RDD Replication

Robert Toto

2/19/2021

One

https://github.com/rvt245/RDD

Two

Summary of Hansen Paper

Research Question

Hansen is asking whether the harsher punishments received for higher blood-alcohol-content (BAC) cutoffs are effective in reducing future drunk driving. In Washington, there are two BAC thresholds for a DUI. If a driver has a BAC above 0.08, he or she receives a certain combination of fines, jail time, and license suspension. However, if the driver has a BAC above 0.15, he or she receives harsher penalties. Hansen wants to know whether the harsher penalties reduce future drunk driving more than the less harsh penalties at the lower BAC cutoff.

Data

Hansen begins with data on DUI stops from the state of Washington from 1995 to 2011, including a total of 512,964 DUI BAC tests. He then narrows this to only DUI stops between 1999 and 2007. In 1999, Washington state legislated a new dual-threshold system (0.08 BAC for DUI and 0.15 BAC for "aggravated" DUI). Furthermore, Hansen stops looking at new DUIs in 2007 because he wants to allow for four years (up to 2011) of additional DUI stops to find repeated offenses. This is crucial because Hanses is analyzing "recidivism" (or repeated offenses). In Hansen's data, recidivism is identified if an offender receives a DUI within four years of his or her initial DUI. The data is also restricted to persons above the drinking age.

Research Design

This is a quasi-experiment because there is no direct randomization of the treatment assignment. More specifically, Hansen uses a Regression Discontinuity (RD) design because the BAC levels that legally determine a DUI provide cutoffs that lead to a natural treatment assignment. Since the DUI law generates a deterministic and discontinuous function of BAC, a Sharp RD can be used. Smoothness (or "continuity") is assumed such that unobservables are expected to remain unchanged across the threshold. This assumption is predicated partly on the assumption that drunk drivers cannot discern their level of impairment prior to driving and thus will not sort themselves out of the running variable by not driving. Identification is established by assuming that it is locally random for a driver to have a BAC either just above or just below the BAC threshold. In other words, some people at the threshold are just lucky, and fall just below the

legal limit. These assumptions are reinforced by the fact that BAC measurements are precise and difficult to manipulate (making sorting at the cutoff very unlikely), yet slight variation between different breathalyzer devices may allow for randomness to occur in the measurement of BAC around the cutoff. The running variable is the minimum BAC measurement at a DUI stop, where two BAC measurements are taken per stop, because the minimum determines legal guilt. For the regression models in the RD design, Hansen uses a local linear regression to estimate the effect of having a BAC above the DUI or aggravated DUI threshold on recidivism. A rectangular kernel is used to weight the regression near the cutoff. The running variable (BAC) is recentered around the relevant threshold (0.08 or 0.15). The outcome variable is recidivism. Pretreatment characteristics related to DUI stops are stable. Bandwidths are varied from 0.05 to 0.025 BAC.

Conclusions

Hansen generally finds that harsher penalties for BAC cutoffs do in fact reduce future drunk driving. When considering the 0.08 BAC threshold, Hansen finds that this threshold reduces recidivism in the following 4 years by 2 percentage points (and is statistically significant at 0.01 level). This finding is consistent across both 0.05 and 0.025 BAC bandwidths. Individuals with at least one prior BAC test reduce recidivism more than those with no prior tests. When considering both the 0.08 and 0.15 thresholds, Hansen finds a notable drop in recidivism for both, indicating that the increased punishments occurring at each increasing threshold does have an effect of reducing recidivism. Across offender with varying levels of prior tests, having a BAC over the legal cutoffs are associated with lower recidivism. At the 0.15 cutoff, for instance, drivers receiving an aggravated DUI reduce recidivism by an additional 1.1 percentage points. Those with at least one prior test reduce recidivism by an additional 2.2 percentage points, further indicating that harshness of punishment leads to less recidivism, as repeated offenses have harsher punishments. The regression is also carried out on disaggregated definitions of recidivism using a rectangular kernel. In these cases, having a BAC over a legal limit continues to be associated with lowered recidivism, except in the case of refusal to breathalyze. This disaggregation also reveals that having a BAC over either of the thresholds is associated with decreased future automobile accidents. Further analysis is conducted on the effect of exceeding a threshold on the punishments, sanctions, and treatments received in court. Hansen finds that surpassing a threshold is associated with harsher punishments and sanctions. Overall, Hansen finds that exceeding a threshold is associated with reduced drunk driving (recidivism) in the short run and the long run.

Three

```
hansen <- hansen %>%
  mutate(dui = ifelse(bac1 >= 0.08 & bac1 < 0.15, 1, 0))

head(hansen)
```

```
##         Date Alcohol1 Alcohol2 low_score male white recidivism acc aged year
## 1 30jul2007        0        0         0    1     1          1   0   48 2007
## 2 20feb2007        0        0         0    1     1          0   0   51 2007
## 3 18mar2003        0        0         0    1     0          0   0   68 2003
## 4 17dec2006        0        0         0    1     1          0   0   51 2006
## 5 07apr1999        0        0         0    1     1          0   0   22 1999
## 6 20mar2003        0        0         0    1     1          0   1   49 2003
##   bac1 bac2 dui
## 1    0    0   0
## 2    0    0   0
## 3    0    0   0
## 4    0    0   0
```

```
## 5      0      0      0
## 6      0      0      0
```

