Candidate number: 49045

Replication of Hansen (2015) Punishment and Deterrence: Evidence from Drunk Driving

Submitted as the summative assessment for
PB4A7: Quantitative Applications for Behavioural Science 2022

In the study *Punishment and Deterrence: Evidence from Drunk Driving* **(Hansen, 2015)**, Hansen investigated the effect of punishments and sanctions on reducing repeat drunk driving. The study implemented a quasi-experimental design, utilising the administrative records from 1995 to 2011 of the state of Washington, U.S., where two thresholds of blood alcohol content (BAC) are used to determine the status of driving under the influence (DUI). Specifically, a driver with a measured BAC over 0.08 is considered a case of DUI and will be punished via measures such as fines, jail time, and driving license suspension. One with a BAC over 0.15 is considered a case of aggravated DUI, to which more severe punishments and sanctions are applied. Since BAC measure has clear-cut numeric thresholds for determining whether a driver will receive harsher punishments, and neither drivers or police can manipulate this measure, Hansen applied a regression discontinuity design to analyse the data **(expand on the justification of using RDD)**. He hypothesized that receiving harsher punishments and sanctions at both thresholds would reduce offenders' future recidivism of DUI. He further applied the RDD to analyse the effect of receiving harsher punishments on the degree of deterrence, incapacitation, and rehabilitation in order to identify the mechanism through which harsher punishments might reduce recidivism **(check the paper to see whether this sentence is precise)**.

Ideally also the results

The present study aims to replicate the findings from **Hansen (2015)** using regression discontinuity design applied to a similar data **(what is it exactly…)**. It will be focused on estimating the effect of receiving punishments on reducing recidivism at the 0.08 BAC threshold only. The paper proceeds as follows. Section 1 discusses the econometric method and assumptions underlying its application to the current data. Section 2 presents the main results, and Section 3 discusses critiques and extensions to the original study. Section 4 concludes.

1   Methods and Assumptions

1.1   Assumptions of the regression discontinuity design

The present study applies the regression discontinuity design to the data provided by the class instructor. Several assumptions need to be met so that the regression discontinuity design can give accurate estimates. First, people need to be randomly assigned to receiving punishment at the BAC threshold. In other words, neither the drivers nor the police can manipulate the BAC level and thus whether one will receive punishment. **Hansen (2015)** has already given a plausible theoretical justification for this assumption. Empirically, I plot the distribution of BAC and test for discontinuity in its distribution at the threshold. Figure 1 shows the distribution of BAC level. Based on a recently developed local polynomial density estimator **(CJM, 2020)**, the hypothesis that the distribution is continuous at 0.08 cannot be rejected at the alpha level of 0.05 ($p = 0.890$). Therefore, there is no evidence for the existence of manipulation on the BAC level.

The second assumption is that the running variable should not contain non-random heaping, which can lead to biased estimates in regression discontinuity models **(Barreca et al., 2011)**. In other words, BAC should not be much more likely to take certain values than others. Based on figure 1, this assumption seems to be violated. Curiously, when BAC is plotted as a discrete variable, non-random heaping disappears (Figure
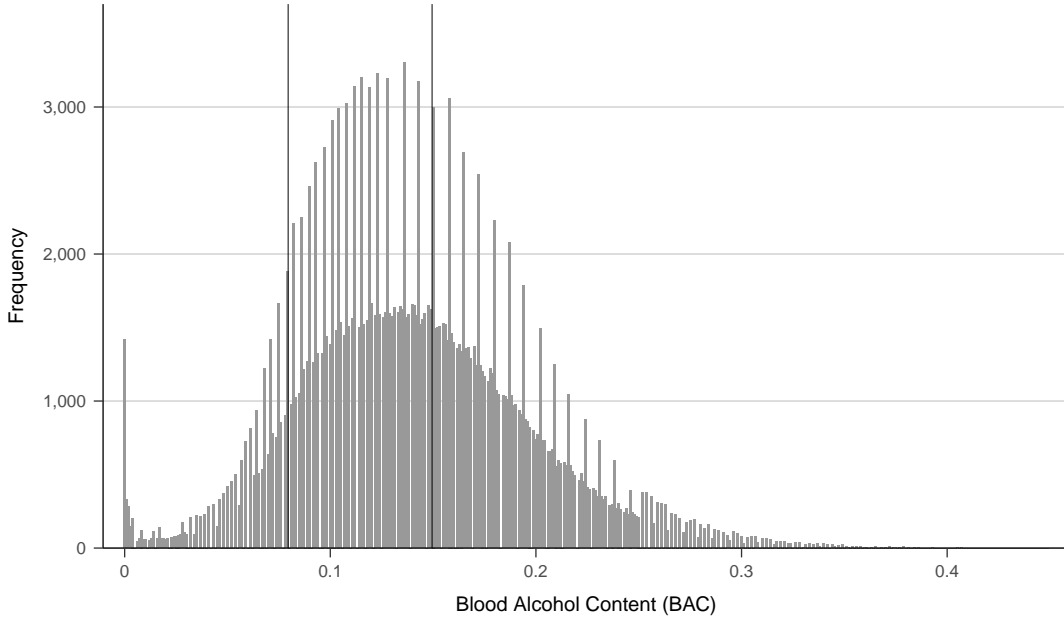
Figure 1: Histogram of blood alcohol content as a continuous variable. The blood alcohol content is plotted as a continuous variable with a bin width of 0.001, the precision used on the breathalysers. The y-axis represents the frequency of observations in each bin. The vertical black lines represent the two thresholds at 0.08 and 0.15.

