)
coef(lm_q5_white)
```

```
## (Intercept)         DUI        bac1    DUI:bac1
## 0.835053886 0.002707367 0.153898558 0.016980500
```

```
lm_q5_acc <- lm_robust(acc ~ DUI+bac1+bac1*DUI, data = BAC)
coef(lm_q5_acc)
```

```
## (Intercept)         DUI        bac1    DUI:bac1
##   0.2007381  -0.2194854  -1.5396801   2.6559952
```

```
texreg::screenreg(list(lm_q5_male,lm_q5_white,lm_q5_acc),type="text")
```

```
##
## ============================================================
##              Model 1         Model 2         Model 3
## ------------------------------------------------------------
## (Intercept)       0.77 *          0.84 *          0.20 *
##              [ 0.76;  0.79]  [ 0.82; 0.85]  [ 0.19;  0.21]
## DUI               0.03 *          0.00           -0.22 *
##              [ 0.02;  0.05]  [-0.01; 0.02]  [-0.23; -0.21]
## bac1              0.22            0.15           -1.54 *
##              [-0.00;  0.44]  [-0.04; 0.35]  [-1.73; -1.35]
## DUI:bac1         -0.31 *          0.02            2.66 *
##              [-0.53; -0.09]  [-0.18; 0.21]  [ 2.46;  2.85]
## ------------------------------------------------------------
## R^2               0.00            0.00            0.02
## Adj. R^2          0.00            0.00            0.02
## Num. obs.    214558          214558          214558
## RMSE              0.41            0.35            0.35
## ============================================================
## * Null hypothesis value outside the confidence interval.
```

They are stable across the DUI punishment thresholds. The coefficients are not statistically significant.

Q6

```
ggplot(Q6, aes(bac1, acc)) +
  geom_point(aes(x = bac1, y = acc), data = Q6) +
  labs(title = 'Panal A', y = 'accident', x = 'bac1') +
  stat_smooth(aes(bac1, acc, group = gg_group), method = "lm") +
  xlim(0,0.5) + ylim(0,1) +
  geom_vline(xintercept = 0.08)
```

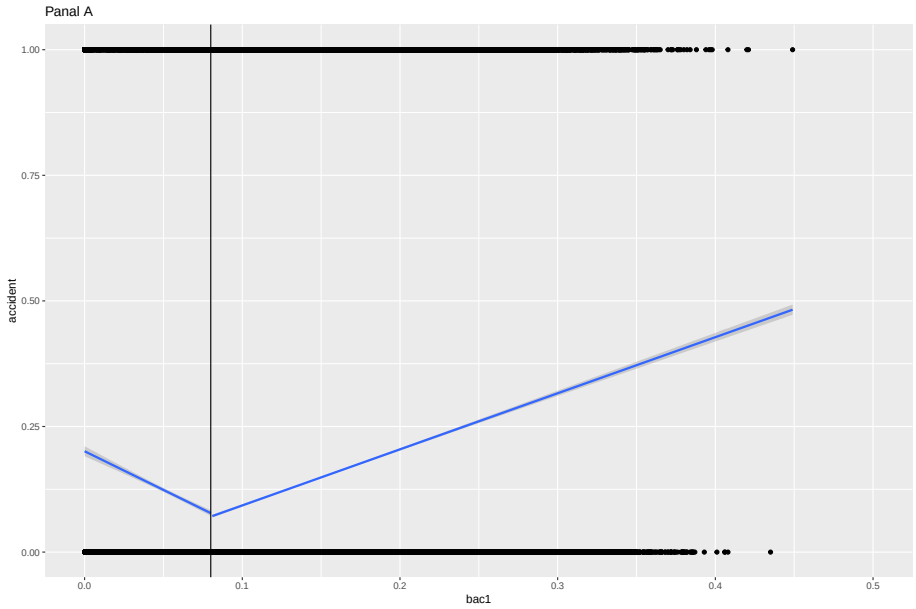


图 1: Q6

```
ggplot(Q6, aes(bac1, acc)) +
  geom_point(aes(x = bac1, y = acc), data = Q6) +
  labs(title = 'Panal A Quadratic', y = 'accident', x = 'bac1')+
```

