Lecture Notes: Causal Inference Fall 2020

Claire Duquennois

7/28/2020

Contents

1 Regression Discontinuity

Regression discontinuity research designs were introduced in other fields as far back as the 1960s but only gained popularity in economics in the past 20 years or so as economist became increasingly focused on causal inference and large administrative datasets became more widely available. This increased popularity is well deserved. When correctly applied, presented and explained, RD designs are very transparent in how they achieve causal identification which makes them very appealing.

RD designs tend to leverage the researchers knowledge of a rule or policy that determines treatment. Identification is then based on the idea that how some rules are applied can be quite arbitrary and that this arbitrary application generates the randomness we crave for the identification of causal effects.

Suppose that we want to estimate the effect of some binary treatment D_i on an outcome Y_i. Using the potential outcomes framework, we write $Y_i(0)$ as the potential untreated outcome and $Y_i(1)$ as the potential treated outcome, with $Y_i = D_iY_i(1) + (1 - D_i)Y_i(0)$. Now suppose that the value of D_i-i.e. whether or not an individual gets treated- is completely (or partially) determined by whether some predictor X_i lies above or below a certain threshold, c. The predictor X_i need not be randomly assigned. In fact, we assume that it is related to the potential outcomes $Y_i(0)$ and $Y_i(1)$, but that this relationship is smooth, i.e. $Y_i(0)$ and $Y_i(1)$ do not jump discontinuously as X_i changes. Any discontinuous change in Y_i as X_i crosses c can thus be interpreted as a causal effect of D_i. We call X_i the "running variable".

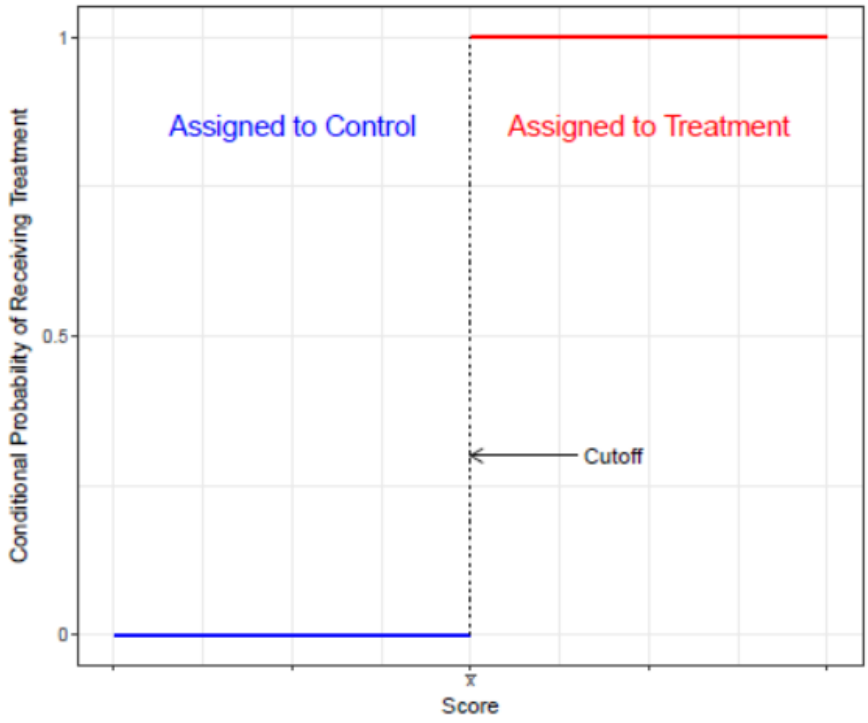
RD designs often arise in administrative situations in which units are assigned a program, treatment or award based on a numerical index being above or below a certain threshold. For example, a politician may be elected if and only if the differential between the vote share that she receives and the vote share that her opponent receives exceeds 0. A student may be assigned to summer school if and only if his performance on a combination of tests falls below a certain threshold. A toxic waste site may receive cleanup funds if and only if it's hazard rating falls above a certain level. In these cases , individuals or units whose indices X lie directly below the threshold c are considered to be comparable to individuals or units whose indices X lie directly above the threshold c and we can estimate the treatment effect by taking a difference in mean outcomes for units directly above the threshold and units directly below the threshold.

