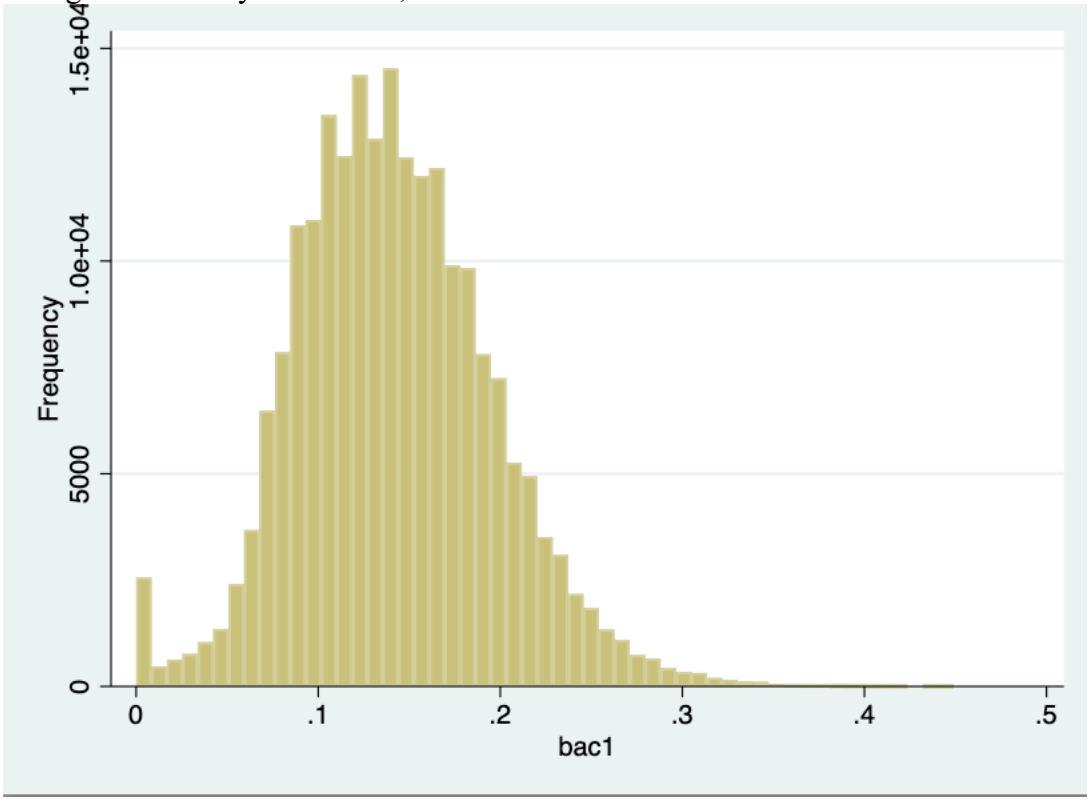
Question 1:
Repo on GitHub

Question 2:
Hansen proposes the research question that harsher punishments will have an effect (most likely to deter) those from driving under the influence more than once. He gathers data from the National Highway Traffic and Safety Administration specific to the state of Washington from 1995 to 2011. His research design is to utilize the data specifically 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 initial BAC test. The paper only includes those 21+. In Washington, a BAC test that registers 0.08 or above is a DUI while a BAC test that register 0.15 or higher is an aggravated DUI. He concludes that harsher punishments and sanctions associated with BAC limits reduce drunk driving further.

Question 3:
Creating the dummy variable

```
generate dumm1 = (bac1>= 0.08)
```

Question 4:
In order to test if there is manipulation of the data, we could check for continuity around the 0.08 BAC level. If it seems evenly distributed on either side of the benchmark, we know drivers can likely not manipulate their BAC levels.
Based on the graph created, there does not seem to be manipulation of the BAC variable. The histogram is evenly distributed, which is similar to what Hansen found.



