Matthew Borelli

Replication 1

Causal Inference

Dr. Scott Cunningham

**Introduction**

Hansen’s project, “Punishment and Deterrence: Evidence from Drunk Driving”, is a study that focuses on drunk driving accidents in the United States. Particularly, Hansen wanted to determine whether harsher punishments and sanctions for DUI convictions cause recidivism rates to drop. The operating theory behind this question comes from a work by Becker (1968), which made the case that criminals are rational agents and commit crimes by weighing expected costs and benefits. Under this model, increasing punishments should lower the amount of crime, which has been backed up in subsequent research by various authors.

This project used administrative records from the Washington (WA), constituting 512,964 DUI stops over the period from 1995 to 2011. Starting on January 1, 1999, WA set the following blood alcohol content (BAC) thresholds for charges: 0.08 for a DUI, and 0.15 for an aggravated DUI. Hansen uses the arrest data from 1999 to 2007. To identify recidivism for suspected offenders, the next four years from their test date are used. To determine the causal effect of these punishments, Hansen utilizes the regression discontinuity method. There is almost no distinguishable difference between a 0.079 and a 0.08 BAC other than the difference in punishments, which means these thresholds make sense as possible discontinuities. In fact, the assumption is made that drivers will randomly fall on either side of the threshold when they are close, with some drivers being luckier than others and getting less punishment.

Hansen found “evidence that having a BAC above either the 0.08 DUI threshold or the 0.15 aggravated DUI is associated with reduced repeat drunk driving both in the short and long term” (1607). Having a BAC above the 0.08 threshold is associated with a 2% decrease in recidivism and having a BAC above the 0.15 threshold is associated with an additional 1% decrease in recidivism. Hansen also estimates the elasticity of sanction amounts: A 10% increase in the sanction for a DUI (fine, sentence length, etc.) is correlated with a 2.3% decrease in drunk driving.

**Sorting on Running Variable**

Before utilizing regression discontinuity, we want to analyze the running variable (bac1) for potential manipulation. If people were able to manipulate their BAC levels we would expect bunching just below the threshold amounts of 0.08 and 0.15 in order to avoid the next level of punishment. Looking at Figure 1, we do not see any evidence of bunching around either of the thresholds. There seems to be a continuous rise and fall that peaks at around 0.13-0.14, without extreme spikes right below 0.08 and 0.15. Thus, we can say that it does not seem like people can manipulate their BAC, which allows us to use regression discontinuity.

Figure 1: Histogram of Blood Alcohol Content

**Covariate Balance**

The next thing that we want to check is whether we have covariate balance. To do this, we estimate regressions with our covariates as the dependent variables. If the we have covariate balance, we would expect to see that the coefficients for DUI should be statistically insignificant. In following Hansen’s methods, these models are estimated using the sample of data in a 0.05 interval from the cutoff point (0.03, 0.13) to make sure the results aren’t biased by outliers. Table 1 shows the four models used for this analysis, utilizing the covariates white, male, age, and accident.

Table 1: Regression results used to check covariance balance

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
| VARIABLES | White | Male | Age | Accident |
|  |  |  |  |  |
| bac1 | 0.0788 | -0.210 | -69.16\*\*\* | -1.096\*\*\* |
|  | (0.224) | (0.234) | (6.862) | (0.176) |
| DUI | 0.00445 | -0.0184 | -6.224\*\*\* | -0.154\*\*\* |
|  | (0.0173) | (0.0198) | (0.585) | (0.0158) |
| bac1\_DUI | 0.0156 | 0.307 | 76.05\*\*\* | 1.888\*\*\* |
|  | (0.238) | (0.263) | (7.647) | (0.203) |
| Constant | 0.840\*\*\* | 0.801\*\*\* | 39.45\*\*\* | 0.171\*\*\* |
|  | (0.0150) | (0.0151) | (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

The coefficient for DUI is significant in 2 of the 4 models, the ones with age and accident indicator as the dependent variables. This means that the indicators for white and male are covariate balanced, and the age and accident indicator variables are not covariate balanced. This is different than Hansen’s findings, which showed that all four of these were covariate balanced. I’m not exactly sure why there’s a difference, but one possibility is that this replication was not done with the same data, thus differences like this could occur. Figures 2A and 2B are graphical representations of their potential discontinuity at the DUI threshold. All the fitted lines were generated using data where the BAC level was less than 0.20, the level than Hansen used in his version of these graphs. For this exercise, there are two different fitted lines for each control variable: Figure 2A shows the linear fit models and figure 2B shows the quadratic fit models.

****

Figure 2A: BAC and Characteristics for Linear Fit Models

****

Figure 2B: BAC and Characteristics for Quadratic Fit Models

In comparing Figure 2A to Hansen’s figures in his study, the slopes are all fairly similar to what Hansen found. Even though Hansen ended up with all of his variables as covariate balanced, a fact which was not replicated in this analysis, he did estimate a large slope difference for accidents and age at the threshold point, and little difference in slope for male and white. The more apparent differences in this analysis come in Figure 2B as the quadratic fit seems to show that all four of the variables are covariate balanced. Hansen did not do a quadratic model to check for covariate balance so no direct comparison could be made, but it should be noted that using the quadratic form of BAC made obvious improvements over the linear fit.