RD designs come in two flavors: Sharp and fuzzy. I will start by looking at sharp RD's and then discuss how things change when things get fuzzy.

1.1 Sharp RD Designs

In a sharp RD design, the probability that $D = 1$ changes from 0 to 1 as the running variable crosses c. In other words, no one with $X < c$ gets treated and everyone with $X \geq c$ gets treated, thus, D_i is a deterministic function of $X_i : D_i = 1$ if $(X_i \geq c)$. (Note: Here we assume that high values of the running variable get treated. For some RD set ups, the reverse will be true.)

Figure 2.1: Conditional Probability of Receiving Treatment in Sharp RD Design



In the image above we see that that the probability of getting treated switches from 0 to 1 when an observation's running variable X crosses the threshold c.

To estimate the causal effect of D_i on some outcome Y_i, we simply take the difference in mean outcome on either side of c. Formally, we estimate

$$lim_{x \to c} E[Y_i | X_i = x] - lim_{x \leftarrow c} E[Y_i | X_i = x] = lim_{x \to c} E[Y_i(1) | X_i = x] - lim_{x \leftarrow c} E[Y_i(0) | X_i = x]$$

This represents the causal effect of D on some outcome Y for individuals with $X_i = c$. We will call this effect τ_{SRD}.

$$\tau_{SRD} = E[Y_i(1) - Y_i(0) | X_i = c]$$

To justify this interpretation, we need it to be the case that $Y_i(0)$ and $Y_i(1)$ are smooth functions of X_i as X_i crosses c. We make this assumption in the form of a conditional expectation:

Assumption 1: The continuity assumption

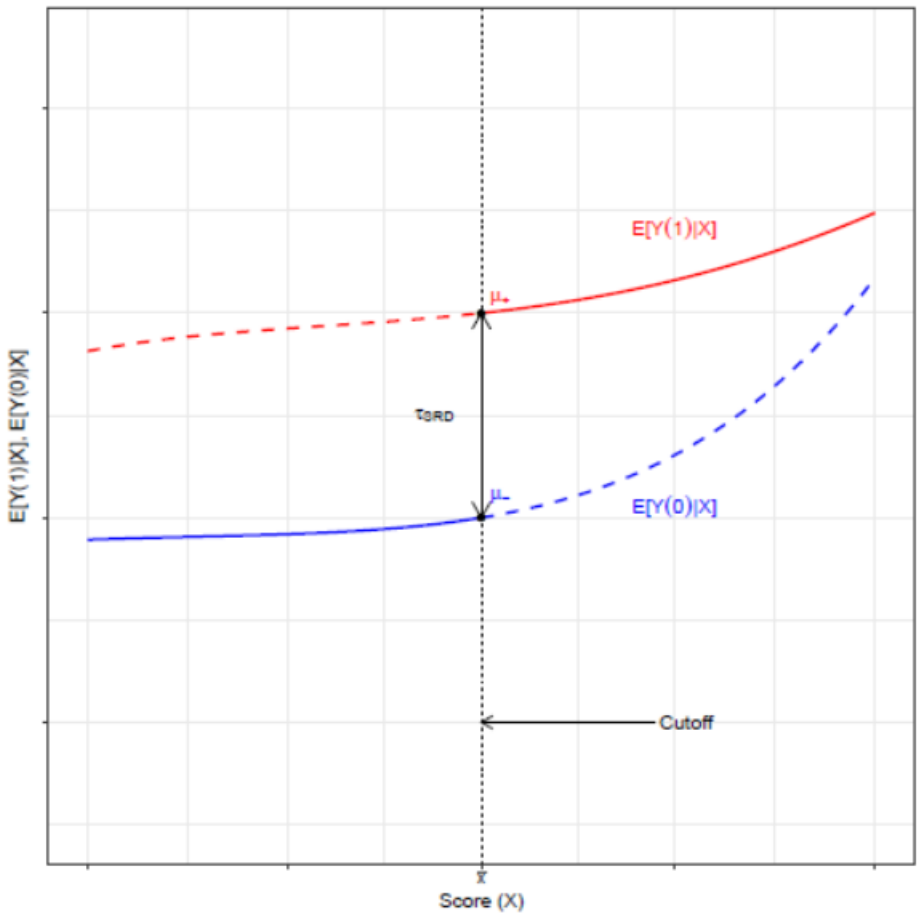
$E[Y_i(0) | X_i = x]$ **and** $E[Y_i(1) | X_i = x]$ **are continuous in** x**.**