Question 5: In this, Hansen used the LATE method, which stata does not have an obviously convenient method to utilize. So our results are drastically different, but nonetheless are in a similar direction of Hansens which is notable.

```
. reg male dui1 bac1 c.dui1##c.bac1
note: dui1 omitted because of collinearity
note: dui1 omitted because of collinearity
note: bac1 omitted because of collinearity
note: c.dui1#c.bac1 omitted because of collinearity
```

Source	SS	df	MS	Number of obs	=	214,558
				F(1, 214556)	=	6.87
Model	1.1412176	1	1.1412176	Prob > F	=	0.0088
Residual	35654.7754	214,556	.166179344	R-squared	=	0.0000
				Adj R-squared	=	0.0000
Total	35655.9166	214,557	.166183889	Root MSE	=	.40765

male	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
dui1	0	(omitted)				
bac1	-.0427876	.0163276	-2.62	0.009	-.0747893	-.0107859
dui1	0	(omitted)				
bac1	0	(omitted)				
c.dui1#c.bac1	0	(omitted)				
_cons	.7955759	.0024759	321.33	0.000	.7907233	.8004286

```
. reg white dui1 bac1 c.dui1##c.bac1
note: dui1 omitted because of collinearity
note: dui1 omitted because of collinearity
note: bac1 omitted because of collinearity
note: c.dui1#c.bac1 omitted because of collinearity
```

Source	SS	df	MS	Number of obs	=	214,558
				F(1, 214556)	=	173.54
Model	20.6789948	1	20.6789948	Prob > F	=	0.0000
Residual	25565.9555	214,556	.119157495	R-squared	=	0.0008
				Adj R-squared	=	0.0008
Total	25586.6345	214,557	.119253319	Root MSE	=	.34519

white	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
dui1	0	(omitted)				
bac1	.1821371	.0138259	13.17	0.000	.1550386	.2092356
dui1	0	(omitted)				
bac1	0	(omitted)				
c.dui1#c.bac1	0	(omitted)				
_cons	.8357748	.0020965	398.65	0.000	.8316656	.8398839

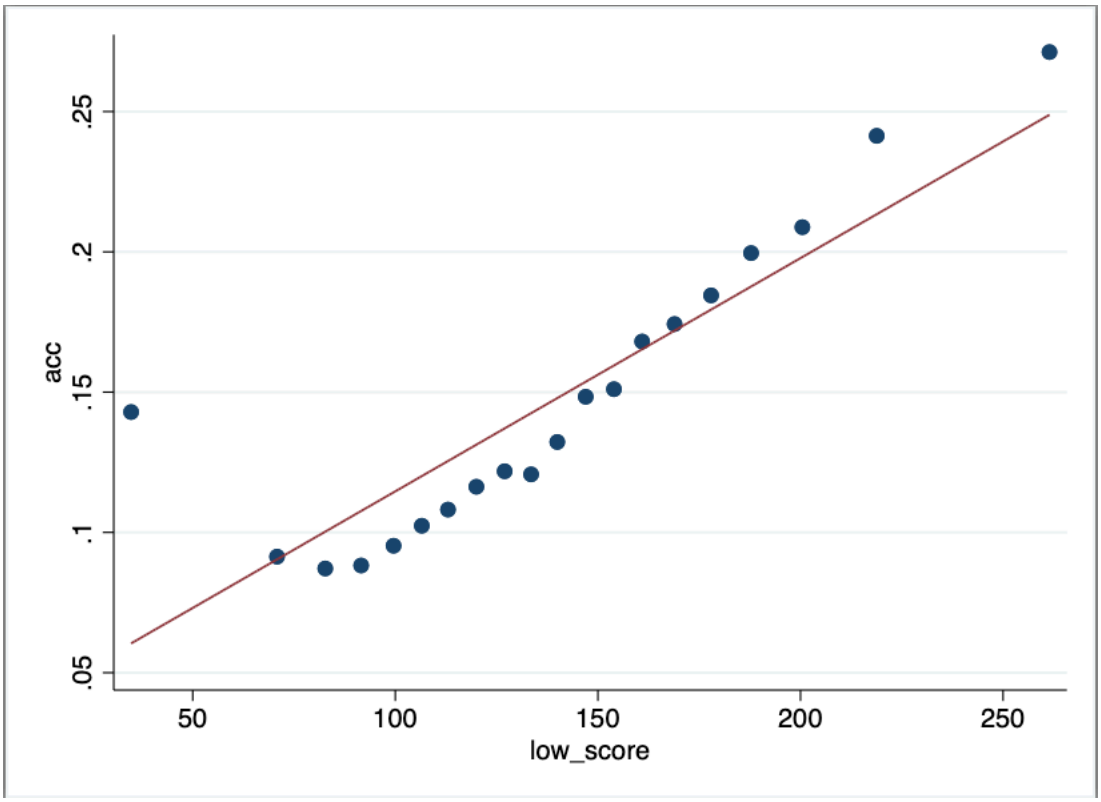
```
. reg acc dui1 bac1 c.dui1##c.bac1
note: dui1 omitted because of collinearity
note: dui1 omitted because of collinearity
note: bac1 omitted because of collinearity
note: c.dui1#c.bac1 omitted because of collinearity
```