**Results**

For this replication, there will be two panels for the regression discontinuity design: One for BAC levels between 0.3 and 0.13 (Panel A) and the other for BAC levels between 0.055 and 0.105 (Panel B). Hansen used panels with differing bandwidths in order to compare results for people close to the threshold with people that might be further away from it. Table 3 displays the regression discontinuity results.

Table 3 – Panel A Regression Discontinuity Estimates for the Effect of Exceeding the 0.08 BAC Threshold on Recidivism

|  |  |  |  |
| --- | --- | --- | --- |
|  | (1) | (2) | (3) |
| VARIABLES | Model 1 | Model 2 | Model 3 |
|  |  |  |  |
| DUI | -0.0271\*\*\* | -0.0576\*\*\* | 0.116 |
|  | (0.00403) | (0.0152) | (0.0843) |
| bac1 | 0.312\*\*\* | -0.0367 | 2.943\* |
|  | (0.0748) | (0.187) | (1.638) |
| bac1\_DUI |  | 0.420\*\* | -4.276\*\* |
|  |  | (0.204) | (2.112) |
| bac1sq |  |  | -25.01\* |
|  |  |  | (13.74) |
| bac1sq\_DUI |  |  | 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 |

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

For Panel A, there are three models: Model 1 has the dummy variable for DUI and the BAC level with controls, Model 2 adds an interaction between DUI and BAC, and Model 3 adds in quadratic terms for the BAC and BAC/DUI interaction term. It is important to note that these controls are slightly different than Hansen’s, primarily due to having no county or state fixed effects and the addition of the accident control. In Model 1 and Model 2, the DUI dummy is significant, meaning that there is a significant discontinuity in recidivism at the 0.08 BAC threshold. However, after controlling for quadratic effects of BAC, there is no longer a significant discontinuity. The linear and quadratic interactions are significant however, which means there is still evidence of a different relationship between BAC and DUI on recidivism for either side of the 0.08 BAC threshold.

Table 3 -- Panel B -- Regression Discontinuity Estimates for the Effect of Exceeding the 0.08 BAC Threshold on Recidivism

|  |  |  |  |
| --- | --- | --- | --- |
|  | (1) | (2) | (3) |
| VARIABLES | Model 1 | Model 2 | Model 3 |
|  |  |  |  |
| DUI | -0.0217\*\*\* | -0.0628\* | 0.379 |
|  | (0.00558) | (0.0350) | (0.422) |
| bac1 | 0.177 | -0.194 | 6.124 |
|  | (0.201) | (0.382) | (8.113) |
| bac1\_DUI |  | 0.530 | -10.66 |
|  |  | (0.449) | (10.60) |
| bac1sq |  |  | -45.74 |
|  |  |  | (58.70) |
| bac1sq\_DUI |  |  | 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 |

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

For BAC levels closer to the threshold, there is no significant effect on recidivism for either the DUI dummy variable or the BAC level after fitting a quadratic model. The only significant independent variables in model 3 are some of the control variables. This starkly contrasts with Hansen’s findings of a significant effect of punishments on recidivism rates, as Model 3 shows that punishments associated with the 0.08 BAC level have no significant impact on recidivism rates.



Figure 3: BAC and Recidivism

The last replication is of Figure 3 from Hansen’s paper, a scatterplot of the regression discontinuity. The dotted red line is at 0.08, the BAC threshold for DUI. In the simple models presented here, there is a significant discontinuity at the threshold for both the linear fit and the quadratic fit. In both instances, the recidivism rate drops right after the threshold, which holds with Hansen’s findings that punishments are correlated with a significant decrease in recidivism. Estimating from the graph, there seems to be approximately a 3% decrease in recidivism rate when the BAC at arrest moves from below the threshold to above the threshold. Another important point to note is that the relationship between BAC and recidivism is different above the threshold. Below the threshold there seems to be little to no correlation between BAC and recidivism, but above the threshold there is a positive relationship between BAC and recidivism. This means that the further an offender was above the DUI threshold, the more likely they are to repeat offend.

**Conclusion**

This replication of Hansen’s does support his findings, albeit with weaker evidence in favor of his conclusions. In his project he found that all of the controls were covariate balanced, but in this replication we found that some were unbalanced. More importantly, the regression discontinuity models in this replication are mixed on whether harsher punishments are significantly correlated with recidivism rates for drunk driving. Using the linear models that Hansen used gives significant coefficients for the DUI threshold BAC but fitting the quadratic form of BAC level onto recidivism completely removes the statistical significance of the threshold. Overall, it seems that simple regression discontinuity models show that punishments are correlated with a sharp decrease in recidivism, but more complex models like quadratic fits do not show this correlation.