With this assumption we can write

$$\tau_{SRD} = lim_{x \to c} E[Y_i | X_i = x] - lim_{x \leftarrow c} E[Y_i | X_i = x]$$

and estimate τ_{SRD} as the difference between the two regression functions estimated in the **neighborhood** of c.

Figure 2.2: RD Treatment Effect in Sharp RD Design



In the image above, τ_{SRD}, the effect of receiving treatment, is estimated as the difference in the mean outcomes of those that are right above the cutoff (who are treated) and those right below the cutoff (who are untreated). This gap will correctly estimate the treatment effect if the continuity assumption holds: had they not received treatment, the treated groups outcomes would be represented by the dashed blue line, and had they received treatment, the untreated groups outcomes would have been the dashed red line.

1.1.1 Simulation

Suppose you are the superintendent of a large school district. Last year you made participation in small reading groups mandatory for all students whose 3rd grade reading score was 75 points or less.You would like to know how these reading groups affected student performance on their 4th grade reading tests. Your data includes the 3rd and 4th grade reading scores for all 5000 students in your school district.

```
set.seed(7000)

sharp<-rnorm(5000, mean=80, sd=5)
sharp<-as.data.frame(sharp)

names(sharp)<-c("read3")
sharp$error<-rnorm(5000, mean=0, sd=5)
```

```r
sharp$pe3<-rnorm(5000, mean=90, sd=4)
sharp$height<-rnorm(5000, mean=130, sd=15)

sharp$treated<-0
sharp$treated[sharp$read3<=75]<-1


#the DGP
sharp$read4<-(-6)+0.8*sharp$read3+10*sharp$treated+sharp$error

sharp<-sharp[sharp$read3<78 & sharp$read3>72,]
```

1.2 Fuzzy RD

The fuzzy RD design (FRD) is similar in concept to the sharp RD except that D_i is no longer a deterministic function of X_i. Instead, the probability of treatment changes by some nonzero amount as the running variable crosses the threshold c, but this change in probability is less than 100 percentage points.

Formally we have that

$$0 < lim_{x \to c} P(D_i = 1 | X_i = x) - lim_{x \leftarrow c} P(D_i = 1 | X_i = x) < 1$$

This scenario is potentially more common than the sharp RD scenario as most things in life are determined by multiple factors and the influence of the running variable as it crosses the threshold c may be just one of these factors. In the fuzzy RD designs, there are now two causal effects to be estimated: the effect of crossing the threshold on the probability of treatment (which is 1 in the sharp RD) and the effect of crossing the threshold on the outcome. Formally, the fuzzy RD estimator is

$$\tau_{FRD} = \frac{lim_{x \to c} E[Y_i | X_i = x] - lim_{x \leftarrow c} E[Y_i | X_i = x]}{lim_{x \to c} E[D_i | X_i = x] - lim_{x \leftarrow c} E[D_i | X_i = x]}.$$