Four

If people could manipulate their BAC, we would want to use a density test to see whether there is bunching or "sorting" at the 0.08 BAC cutoff. The McCrary Density Test if a popular test to identify such manipulation. The McCrary Density Test gives a null hypothesis that the density in the data is continuous at the cutoff. The alternative hypothesis is that there is greater density (i.e. bunching) around the cutoff. The test separated the data into bins and calculates the frequency of observations in each bin. To reject the null, one would have to see a greater frequency of observations in the bins next to the cutoff, compared to bins further away. In R, we can use "rddensity" to run this test. We can also inspect the histogram of BAC data to look for visual bunching in the running variable near the cutoff.

```r
#BAC Histogram
ggplot(data=hansen) +
  geom_histogram(aes(x=bac1, after_stat(density)), binwidth = .003) +
  labs(x = 'Blood Alcohol Content (%)',
       y = 'Frequency',
       title = 'BAC Histogram') +
  theme_bw() +
  scale_x_continuous(breaks=seq(0,0.4,0.04)) +
  geom_vline(xintercept = 0.08)
```
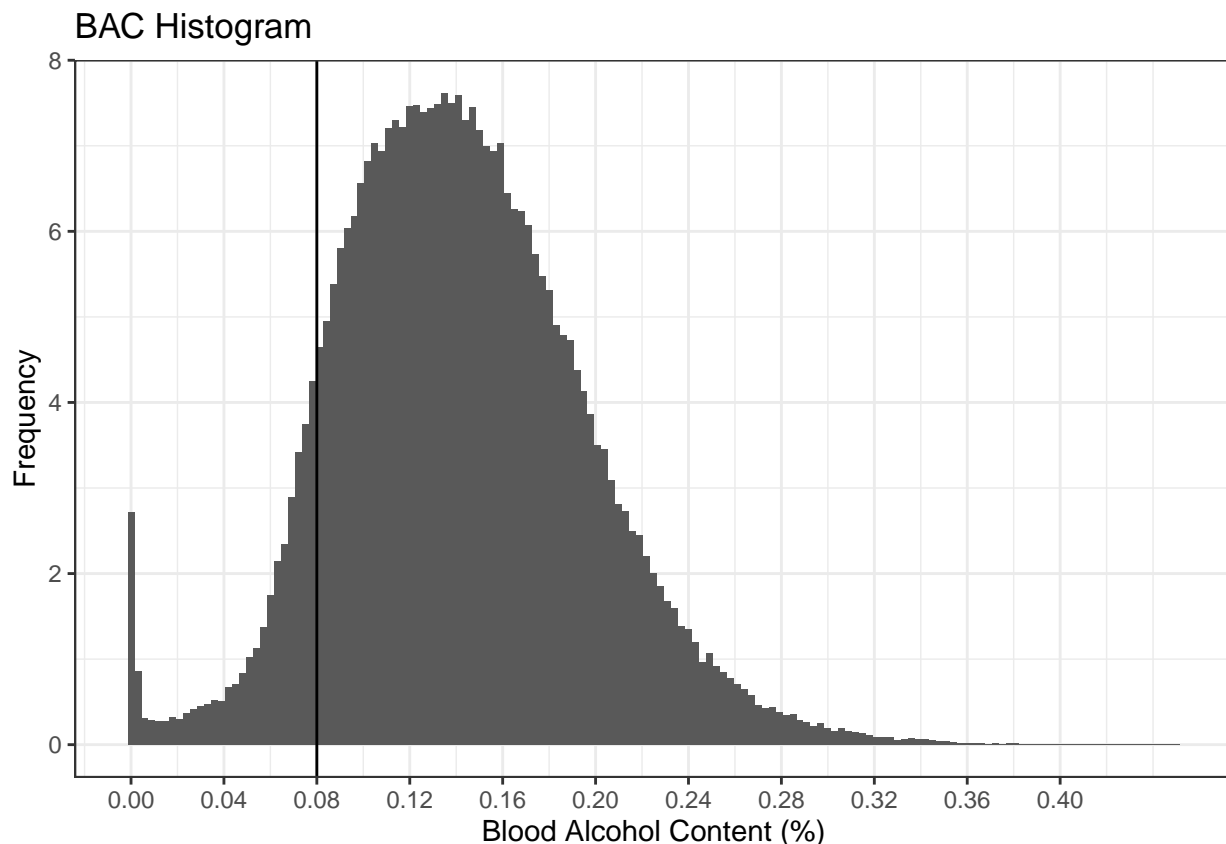

The histogram of BAC observations shows no visual sign of bunching at the 0.08 BAC cutoff for a DUI.

The distribution illustrates a smooth increase in bin frequencies as BAC increases with no noticeable distontinuous jump in the frequency of BAC observations near the cutoff.

```
library(rdd)
library(rddensity)

#Density Test
density <- rddensity(hansen$bac1, c = 0.08, massPoints=FALSE)
#rdplotdensity(density, hansen$bac1)
summary(density)
```

```
##
## Manipulation testing using local polynomial density estimation.
##
## Number of obs =          214558
## Model =                  unrestricted
## Kernel =                 triangular
## BW method =              estimated
## VCE method =             jackknife
##
## c = 0.08                 Left of c          Right of c
## Number of obs            23010              191548
## Eff. Number of obs       14442              28946
## Order est. (p)           2                  2
## Order bias (q)           3                  3
## BW est. (h)              0.023              0.023
##
## Method                   T                  P > |T|
## Robust                   -0.2299            0.8182
##
##
## P-values of binomial tests (H0: p=0.5).
##
## 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
```

The density test confirms that there is no manipulation of BAC at the cutoff. The test gives a p-value of 0.8182, indicating we should not reject the null (that the density in the data is continuous at the cutoff). While there may be some heaping due to rounding and the choice of bin-widths, the point estimates model of BAC frequency moves smoothly across the cutoff with no discontinuous jump. Overall, this is strong evidence that drivers did not manipulate BAC results. This also aligns with the real-world understanding that breathalyzers are difficult to cheat. This conclusion is also consistent with Hansen's McCrary density test results which showed "little evidence of endogenous sorting to one side of either of the thresholds studied." Although, the p-value for the 0.08 cutoff in Hansen's results was lower at 0.59.