```
  stat_smooth(aes(bac1, acc, group = gg_group), method = "lm",
              formula = y ~ x + I(x^2)) +
  xlim(0,0.5) + ylim(0,1) +
  geom_vline(xintercept = 0.08)
```

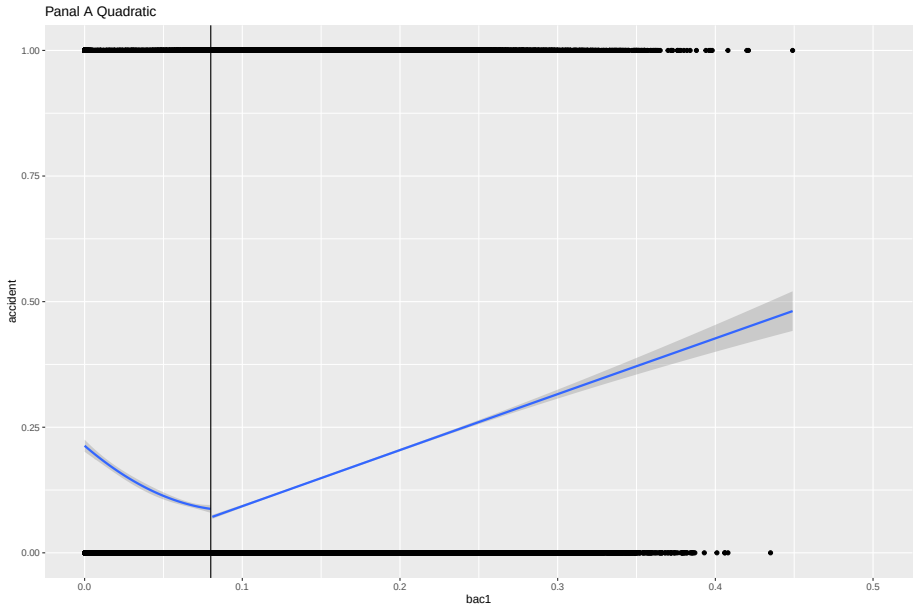


图 2: Q6

```
ggplot(Q6, aes(bac1, male)) +
  geom_point(aes(x = bac1, y = male), data = Q6) +
  labs(title = 'Panal B', y = 'male', x = 'bac1') +
  stat_smooth(aes(bac1, male, group = gg_group), method = "lm") +
  xlim(0,0.5) + ylim(0,1) +
  geom_vline(xintercept = 0.08)
```

```
ggplot(Q6, aes(bac1, male)) +
  geom_point(aes(x = bac1, y = male), data = Q6) +
  labs(title = 'Panal B Quadratic', y = 'male', x = 'bac1')+
  stat_smooth(aes(bac1, male, group = gg_group), method = "lm",
```

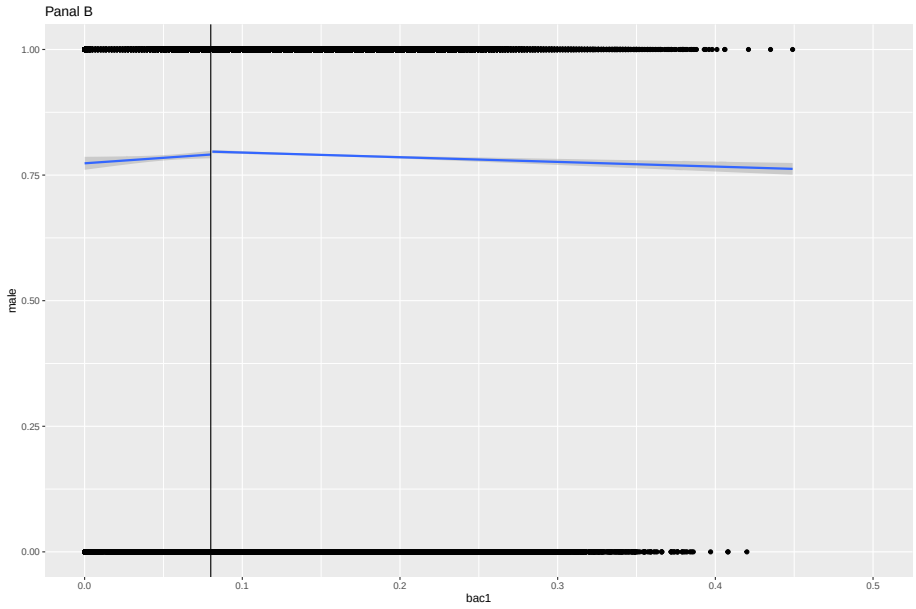


图 3: Q6