If this looks familiar, that's because it should. It's the direct analog of an IV estimator in which the instrument is an indicator for whether X_i lies directly above c. The way the IV estimator is being used, and should be interpreted here is very similar to how the IV estimator allowed us to recover the LATE estimate in an RCT. In a fuzzy RD there is the intend-to-treat group (say who have $X_i \geq c$ for example) and the control group ($X_i < c$). Because we are in fuzzy land, just because you are in the intend-to-treat group does not necessarily mean you get treated (there are never takers) and similarly, some in the control get treated (always takers). But there is a group of observations, the compliers, whose treatment status would change if they went from control to intend-to-treat (ie if they were moved across the threshold). This is analogous to an RCT with never and always takers. Thus, just like in the RCT context, if I just compare the outcomes of those that are above and below the threshold ($lim_{x \to c} E[Y_i | X_i = x] - lim_{x \leftarrow c} E[Y_i | X_i = x]$) the effect is "diluted" by the fact that for many observations, crossing the threshold has no effect (since they are always or never takers). Therefore, just as we were able to recover the local average treatment effect (LATE) effect in the RCT by scaling the average treatment effect by the change in the probability of treatment, here too we scale our estimate by the change in probability of receiving treatment ($lim_{x \to c} E[D_i | X_i = x] - lim_{x \leftarrow c} E[D_i | X_i = x]$). Thus, intuitively, the fuzzy RD design measures the average treatment effect for RD compliers at the threshold,

$$\tau_{FRD} = E[Y_i(1) - Y_i(0) | \text{unit } i \text{ is a complier and } X_i = c].$$

1.2.1 Simulation

Suppose you are the superintendent of a large school district. Last year you strongly encouraged students to participate in small reading groups if their 3rd grade reading score fell below 75 points.You would like to know how these reading groups affected student performance on their 4th grade reading tests. Your data

includes the 3rd and 4th grade reading scores for all 5000 students in your school district and whether or not a student participated in the small reading groups.

```
set.seed(2000)

fuzzy<-rnorm(5000, mean=80, sd=5)
fuzzy<-as.data.frame(fuzzy)

names(fuzzy)<-c("read3")
fuzzy$error<-rnorm(5000, mean=0, sd=5)
fuzzy$pe3<-rnorm(5000, mean=90, sd=4)
fuzzy$height<-rnorm(5000, mean=130, sd=15)


fuzzy$lowprob<-rbinom(5000,1,0.3)
fuzzy$highprob<-rbinom(5000,1,0.8)
fuzzy$treated<-NA
fuzzy$treated[fuzzy$read3>75]<-fuzzy$lowprob[fuzzy$read3>75]
fuzzy$treated[fuzzy$read3<=75]<-fuzzy$highprob[fuzzy$read3<=75]




#the DGP
fuzzy$read4<-(-6)+0.8*fuzzy$read3+10*fuzzy$treated+fuzzy$error

fuzzy<-fuzzy[fuzzy$read3<78 & fuzzy$read3>72,]
```

1.2.2 External Validity

The conditioning in the equation above does suggest that there are limitation to the applicability of RD estimates. First of all, we focus on observations that are in the neighborhood of the threshold since observations that are far below and far above the threshold likely differ from each other in many observable and unobservable ways. In an RD design, we are implicitly assuming that whether an observation fell just above or just below the threshold is effectively arbitrary and these observations are identical in all observable and non-observable characteristics and conditions, except for their treatment status. In this case we can measure a Local Average Treatment Effect (LATE) that is valid around the threshold. What this means is that RD estimates are inherently localized since the effects are estimated for a sub population where their X_i is in the neighborhood of c. Because of this, even though RD estimates have a relatively high degree of internal validity, it is important to think about their external validity since treatment effects could be quite different for observation where X_i is quite different from c. When we are considering a fuzzy RD, we must also be aware that we are estimating effects on compliers, which is yet another further subsample of the population.

1.3 The Graphs

At the heart of any good RD paper is the graphical analysis. The strength of the RD design is that the treatment assignment rule is known (or at least partially known). We should therefore be able to see discontinuous changes in the treatment and the outcome (if there is an effect) as the running variable crosses the threshold c. Any RD that fails to exhibit a visually perceptible break in treatment probability at the discontinuity threshold is basically not credible, regardless of the regression results. Conversely, a break that is visually perceptible will almost surely be statistically significant. So with RD papers, the statistical results really take a back seat to the graphical analysis.