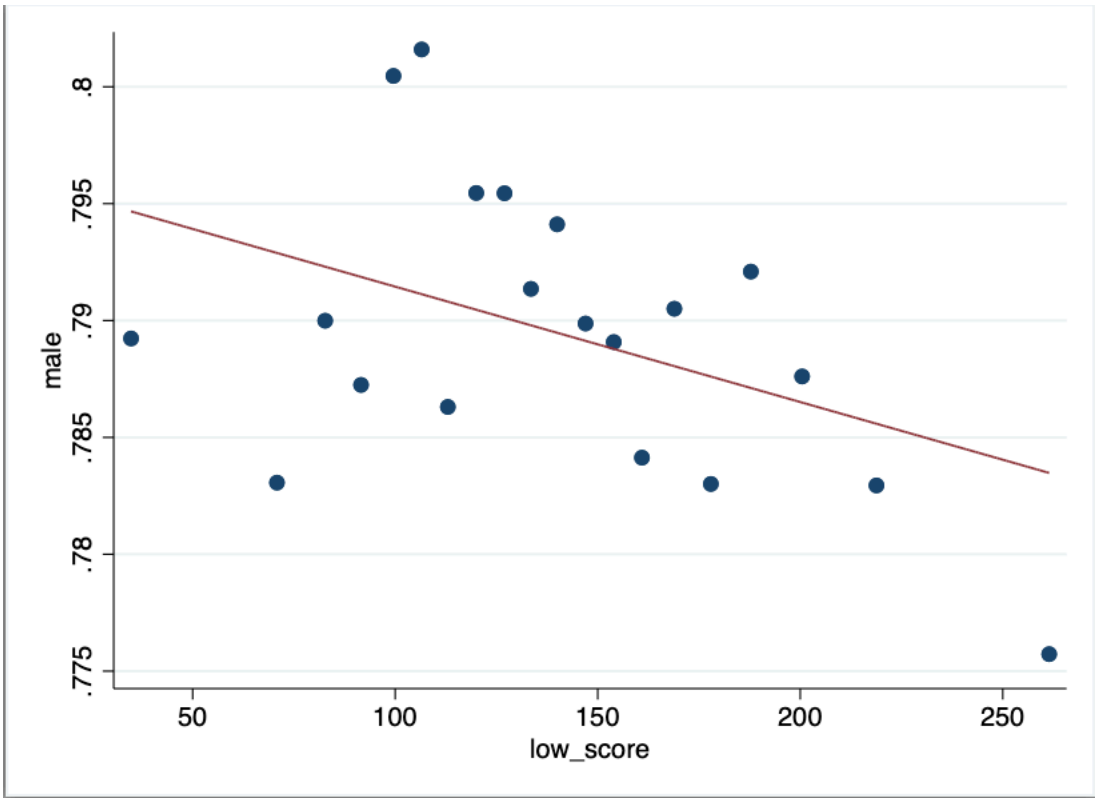
Source	SS	df	MS	Number of obs	=	214,558
				F(1, 214556)	=	3415.25
Model	422.23161	1	422.23161	Prob > F	=	0.0000
Residual	26525.8517	214,556	.123631368	R-squared	=	0.0157
				Adj R-squared	=	0.0157
Total	26948.0833	214,557	.125598714	Root MSE	=	.35161