4

Five

```r
#Check if covariates are balanced at the cutoff
library(stargazer)

#Recenter the running variable first
hansen_c <- hansen %>%
  mutate(bac1_c = bac1 - 0.08)

#Isolate data to 0.05 Bandwidth
hansen1 <- hansen_c %>%
  filter(bac1_c >=-0.05 & bac1_c <= 0.05)

#Check if covariates are balanced at the cutoff
m1 = lm(male ~ dui*bac1_c, data=hansen1)
stargazer(m1, type = "text")
```

```
##
## =================================================
## 					Dependent variable:
## 				 ----------------------------
## 							male
## -------------------------------------------------
## dui							0.006
## 							(0.006)
##
## bac1_c						-0.210
## 							(0.240)
##
## dui:bac1_c					0.307
## 							(0.263)
##
## Constant					0.784***
## 							(0.005)
##
## -------------------------------------------------
## Observations				89,967
## R2							0.0001
## Adjusted R2					0.00002
## Residual Std. Error	0.406 (df = 89963)
## F Statistic			1.519 (df = 3; 89963)
## =================================================
## Note:				*p<0.1; **p<0.05; ***p<0.01
```

```r
m2 = lm(white ~ dui*bac1_c, data=hansen1)
stargazer(m2, type = "text")
```

```
##
## =================================================
## 					Dependent variable:
## 				 ----------------------------
## 							white
```

```
## ----------------------------------------------
## dui                               0.006
##                                  (0.005)
##
## bac1_c                            0.079
##                                  (0.210)
##
## dui:bac1_c                        0.016
##                                  (0.230)
##
## Constant                         0.846***
##                                  (0.004)
##
## ----------------------------------------------
## Observations                     89,967
## R2                               0.0001
## Adjusted R2                      0.0001
## Residual Std. Error     0.355 (df = 89963)
## F Statistic          3.897*** (df = 3; 89963)
## ==============================================
## Note:                *p<0.1; **p<0.05; ***p<0.01
```

```
m3 = lm(aged ~ dui*bac1_c, data=hansen1)
stargazer(m3, type = "text")
```

```
##
## ================================================
##                      Dependent variable:
##                  ----------------------------
##                             aged
## ------------------------------------------------
## dui                        -0.140
##                           (0.162)
##
## bac1_c                   -69.164***
##                           (6.844)
##
## dui:bac1_c                76.049***
##                           (7.508)
##
## Constant                 33.920***
##                           (0.131)
##
## ------------------------------------------------
## Observations               89,967
## R2                         0.002
## Adjusted R2                0.002
## Residual Std. Error   11.573 (df = 89963)
## F Statistic         70.505*** (df = 3; 89963)
## ================================================
## Note:                *p<0.1; **p<0.05; ***p<0.01
```

```
m4 = lm(acc ~ dui*bac1_c, data=hansen1)
stargazer(m4, type = "text")
```

```
##
## ===============================================
##                      Dependent variable:
##                ----------------------------
##                             acc
## -----------------------------------------------
## dui                        -0.003
##                            (0.004)
##
## bac1_c                    -1.096***
##                            (0.178)
##
## dui:bac1_c                 1.888***
##                            (0.196)
##
## Constant                   0.083***
##                            (0.003)
##
## -----------------------------------------------
## Observations                89,967
## R2                          0.002
## Adjusted R2                 0.001
## Residual Std. Error    0.301 (df = 89963)
## F Statistic         45.207*** (df = 3; 89963)
## ===============================================
## Note:              *p<0.1; **p<0.05; ***p<0.01
```

The BAC was recentered around 0.08 and limited to a bandwidth of 0.05. A local linear regression was applied, which naturally uses a rectangular kernel. These regression results show the regression discontinuity estimates for the effect of exceeding the 0.08 BAC threshold on pretreatment characteristics (i.e. male, white, age, accident). These regressions fail to reject the null that the pretreatment characteristics are unrelated to the 0.08 BAC cutoff for a DUI. This indicates that these demographic factors are stable ("balanced") across the BAC cutoff. These results mirror the findings in Table 2, Panel A of Hansen 2015. The coefficient values are slightly different from those in Hansen because our analysis only uses the 0.08 BAC cutoff for a DUI, whereas Hansen uses both cutoffs (0.08 and 0.15). The stability of these demographic factors lends credibility to the regression discontinuity model providing unbiased estimates on recidivism.