There are several types of graphs that make appearances in RD analyses. All of them require some data

preperation and design as a simple scatterplot of your data is unlikely to reveal the patterns you wish to illustrate. The main graphs of an RD design are basically a histogram-type plot that presents the average value of the outcome, treatment status and covariates at evenly spaced values of the running variables. Gernerating these plots requires choosing two key parameters: the binwidth, h, and the number of bins shown to the left and right of the threshold value, K_0 and K_1. Once these choices are made, you construct $K_0 + K_1$ bins: K_0 evenly spaced bins of width h below the threshold value and K_1 evenly spaced bins of width h above the threshold value. Note, you should avoid having any bin crossing the threshold value c as this will make the discontinuities we hope to ovserve less easy to identify visually.

1.3.1 Treatment status

RD papers often include a graph that plots treatment by the running variable. We expect to see a visually perceptible discontinuity in the probability of treatment as the running variable crosses the threshold.

After constructing the bins described above, plotting this graph requires calculating, $\bar{D}_k$, the average treatment level in the bin

$$\bar{D}_k = \frac{1}{N_k} \sum_{i=1}^{N} D_i * 1(b_k < X_i \leq b_{k+1})$$

and plotting these values against the midpoint of each of the bins.

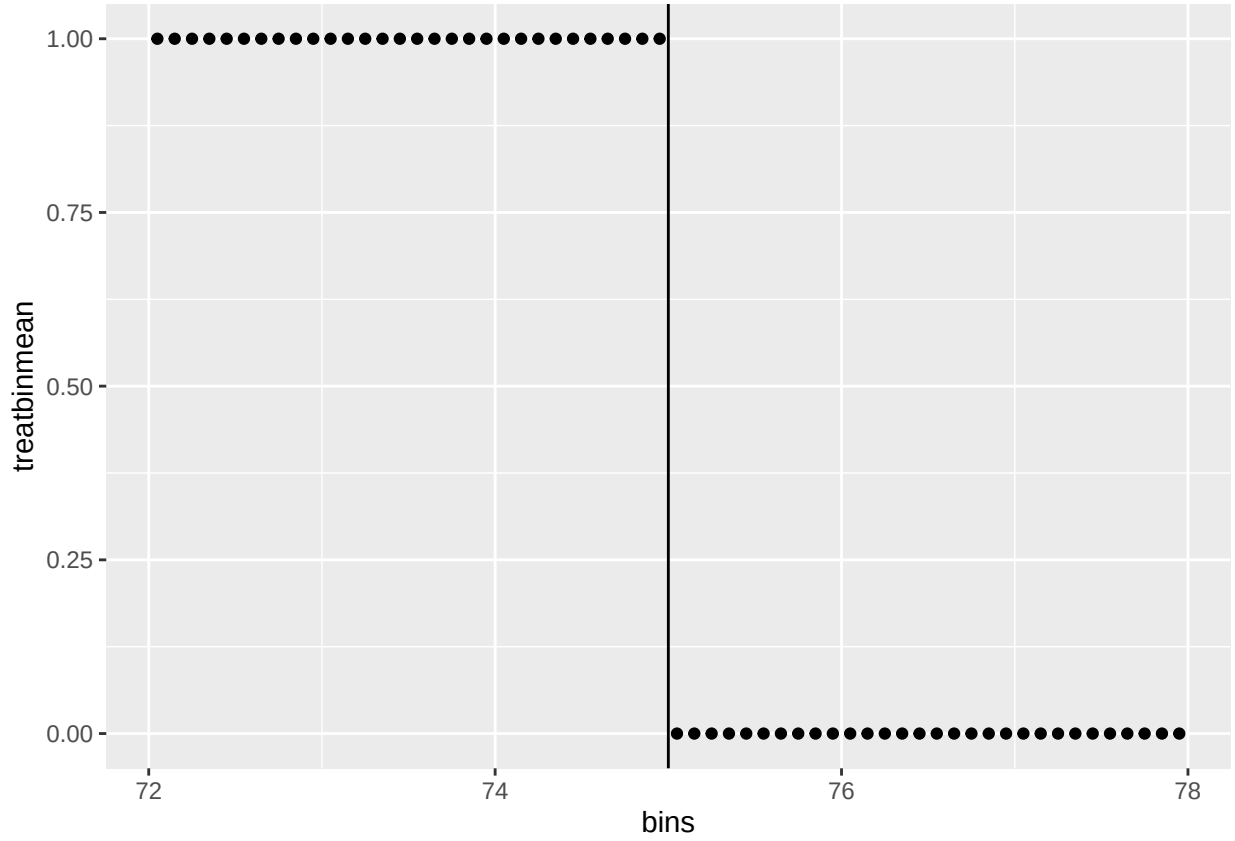
In a sharp RD, $\bar{D}_k$ should be either 0 or 1 depending on whether the bin is above or below the threshold. Thus this graph is not particularly interesting in a sharp RD (and is therefore sometimes relegated to the appendix), and primarily serves to validate the RD design.

```
library(dplyr)
library(ggplot2)
#I will break up the data into 60 bins (30 above and 30 below the threshold)

cuts<-c(72,72.1,72.2,72.3,72.4,72.5,72.6,72.7,72.8,72.9,73,
        73.1,73.2,73.3,73.4,73.5,73.6,73.7,73.8,73.9,74,
        74.1,74.2,74.3,74.4,74.5,74.6,74.7,74.8,74.9,75,
        75.1,75.2,75.3,75.4,75.5,75.6,75.7,75.8,75.9,76,
        76.1,76.2,76.3,76.4,76.5,76.6,76.7,76.8,76.9,77,
        77.1,77.2,77.3,77.4,77.5,77.6,77.7,77.8,77.9,78)
midpoints<-cuts[2:61]-0.05

sharp$bins <- cut(sharp$read3,
                  breaks=cuts,
                  include.lowest=TRUE,
                  right=FALSE,
                  labels=midpoints)


sharp_mean<-sharp %>%
    group_by(bins) %>%
    dplyr::summarize(outbinmean = mean(read4, na.rm=TRUE), treatbinmean=mean(treated, na.rm=TRUE), pebi
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
sharp_mean$bins<-as.numeric(as.character(sharp_mean$bins))

plot1shp<-ggplot(sharp_mean, aes(x=bins, y=treatbinmean))+
         geom_point()+
```

```
        geom_vline(xintercept = 75)
plot1shp
```