```
              formula = y ~ x + I(x^2)) +
  xlim(0,0.5) + ylim(0,1) +
  geom_vline(xintercept = 0.08)
```

```
ggplot(Q6, aes(bac1, aged)) +
  geom_point(aes(x = bac1, y = aged), data = Q6) +
  labs(title = 'Panal C', y = 'aged', x = 'bac1') +
  stat_smooth(aes(bac1, aged, group = gg_group), method = "lm") +
  xlim(0,0.5) + ylim(0,80) +
  geom_vline(xintercept = 0.08)
```

```
ggplot(Q6, aes(bac1, aged)) +
  geom_point(aes(x = bac1, y = aged), data = Q6) +
  labs(title = 'Panal C Quadratic', y = 'aged', x = 'bac1')+
  stat_smooth(aes(bac1, aged, group = gg_group), method = "lm",
```

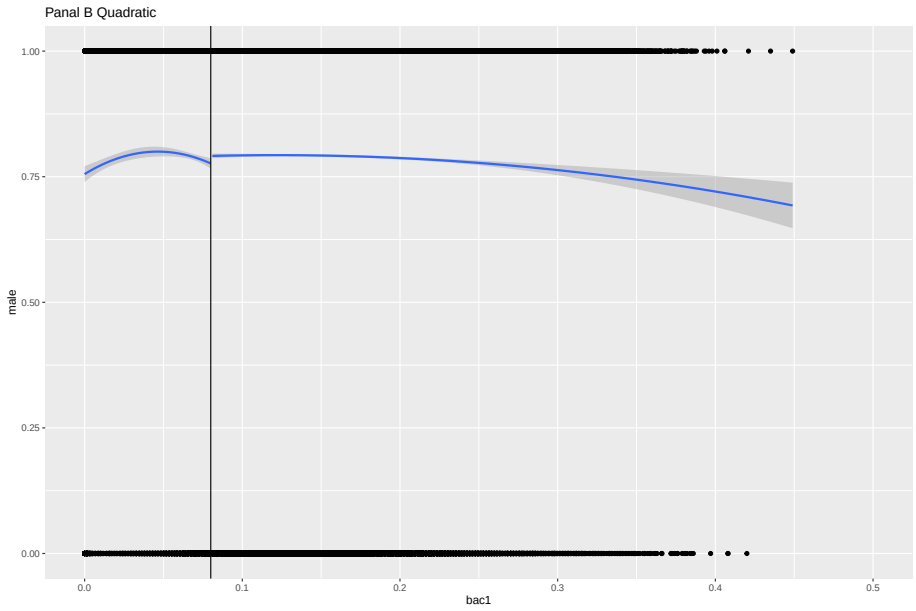

Panal B Quadratic
1.00
0.75
0.50
0.25
0.00
male
0.0
0.1
0.2
0.3
0.4
0.5
bac1

图 4: Q6

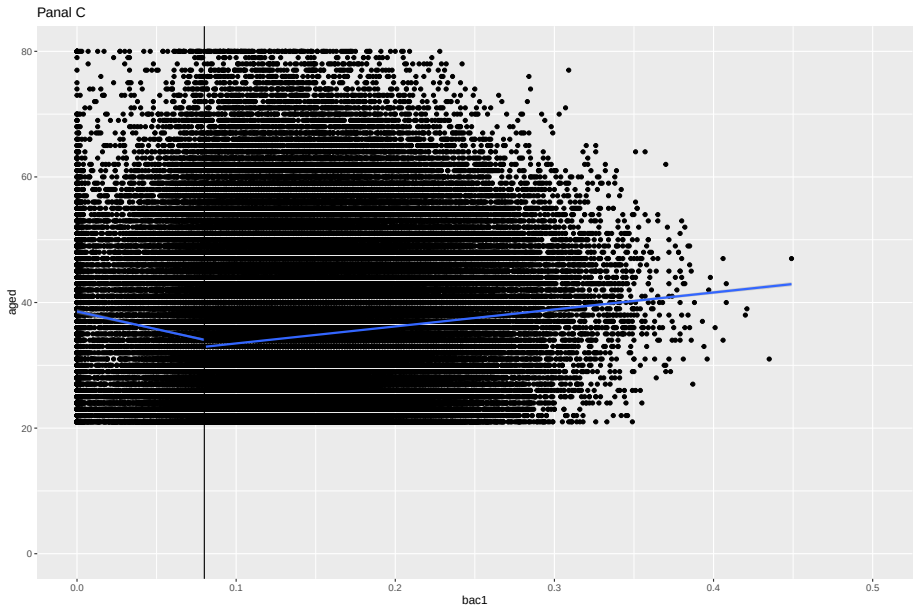

Panal C
80
60
40
20
0
aged
0.0
0.1
0.2
0.3
0.4
0.5
bac1

图 5: Q6