Six

```r
#Prepare the data for discontinuity modeling of pretreatment characteristics
categories <- hansen$bac1

hansen2 <- hansen %>%
  filter(bac1 >= 0 & bac1 <= 0.2)

means_male <- split(hansen2$male, cut(hansen2$bac1, 191)) %>%
  lapply(mean) %>%
  unlist()
agg_hansen_male <- data.frame(male = means_male, bac1 = seq(0.01, 0.2, by = 0.001))

means_white <- split(hansen2$white, cut(hansen2$bac1, 191)) %>%
  lapply(mean) %>%
  unlist()
agg_hansen_white <- data.frame(white = means_white, bac1 = seq(0.01, 0.2, by = 0.001))

means_acc <- split(hansen2$acc, cut(hansen2$bac1, 191)) %>%
  lapply(mean) %>%
  unlist()
agg_hansen_acc <- data.frame(acc = means_acc, bac1 = seq(0.01, 0.2, by = 0.001))

means_aged <- split(hansen2$aged, cut(hansen2$bac1, 191)) %>%
  lapply(mean) %>%
  unlist()
agg_hansen_aged <- data.frame(aged = means_aged, bac1 = seq(0.01, 0.2, by = 0.001))

hansen2 <- hansen2 %>%
  mutate(gg_group = case_when(bac1 > 0.08 ~ 1, TRUE ~ 0))
```

```r
#Model "male" Stability Using Linear Model and Quadratic models
lmale <- ggplot(hansen2, aes(bac1, male)) +
  geom_point(aes(x = bac1, y = male), data = agg_hansen_male) +
  geom_vline(xintercept = 0.08, colour = "grey", linetype = 2) +
  stat_smooth(aes(bac1, male, group = gg_group), method = "lm") +
  labs(x = "BAC")

qmale <- ggplot(hansen2, aes(bac1, male)) +
  geom_point(aes(x = bac1, y = male), data = agg_hansen_male) +
  geom_vline(xintercept = 0.08, colour = "grey", linetype = 2) +
  stat_smooth(aes(bac1, male, group = gg_group), method = "lm", formula = y ~ x + I(x^2)) +
  labs(x = "BAC")

#Model "white" Stability Using Linear Model and Quadratic models
lwhite <- ggplot(hansen2, aes(bac1, white)) +
  geom_point(aes(x = bac1, y = white), data = agg_hansen_white) +
  geom_vline(xintercept = 0.08, colour = "grey", linetype = 2) +