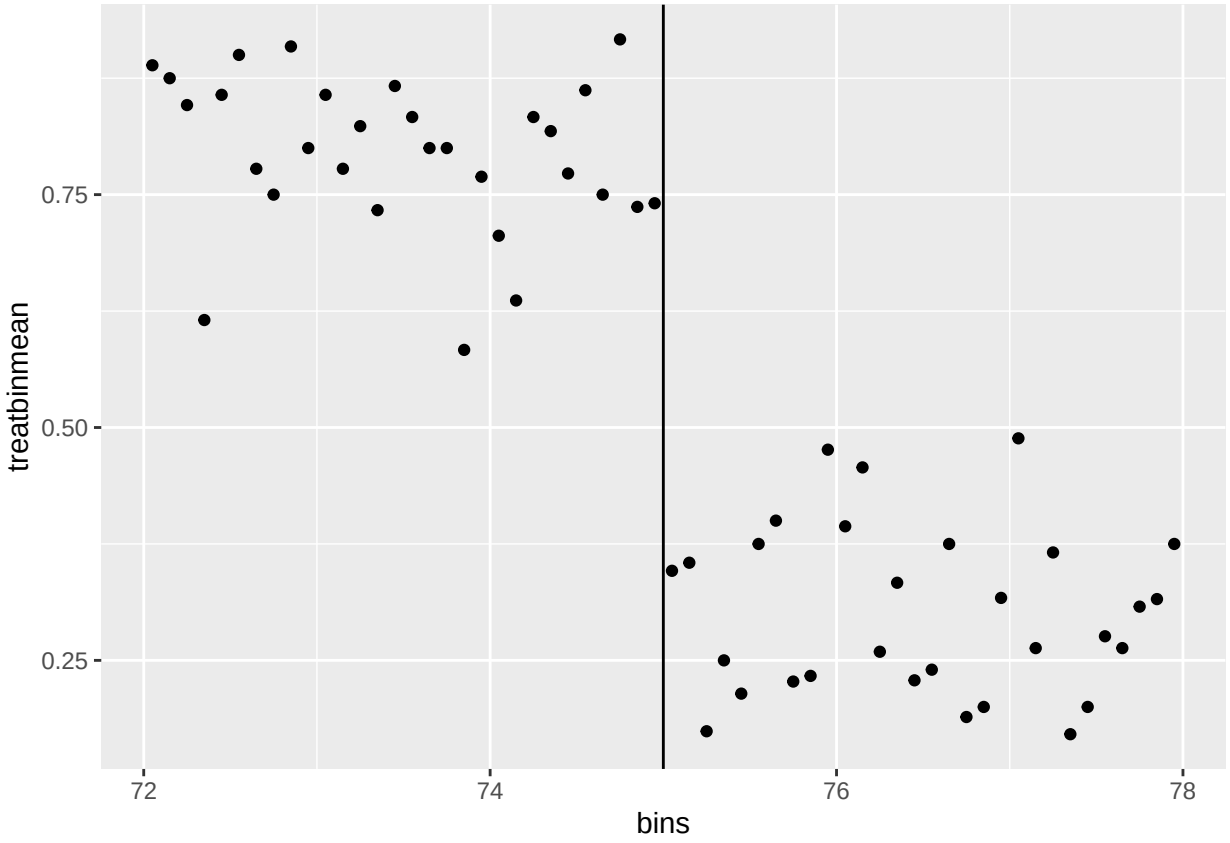
In a fuzzy RD, $\bar{D}_k$ can take on many possible values. This plot should show that there is a visual discontinuity in the probability of getting treated at the threshold c. A visual break implies that crossing the threshold has a significant effect on the probability of treatment. In fuzzy RD designs, this graph is equivalent to the first stage in an IV analysis. It shows that we have found a tool that generates some random variation we can leverage to estimate unbiased treatment effects.

```
library(dplyr)
library(ggplot2)
#I will break up the data into 60 bins (30 above and 30 below the threshold)

cuts<-c(72,72.1,72.2,72.3,72.4,72.5,72.6,72.7,72.8,72.9,73,
        73.1,73.2,73.3,73.4,73.5,73.6,73.7,73.8,73.9,74,
        74.1,74.2,74.3,74.4,74.5,74.6,74.7,74.8,74.9,75,
        75.1,75.2,75.3,75.4,75.5,75.6,75.7,75.8,75.9,76,
        76.1,76.2,76.3,76.4,76.5,76.6,76.7,76.8,76.9,77,
        77.1,77.2,77.3,77.4,77.5,77.6,77.7,77.8,77.9,78)
midpoints<-cuts[2:61]-0.05

fuzzy$bins <- cut(fuzzy$read3,
                  breaks=cuts,
                  include.lowest=TRUE,
                  right=FALSE,
                  labels=midpoints)
```

```
fuzzy_mean<-fuzzy %>%
    group_by(bins) %>%
    dplyr::summarize(outbinmean = mean(read4, na.rm=TRUE), treatbinmean=mean(treated, na.rm=TRUE), pebi
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
fuzzy_mean$bins<-as.numeric(as.character(fuzzy_mean$bins))

plot1fuz<-ggplot(fuzzy_mean, aes(x=bins, y=treatbinmean))+
         geom_point()+
         geom_vline(xintercept = 75)
plot1fuz
```



1.3.2 Outcomes

The main course of an RD paper is a plot of the outcome by the running variable. If there is a treatment effect, we would expect to see a discontinuity here too.