```
              formula = y ~ x + I(x^2)) +
  xlim(0,1) + ylim(0,80) +
  geom_vline(xintercept = 0.08)
```

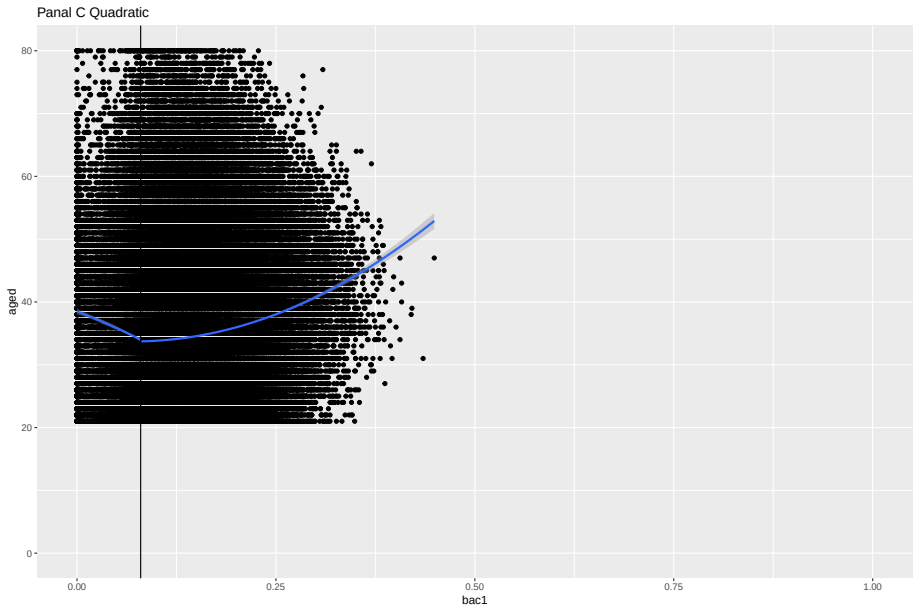


图 6: Q6

```
ggplot(Q6, aes(bac1, white)) +
  geom_point(aes(x = bac1, y = white), data = Q6) +
  labs(title = 'Panal D', y = 'white', x = 'bac1') +
  stat_smooth(aes(bac1, white, group = gg_group), method = "lm") +
  xlim(0,1) + ylim(0,1) +
  geom_vline(xintercept = 0.08)
```