  stat_smooth(aes(bac1, white, group = gg_group), method = "lm") +
  labs(x = "BAC")

qwhite <- ggplot(hansen2, aes(bac1, white)) +
  geom_point(aes(x = bac1, y = white), data = agg_hansen_white) +
  geom_vline(xintercept = 0.08, colour = "grey", linetype = 2) +
  stat_smooth(aes(bac1, white, group = gg_group), method = "lm", formula = y ~ x + I(x^2)) +
  labs(x = "BAC")

#Model "accident" Stability Using Linear Model and Quadratic models
lacc <- ggplot(hansen2, aes(bac1, acc)) +
  geom_point(aes(x = bac1, y = acc), data = agg_hansen_acc) +
  geom_vline(xintercept = 0.08, colour = "grey", linetype = 2) +
  stat_smooth(aes(bac1, acc, group = gg_group), method = "lm") +
  labs(x = "BAC")

qacc <- ggplot(hansen2, aes(bac1, acc)) +
  geom_point(aes(x = bac1, y = acc), data = agg_hansen_acc) +
  geom_vline(xintercept = 0.08, colour = "grey", linetype = 2) +
  stat_smooth(aes(bac1, acc, group = gg_group), method = "lm", formula = y ~ x + I(x^2)) +
  labs(x = "BAC")

#Model "age" Stability Using Linear Model and Quadratic models
lage <- ggplot(hansen2, aes(bac1, aged)) +
  geom_point(aes(x = bac1, y = aged), data = agg_hansen_aged) +
  geom_vline(xintercept = 0.08, colour = "grey", linetype = 2) +
  stat_smooth(aes(bac1, aged, group = gg_group), method = "lm") +
  labs(x = "BAC")

qage <- ggplot(hansen2, aes(bac1, aged)) +
  geom_point(aes(x = bac1, y = aged), data = agg_hansen_aged) +
  geom_vline(xintercept = 0.08, colour = "grey", linetype = 2) +
  stat_smooth(aes(bac1, aged, group = gg_group), method = "lm", formula = y ~ x + I(x^2)) +
  labs(x = "BAC")

library(gridExtra)
```
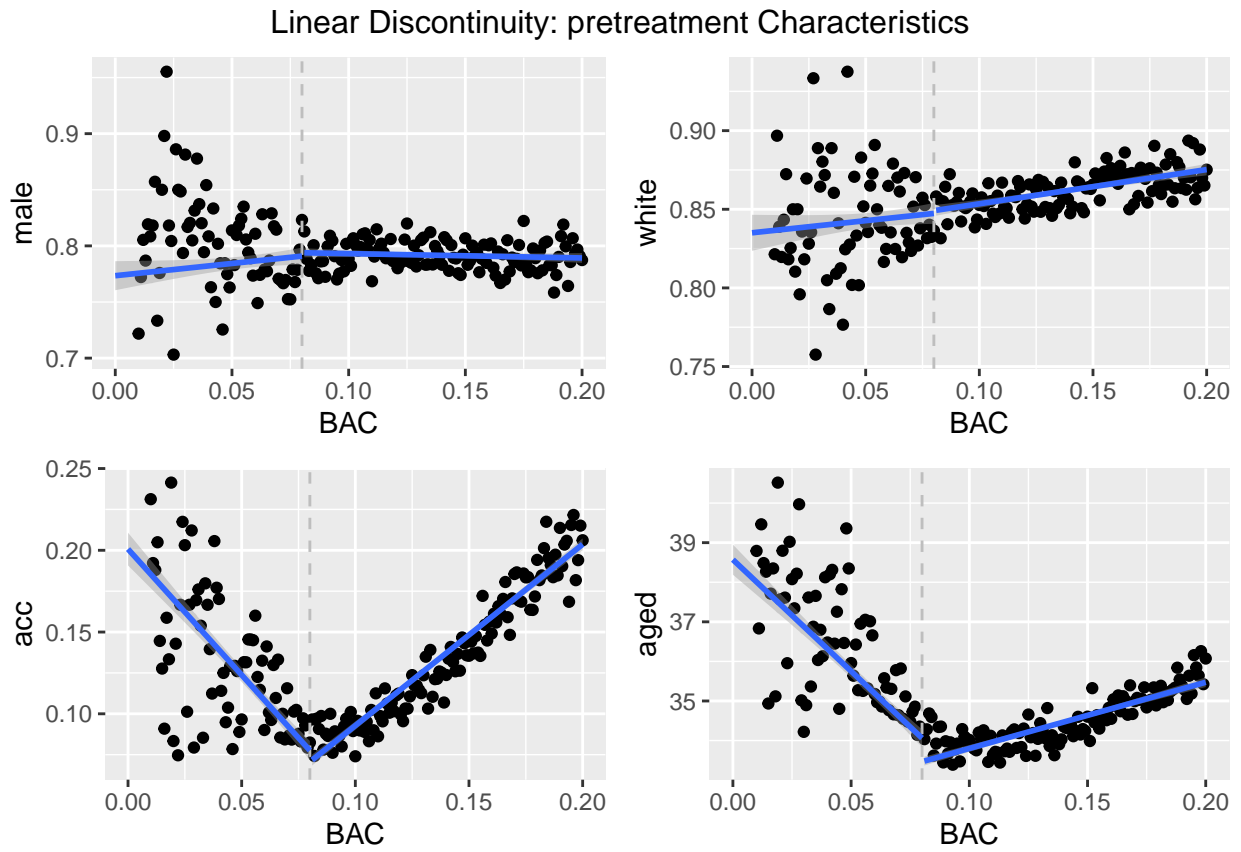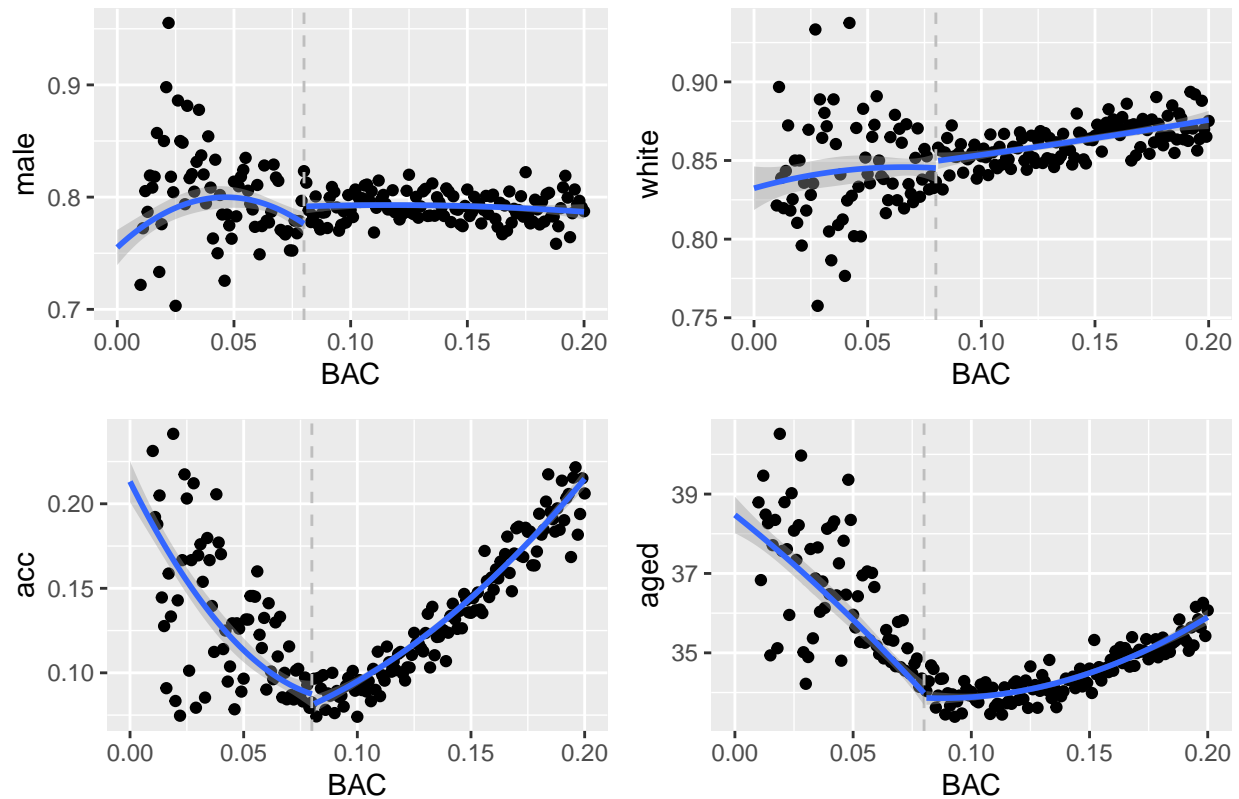
```
grid.arrange(lmale, lwhite, lacc, lage, nrow=2, top="Linear Discontinuity: pretreatment Characteristics
```