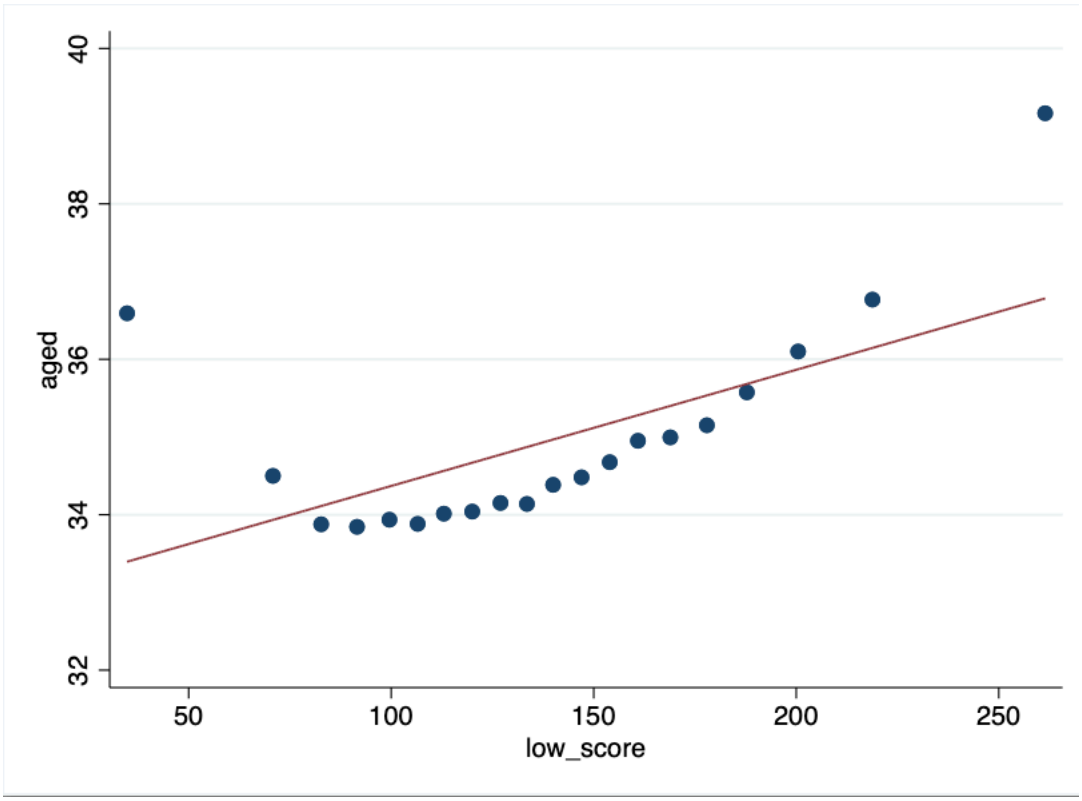
acc	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
dui1	0	(omitted)				
bac1	.8230175	.0140831	58.44	0.000	.795415	.85062
dui1	0	(omitted)				
bac1	0	(omitted)				
c.dui1#c.bac1	0	(omitted)				
_cons	.0306435	.0021355	14.35	0.000	.026458	.0348291