```
ggplot(Q6, aes(bac1, white)) +
  geom_point(aes(x = bac1, y = white), data = Q6) +
  labs(title = 'Panal D Quadratic', y = 'white', x = 'bac1')+
  stat_smooth(aes(bac1, white, group = gg_group), method = "lm",
              formula = y ~ x + I(x^2)) +
```

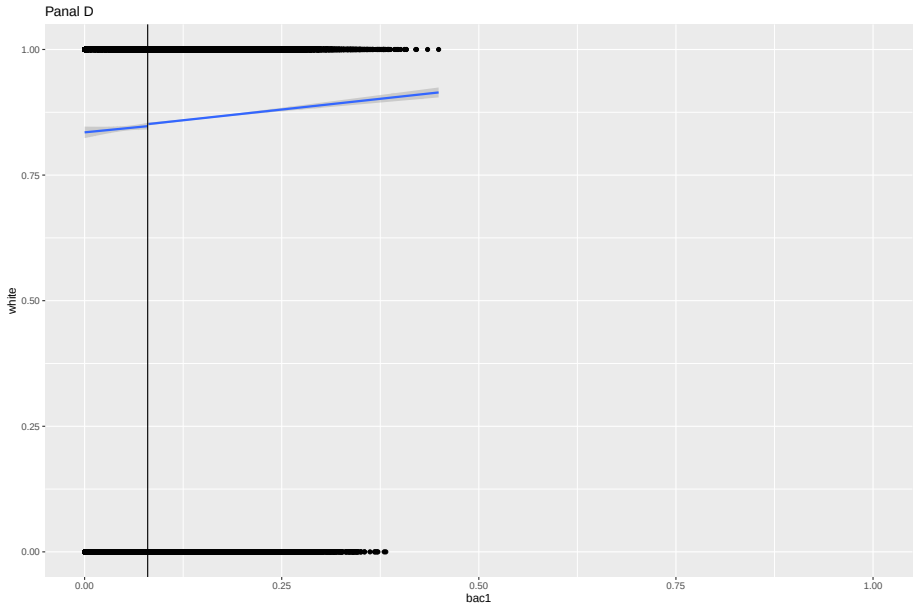
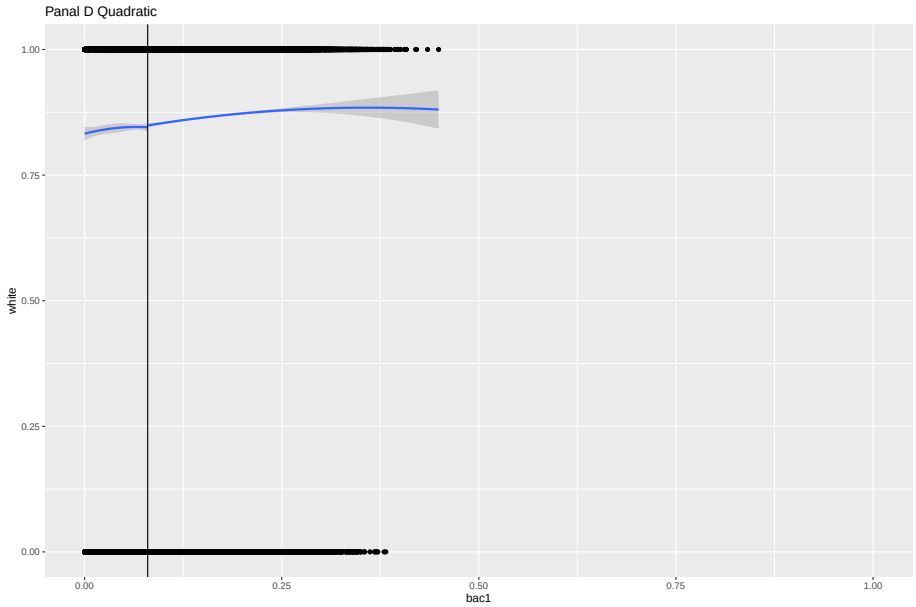


图 7: Q6

```
  xlim(0,1) + ylim(0,1) +
  geom_vline(xintercept = 0.08)
```



Through these figures, I find these predetermined characteristics remain unchanged across the punishment thresholds. It's same with what hansen has found.

Q7

simple linear regression. I try to use loess(), but I find texreg::screenreg() cannot show the results

```
lm_q7_wide1 = lm_robust(recidivism ~ bac1, data = Q7_wide)
coef(lm_q7_wide1)
```

```
## (Intercept)        bac1
##  0.11342691 -0.06798084
```

SIMPLE LINEAR REGRESSION. I TRY TO USE LOESS(), BUT I FIND TEXREG::SCREENREG() C

```
lm_q7_wide2 = lm_robust(recidivism ~ bac1*DUI, data = Q7_wide)
coef(lm_q7_wide2)
```

```
##  (Intercept)         bac1          DUI     bac1:DUI
##  0.116596893  0.005875803 -0.054180587  0.383849914
```

```
lm_q7_wide3 = lm_robust(recidivism ~ bac1*DUI+bac1_sq*DUI, data = Q7_wide)
coef(lm_q7_wide3)
```