Linear Discontinuity: pretreatment Characteristics


```
grid.arrange(qmale, qwhite, qacc, qage, nrow=2, top="Quadratic Discontinuity: pretreatment Characterist
```

Quadratic Discontinuity: pretreatment Characteristics

There is not an observable discontinuous change in the average values of the four pretreatment characteristic covariates (male, white, accidents, and age) around the 0.08 BAC cutoff. This provides indirect evidence to support the continuity assumption of the RD design on recidivism. In other words, these covariates are appear to be invariant to change in treatment assignment.

[COMPARE TO HANSEN.....]

Seven

```
library(RDHonest)

##--BINWIDTH 0.05--##
#Binwidth 0.05: control for bac1 linearly
RDestimate(recidivism ~ bac1, cutpoint = 0.08, bw = 0.05, data=hansen)
```

```
##
## Call:
## RDestimate(formula = recidivism ~ bac1, data = hansen, cutpoint = 0.08,
##     bw = 0.05)
##
## Coefficients:
##      LATE    Half-BW   Double-BW
##  -0.01988   -0.01808    -0.02151
```

```
RDHonest(recidivism ~ bac1, cutoff = 0.08, h=0.05, kern="triangular", M=0.1, sclass="T", order = 1, data
```

```
## Call:
## RDHonest(formula = recidivism ~ bac1, data = hansen, cutoff = 0.08,
##     M = 0.1, kern = "triangular", h = 0.05, sclass = "T", order = 1)
##
##
## Inference by se.method:
##       Estimate Maximum Bias Std. Error
## nn -0.01987723 4.049288e-05 0.00471454
##
## Confidence intervals:
## nn    (-0.0291179, -0.01063656), (-0.02767245, Inf), (-Inf, -0.01208201)
##
## Bandwidth: 0.05
## Number of effective observations: 18911.67
```

```
#Binwidth 0.05: interact bac1 with cutoff linearly
RDestimate(recidivism ~ bac1 | bac1*dui, cutpoint = 0.08, bw = 0.05, data=hansen)
```

```
##
## Call:
## RDestimate(formula = recidivism ~ bac1 | bac1 * dui, data = hansen,
##     cutpoint = 0.08, bw = 0.05)
##
## Coefficients:
##      LATE    Half-BW   Double-BW
##   -0.01988   -0.01808    -0.02765
```

```
RDHonest(recidivism ~ bac1 + bac1*dui, cutoff = 0.08, h=0.05, kern="triangular", M=0.1, sclass="T", ord
```

```
## Call:
## RDHonest(formula = recidivism ~ bac1 + bac1 * dui, data = hansen,
##     cutoff = 0.08, M = 0.1, kern = "triangular", h = 0.05, sclass = "T",
##     order = 1)
##
##
## Inference by se.method:
##       Estimate Maximum Bias Std. Error
## nn -0.01987723 4.049288e-05 0.00471454
##
## Confidence intervals:
## nn    (-0.0291179, -0.01063656), (-0.02767245, Inf), (-Inf, -0.01208201)
##
## Bandwidth: 0.05
## Number of effective observations: 18911.67
```

```
#Binwidth 0.05: interact bac1 with cutoff linearly and as a quadratic
RDestimate(recidivism ~ bac1 | bac1*dui + (bac1^2)*dui, cutpoint = 0.08, bw = 0.05, data=hansen)
```

```
##
```

```
## Call:
## RDestimate(formula = recidivism ~ bac1 | bac1 * dui + (bac1^2) *
##     dui, data = hansen, cutpoint = 0.08, bw = 0.05)
##
## Coefficients:
##       LATE    Half-BW   Double-BW
##   -0.01988   -0.01808   -0.02765
```

RDHonest(recidivism ~ bac1 + bac1*dui + (bac1^2)*dui, cutoff = 0.08, h=0.05, kern="triangular", M=0.1,

```
## Call:
## RDHonest(formula = recidivism ~ bac1 + bac1 * dui + (bac1^2) *
##     dui, data = hansen, cutoff = 0.08, M = 0.1, kern = "triangular",
##     h = 0.05, sclass = "T", order = 1)
##
##
## Inference by se.method:
##       Estimate Maximum Bias Std. Error
## nn -0.01987723 4.049288e-05 0.00471454
##
## Confidence intervals:
## nn    (-0.0291179, -0.01063656), (-0.02767245, Inf), (-Inf, -0.01208201)
##
## Bandwidth: 0.05
## Number of effective observations: 18911.67
```

##--BINWIDTH 0.025--##
#Binwidth 0.025: control for bac1 linearly
RDestimate(recidivism ~ bac1, cutpoint = 0.08, bw = 0.025, data=hansen)

```
##
## Call:
## RDestimate(formula = recidivism ~ bac1, data = hansen, cutpoint = 0.08,
##     bw = 0.025)
##
## Coefficients:
##       LATE    Half-BW   Double-BW
##   -0.01808   -0.01650   -0.01988
```