2). A closer inspection of the data suggests that this is probably due to the lack of precision in BAC in the current data. Many BAC values deviate by a very small amount from the values that are supposed to be given by the breathalysers (i.e., precise to three digits after the decimal point). Some bins have a value at the left boundary deviating upwards and a value at the right boundary deviating downwards. Such bins will contain a larger number of observations than other bins and thus leads to heaping. This would not occur if the data records the values precisely so that the number of observations are rightly separated into and plotted by two bins, or if we treat BAC as discrete and give each value its own bin. **(consider using a simulation to justify this claim?)**

1.2 The model

The present study utilises a local linear regression model with a triangular kernel weight function to estimate the effect of receiving punishments at the threshold on recidivism. For sensitivity analyses, the models are re-estimated using second-order polynomials and other kernel functions (universal and rectangular). The model is specified by the following function:

$$R_i = \alpha + \beta_1 DUI_i + \beta_2 BAC_i + \beta_3 DUI_i \times BAC_i + \epsilon_i \tag{1}$$

where the variable DUI_i is an indicator of whether BAC is above the 0.08 threshold, BAC_i is the measure of BAC level, and R_i is an indicator of recidivism. The variable BAC_i in the model is centered around the
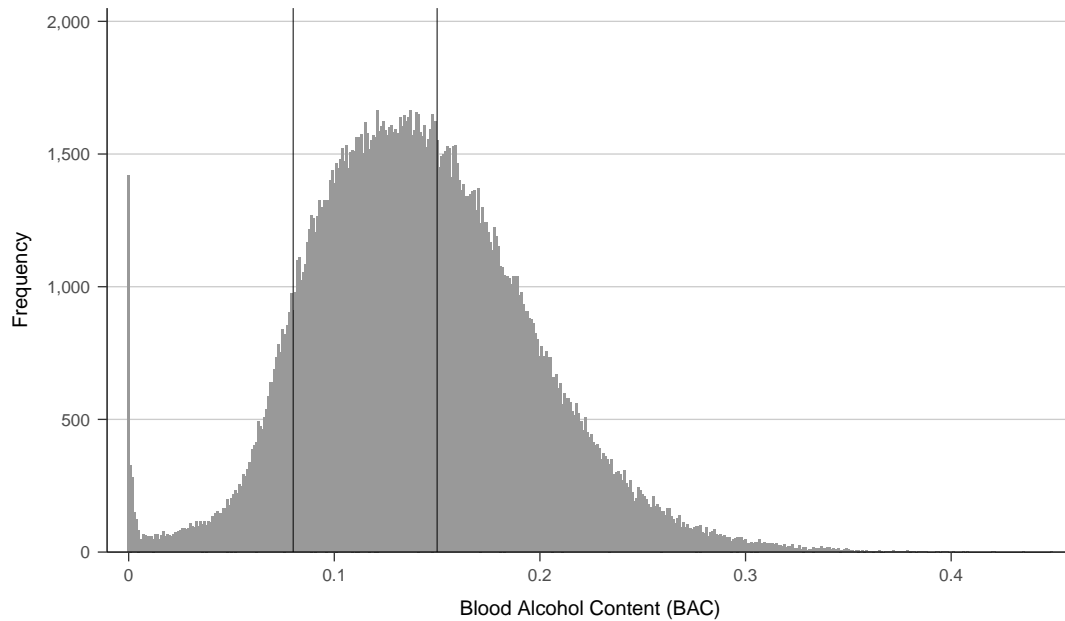
2

Figure 2: Histogram of blood alcohol content as a discrete variable. The blood alcohol content is plotted as a discrete variable with a bin width of 0.001, the precision used on the breathalysers. The y-axis represents the frequency of observations in each bin. The vertical black lines represent the two thresholds at 0.08 and 0.15.

threshold value of 0.08.

1.3 Inclusion of covariates

In order to determine whether any covariates should be included in the model, I run preliminary analyses on the effect of receiving punishments at the threshold on four predetermined characteristics, including three demographic variables (gender, race, and age) and the BAC test being conducted in a traffic accident. The analyses use the same local linear regression model specified in equation (1), with a bandwidth of 0.05 and no kernel weight function. As Table 1 shows, I fail to reject any of the null hypotheses that predetermined characteristics remained the same at the threshold. This indicates that gender, race, age, and the BAC test being conducted in a traffic accident, on average, were not different between people who did not receive punishments and those who received punishments at the threshold. According to **Calonico et al. (2019)**, I will include the four predetermined characteristics in regression discontinuity models to improve the estimation precision of the effects of receiving punishments at the threshold on recidivism.

2   Results

3   Critique and Extension

4   Conclusion

Table 1: Regression Discontinuity Estimates of the Effect of Receiving Punishments at the Threshold on Predetermined Characteristics

	Male	White	Age	Accident
DUI	0.006	0.006	−0.141	−0.003
	(0.006)	(0.005)	(0.164)	(0.004)
Mean (at 0.079)	0.768	0.852	−0.003	28.5
Numb. of Obs.	89,967	89,967	89,967	89,967

Note. Regression discontinuity based estimates of the effect of receiving punishments at the threshold on four predetermined characteristics. All models use a bandwidth of 0.05 and a universal kernel weight function. Heteroscedasticity-robust standard errors are in parentheses. $^{*} p < 0.1$, $^{**} p < 0.05$, $^{***} p < 0.01$.