```
##  (Intercept)         bac1          DUI      bac1_sq     bac1:DUI  DUI:bac1_sq
##   0.03377996   2.93861639   0.10811349 -24.61441633  -4.07793624  31.84326212
```

```
lm_q7_narrow1 = lm_robust(recidivism ~ bac1, data = Q7_narrow)
coef(lm_q7_narrow1)
```

```
## (Intercept)        bac1
##   0.1472530  -0.4917993
```

```
lm_q7_narrow2  = lm_robust(recidivism ~ bac1*DUI, data = Q7_narrow)
coef(lm_q7_narrow2)
```

```
## (Intercept)        bac1         DUI    bac1:DUI
##  0.12719210 -0.15031633 -0.05593986  0.45190372
```

```
lm_q7_narrow3  = lm_robust(recidivism ~ bac1*DUI + bac1_sq*DUI, data = Q7_narrow)
coef(lm_q7_narrow3)
```

```
##  (Intercept)         bac1          DUI      bac1_sq     bac1:DUI  DUI:bac1_sq
##  -0.05279834   5.11557262   0.25322793 -38.12431487  -7.61232266  53.19613891
```

```
texreg::screenreg(list(lm_q7_wide1,lm_q7_wide2,lm_q7_wide3),type="text")
```

```
##
## ============================================================
##              Model 1        Model 2         Model 3
## ------------------------------------------------------------
## (Intercept)      0.11 *         0.12 *          0.03
##              [ 0.10; 0.12]  [ 0.09;  0.14]  [ -0.06;  0.13]
## bac1            -0.07           0.01            2.94
##              [-0.16; 0.03]  [-0.36;  0.37]  [ -0.28;  6.16]
## DUI                           -0.05 *          0.11
##                              [-0.08; -0.02]  [ -0.06;  0.28]
## bac1:DUI                      0.38           -4.08
##                              [-0.02;  0.79]  [ -8.30;  0.15]
## bac1_sq                                      -24.61
##                                             [-51.61;  2.38]
## DUI:bac1_sq                                    31.84 *
##                                             [  1.90; 61.78]
## ------------------------------------------------------------
## R^2              0.00           0.00            0.00
## Adj. R^2         0.00           0.00            0.00
## Num. obs.    88373          88373           88373
## RMSE             0.31           0.31            0.31
## ============================================================
## * Null hypothesis value outside the confidence interval.
```

```
texreg::screenreg(list(lm_q7_narrow1,lm_q7_narrow2,lm_q7_narrow3),type="text")
```

```
##
## ============================================================
##              Model 1         Model 2        Model 3
## ------------------------------------------------------------
## (Intercept)      0.15 *          0.13 *        -0.05
##              [ 0.13;  0.17]  [ 0.07; 0.18]  [  -0.60;   0.49]
## bac1            -0.49 *         -0.15           5.12
##              [-0.72; -0.27]  [-0.90; 0.60]  [ -10.83;  21.06]
```

```
## DUI                              -0.06           0.25
##                                [-0.13; 0.01]  [  -0.62;   1.13]
## bac1:DUI                          0.45          -7.61
##                                [-0.45; 1.35]  [ -29.29;  14.07]
## bac1_sq                                         -38.12
##                                               [-153.46;  77.21]
## DUI:bac1_sq                                      53.20
##                                               [ -86.68; 193.07]
## ------------------------------------------------------------------
## R^2              0.00            0.00            0.00
## Adj. R^2         0.00            0.00            0.00
## Num. obs.    45484           45484           45484
## RMSE             0.31            0.31            0.31
## ==================================================================
## * Null hypothesis value outside the confidence interval.
```

Q8

```
ggplot(Q8, aes(bac1, recidivism)) +
  geom_point(aes(x = bac1, y = recidivism), data = Q8) +
  stat_smooth(aes(bac1, recidivism, group = gg_group), method = "lm") +
  xlim(0,0.15) + ylim(0,1) +
  geom_vline(xintercept = 0.08)
```