RDHonest(recidivism ~ bac1, cutoff = 0.08, h=0.025, kern="triangular", M=0.1, sclass="T", order = 1, da

```
## Call:
## RDHonest(formula = recidivism ~ bac1, data = hansen, cutoff = 0.08,
##     M = 0.1, kern = "triangular", h = 0.025, sclass = "T", order = 1)
##
##
## Inference by se.method:
##     Estimate Maximum Bias  Std. Error
## nn -0.018083 1.117077e-05 0.006263225
##
## Confidence intervals:
## nn    (-0.03035872, -0.005807286), (-0.02839626, Inf), (-Inf, -0.007769742)
```

```
## 
## Bandwidth: 0.025
## Number of effective observations: 9811.295
```

```
#Binwidth 0.025: interact bac1 with cutoff linearly
RDestimate(recidivism ~ bac1 | bac1*dui, cutpoint = 0.08, bw = 0.025, data=hansen)
```

```
## 
## Call:
## RDestimate(formula = recidivism ~ bac1 | bac1 * dui, data = hansen,
##     cutpoint = 0.08, bw = 0.025)
## 
## Coefficients:
##      LATE    Half-BW   Double-BW
## -0.01808   -0.01650    -0.01988
```

```
RDHonest(recidivism ~ bac1 + bac1*dui, cutoff = 0.08, h=0.025, kern="triangular", M=0.1, sclass="T", or
```

```
## Call:
## RDHonest(formula = recidivism ~ bac1 + bac1 * dui, data = hansen,
##     cutoff = 0.08, M = 0.1, kern = "triangular", h = 0.025, sclass = "T",
##     order = 1)
## 
## 
## Inference by se.method:
##     Estimate Maximum Bias  Std. Error
## nn -0.018083 1.117077e-05 0.006263225
## 
## Confidence intervals:
## nn    (-0.03035872, -0.005807286), (-0.02839626, Inf), (-Inf, -0.007769742)
## 
## Bandwidth: 0.025
## Number of effective observations: 9811.295
```

```
#Binwidth 0.025: interact bac1 with cutoff linearly and as a quadratic
RDestimate(recidivism ~ bac1 | bac1*dui + (bac1^2)*dui, cutpoint = 0.08, bw = 0.025, data=hansen)
```

```
## 
## Call:
## RDestimate(formula = recidivism ~ bac1 | bac1 * dui + (bac1^2) *
##     dui, data = hansen, cutpoint = 0.08, bw = 0.025)
## 
## Coefficients:
##      LATE    Half-BW   Double-BW
## -0.01808   -0.01650    -0.01988
```