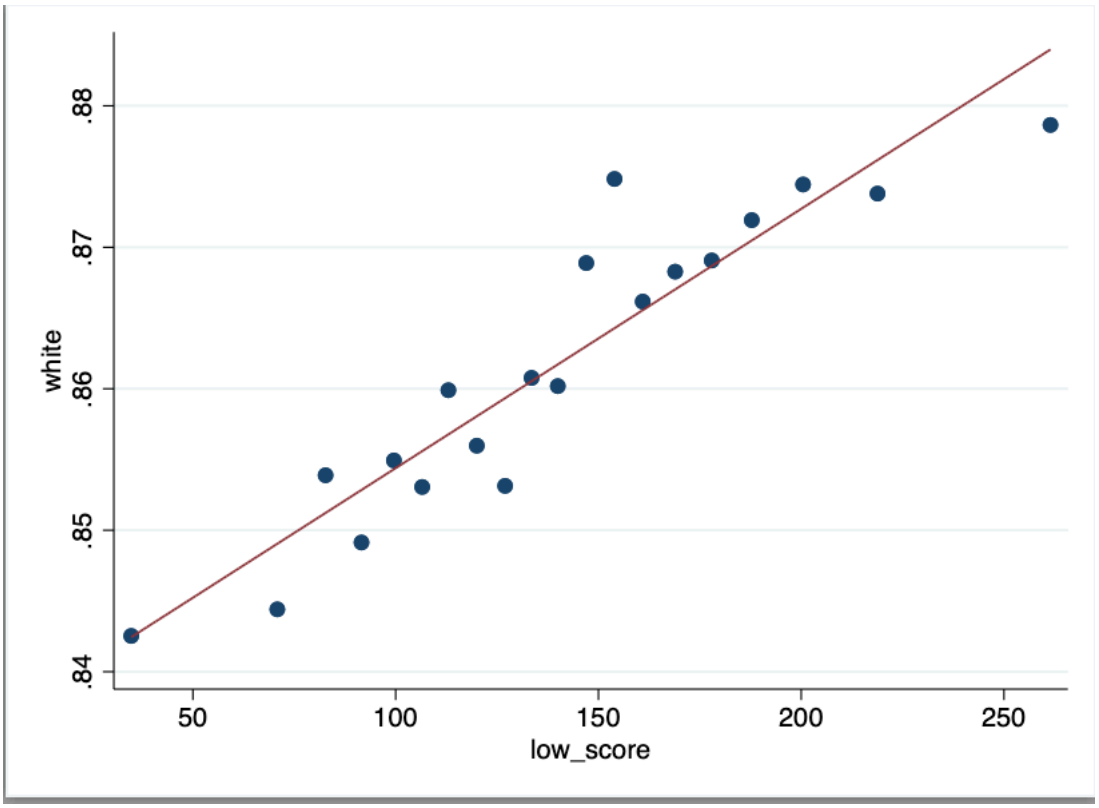
Question 6:

I found that the higher the initial BAC, the higher the likelihood there was an accident. This is similar to Hansens results. Once again Hansen was able to manipulate his results around the DUI cutoffs of 0.08 and 0.15 whereas I had a difficult time in doing so. I was only able to fit one line, but the results somewhat mirror Hansens results.




male
.775
.78
.785
.79
.795
.8
50
100
150
200
250
low_score


aged
32
34
36
38
40
50
100
150
200
250
low_score



Question 7:
Hansen's results obviously produced a table whereas mine are coming in a printed STATA output. The estout command was being fussy. But, I believe that Hansen used LATE methods whereas I did not, which is the reason the results are different.
Row 1, Column 1

recidivism	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
bac1	-.0545834	.	.	.	.	.
0.dui1	0	(omitted)				
dui1#c.bac1						
0	0	(omitted)				
_cons	.1123115	.	.	.	.	.