```
ggplot(Q8, aes(bac1, recidivism)) +
  geom_point(aes(x = bac1, y = recidivism), data = Q8) +
  stat_smooth(aes(bac1, recidivism, group = gg_group), method = "lm",
              formula = y ~ x + I(x^2)) +
  xlim(0,0.15) + ylim(0,1) +
  geom_vline(xintercept = 0.08)
```

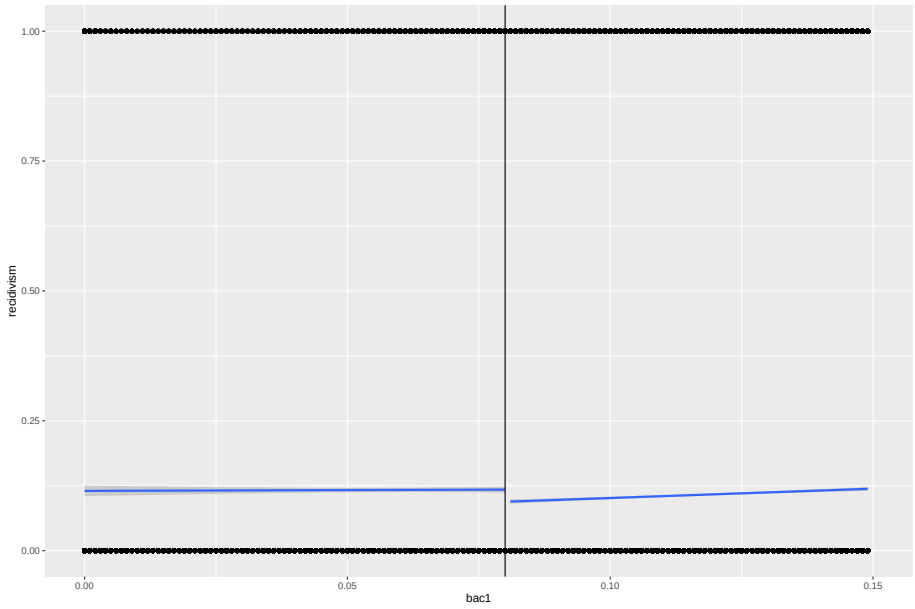

1.00
0.75
0.50
0.25
0.00
recidivism
0.00
0.05
0.10
0.15
bac1

图 8: Q8

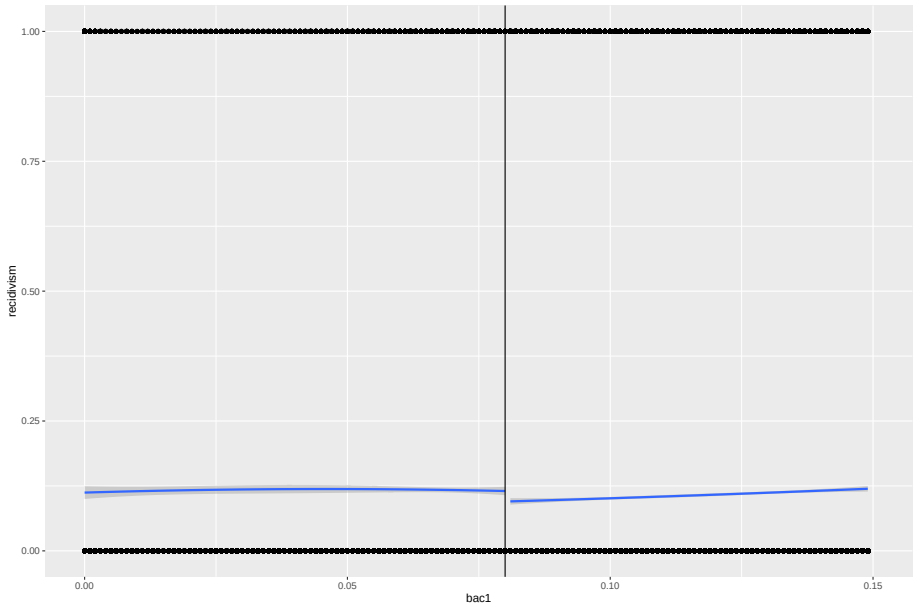

1.00
0.75
0.50
0.25
0.00
recidivism
0.00
0.05
0.10
0.15
bac1

图 9: Q8

Q9

Regression discontinuity design gives causal effect of treatment on the outcome. In this exercise, I use BAC=0.08 as a cutoff and use density test to find whether agents can manipulate running variable score. By recreating Table 2 and Figure, I fail to reject the null hypothesis that predetermined characters are unrelated to DUI. That's the same with hansen's paper.