```
RDHonest(recidivism ~ bac1 + bac1*dui + (bac1^2)*dui, cutoff = 0.08, h=0.025, kern="triangular", M=0.1,
```

```
## Call:
## RDHonest(formula = recidivism ~ bac1 + bac1 * dui + (bac1^2) *
##     dui, data = hansen, cutoff = 0.08, M = 0.1, kern = "triangular",
```

```
##       h = 0.025, sclass = "T", order = 1)
##
##
## Inference by se.method:
##      Estimate Maximum Bias  Std. Error
## nn -0.018083 1.117077e-05 0.006263225
##
## Confidence intervals:
## nn    (-0.03035872, -0.005807286), (-0.02839626, Inf), (-Inf, -0.007769742)
##
## Bandwidth: 0.025
## Number of effective observations: 9811.295
```

```r
categories <- hansen$bac1

hansen3 <- hansen %>%
  filter(bac1 >= 0 & bac1 <= 0.15)

means_recid <- split(hansen3$recidivism, cut(hansen3$bac1, 75)) %>%
  lapply(mean) %>%
  unlist()
agg_hansen_recid <- data.frame(recidivism = means_recid, bac1 = seq(0.001, 0.15, by = 0.002))

hansen3 <- hansen3 %>%
  mutate(gg_group2 = case_when(bac1 > 0.08 ~ 1, TRUE ~ 0))

#Model "recidivism" RDD Using Linear Model and Quadratic models
lrecid <- ggplot(hansen3, aes(bac1, recidivism)) +
  geom_point(aes(x = bac1, y = recidivism), data = agg_hansen_recid) +
  geom_vline(xintercept = 0.08, colour = "grey", linetype = 2) +
  stat_smooth(aes(bac1, recidivism, group = gg_group2), method = "lm") +
  labs(x = "BAC", title = "BAC and Recidivism (quadratic)")
lrecid
```
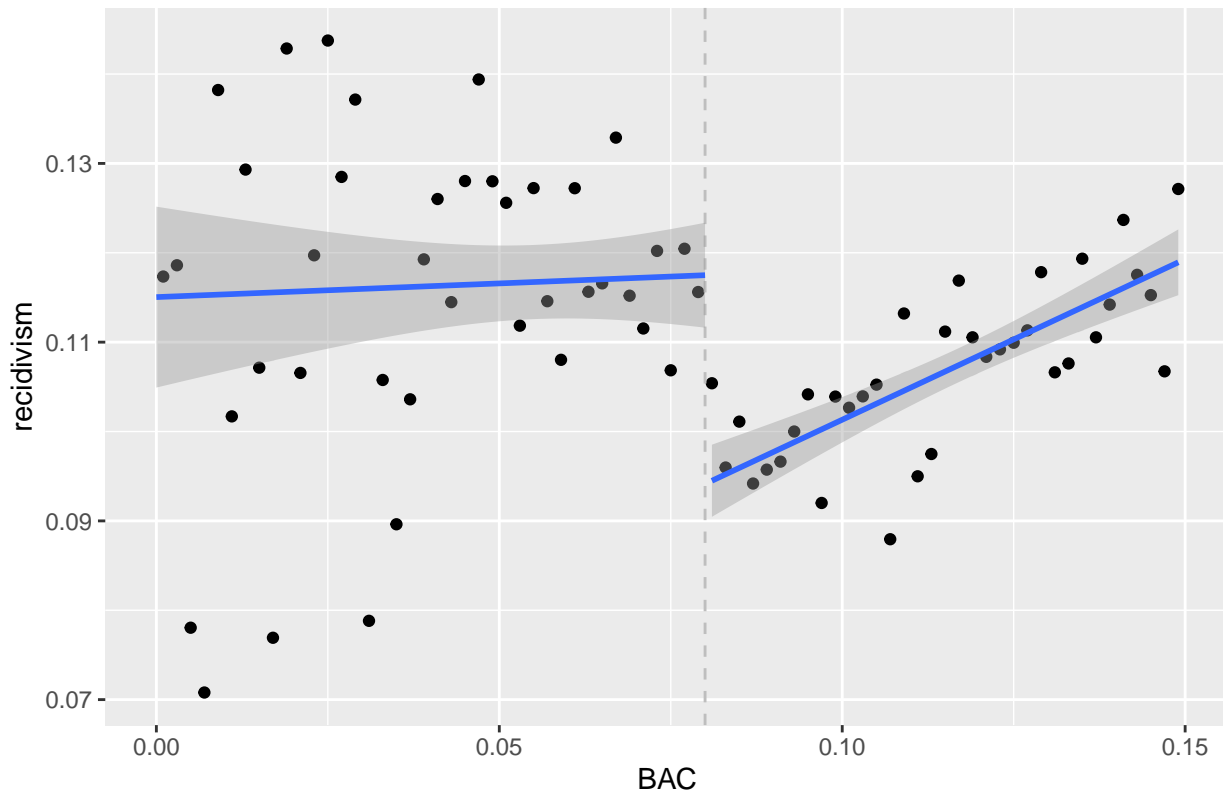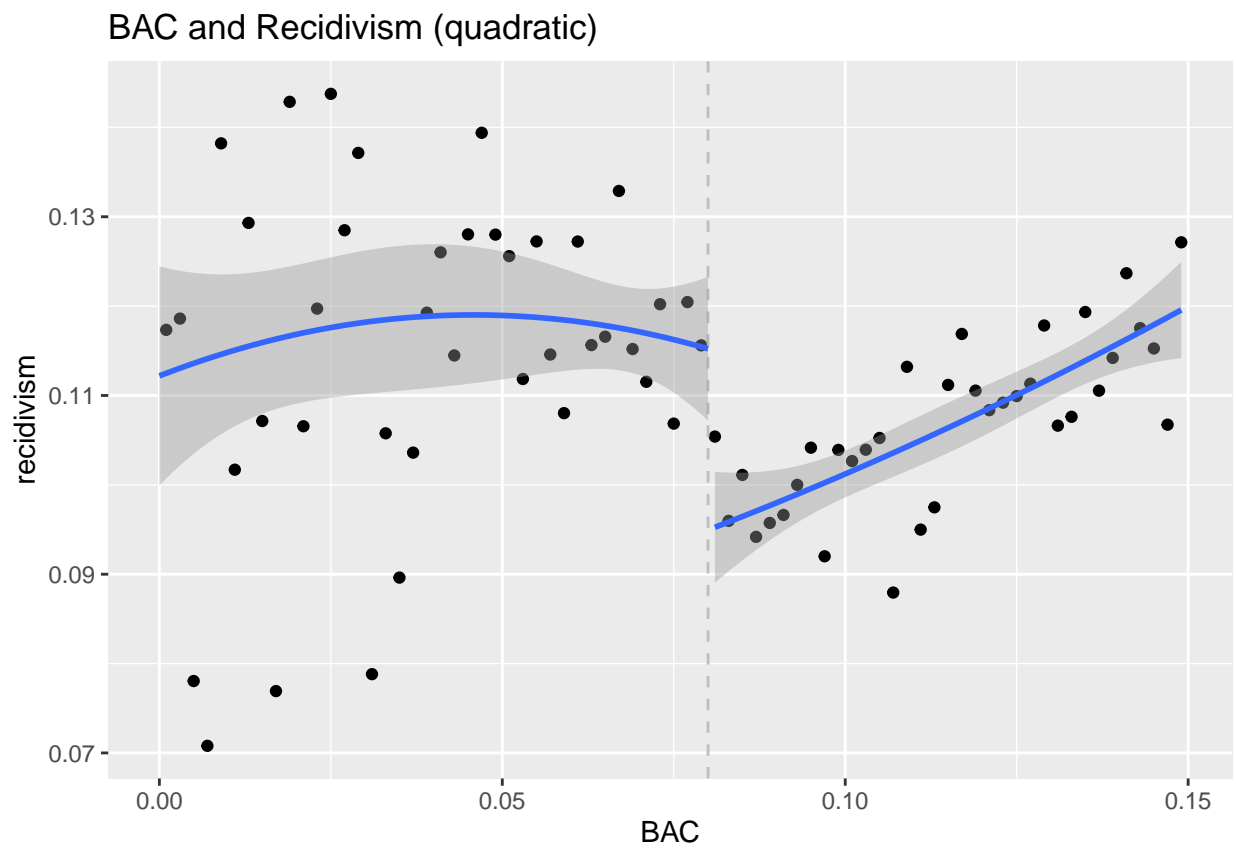

BAC and Recidivism (quadratic)

```r
qrecid<- ggplot(hansen3, aes(bac1, recidivism)) +
  geom_point(aes(x = bac1, y = recidivism), data = agg_hansen_recid) +
  geom_vline(xintercept = 0.08, colour = "grey", linetype = 2) +
```

```
    stat_smooth(aes(bac1, recidivism, group = gg_group2), method = "lm", formula = y ~ x + I(x^2)) +
    labs(x = "BAC", title = "BAC and Recidivism (quadratic)")
qrecid
```

BAC and Recidivism (quadratic)

#Nine [DISCUSS & COMPARE]