Row 1, Column 2

recidivism	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
bac1	-.0545834	.	.	.	.	.
dui1	0	(omitted)				
c.bac1#c.dui1	0	(omitted)				
_cons	.1123115	.	.	.	.	.

recidivism	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
bac4	8.326036	.	.	.	.	.
dui1	0	(omitted)				
c.bac4#c.dui1	0	(omitted)				
bac1	-1.575768	.	.	.	.	.
dui1	0	(omitted)				
c.bac1#c.dui1	0	(omitted)				
_cons	.1773877	.	.	.	.	.

0.055 to 0.105

Row 2, Column 1

recidivism	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
bac1	-.4417888	.	.	.	.	.
_cons	.1433839	.	.	.	.	.

Row 2, Column 2

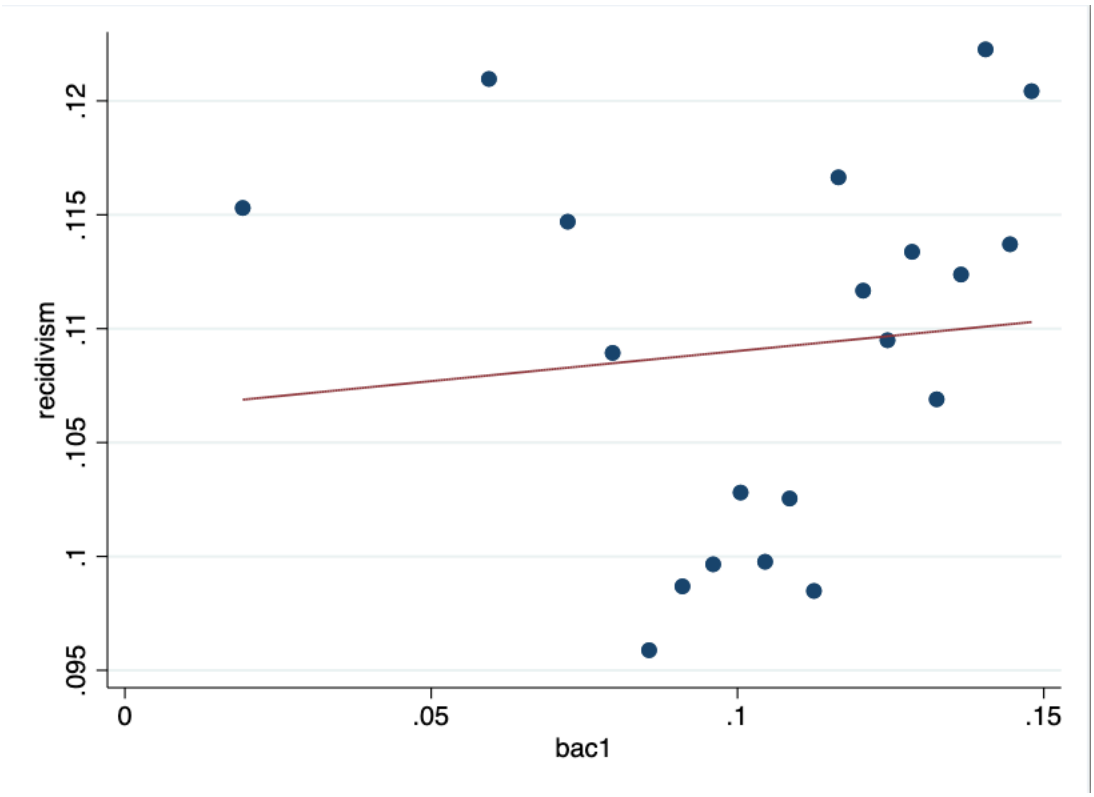
recidivism	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
bac1	-.2966053	.	.	.	.	.
dui1	0	(omitted)				
c.bac1#c.dui1	0	(omitted)				
_cons	.1306692	.	.	.	.	.

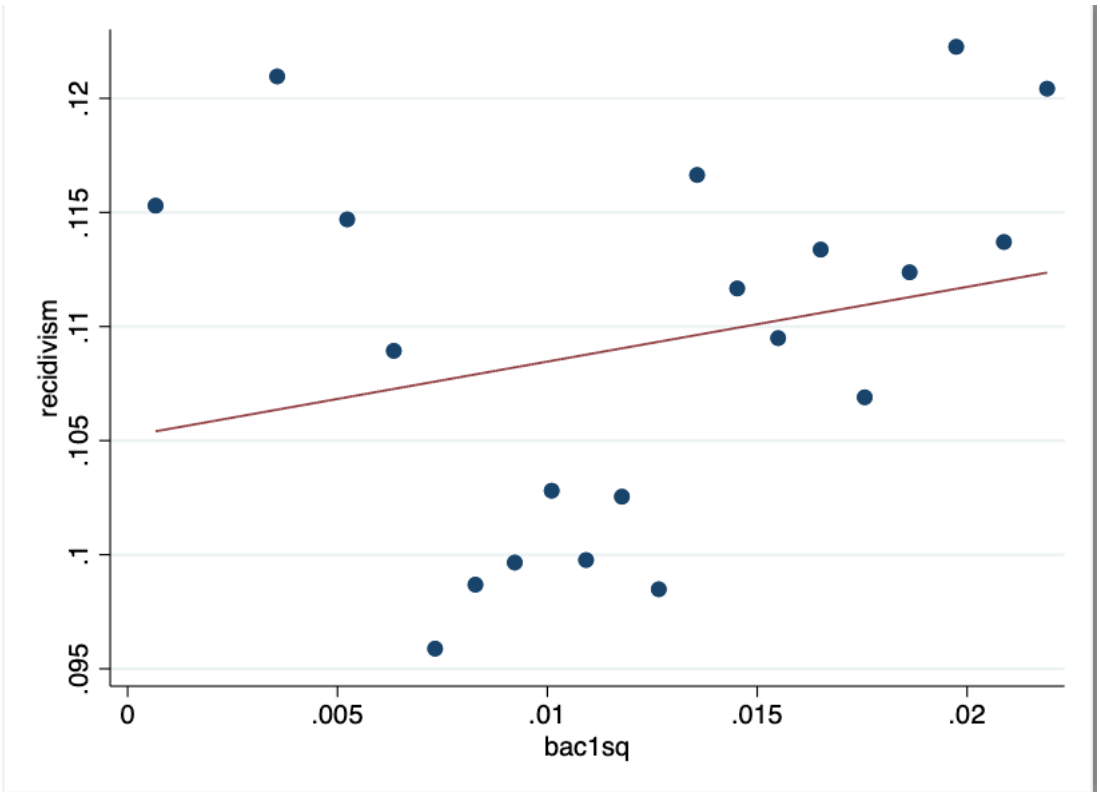
Row 2, column 3

recidivism	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
bac1	.2597717	.	.	.	.	.
dui1	0	(omitted)				
c.bac1#c.dui1	0	(omitted)				
bac4	-4.079976	.	.	.	.	.
dui1	0	(omitted)				
c.bac4#c.dui1	0	(omitted)				
_cons	.114023	.	.	.	.	.

Question 8:

Hansen once again was able to produce a graph with a much better result, allowing for the estimation to differ around the DUI cutoff of 0.08 whereas I had a difficult time allowing for this result. Nonetheless, we can see that our results are somewhat similar to Hansens in nature with the 0.1 to 0.15 range respectively.


recidivism
.095
.1
.105
.11
.115
.12
0
.05
.1
.15
bac1


recidivism
.095
.1
.105
.11
.115
.12
0
.005
.01
.015
.02
bac1sq

Question 9:

We were testing the hypothesis that harsher punishment deters from recidivism. I believe that this hypothesis can be concluded and Hansens' decision is correct. He had a more precise method than mine and yet I noticed these results as well.
I think I learned from this exercise that RDD replications can be difficutl in both STATA and R and it may be useful to utilize both.