Plotting this graph requires calculating, $\bar{D}_k$, the average outcome in each bin

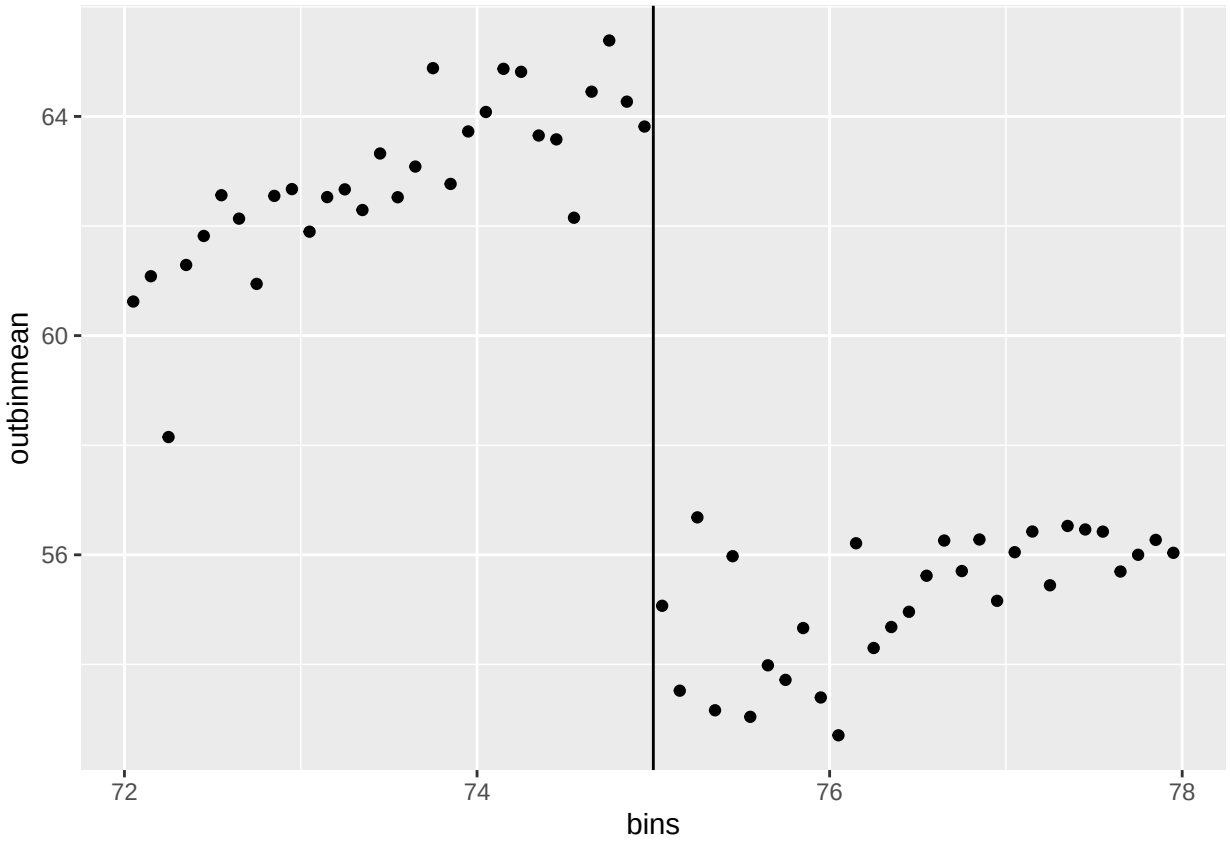
$$\bar{Y}_k = \frac{1}{N_k} \sum_{i=1}^{N} Y_i * 1(b_k < X_i \leq b_{k+1})$$

and plotting these values against the midpoint of each of the bins.

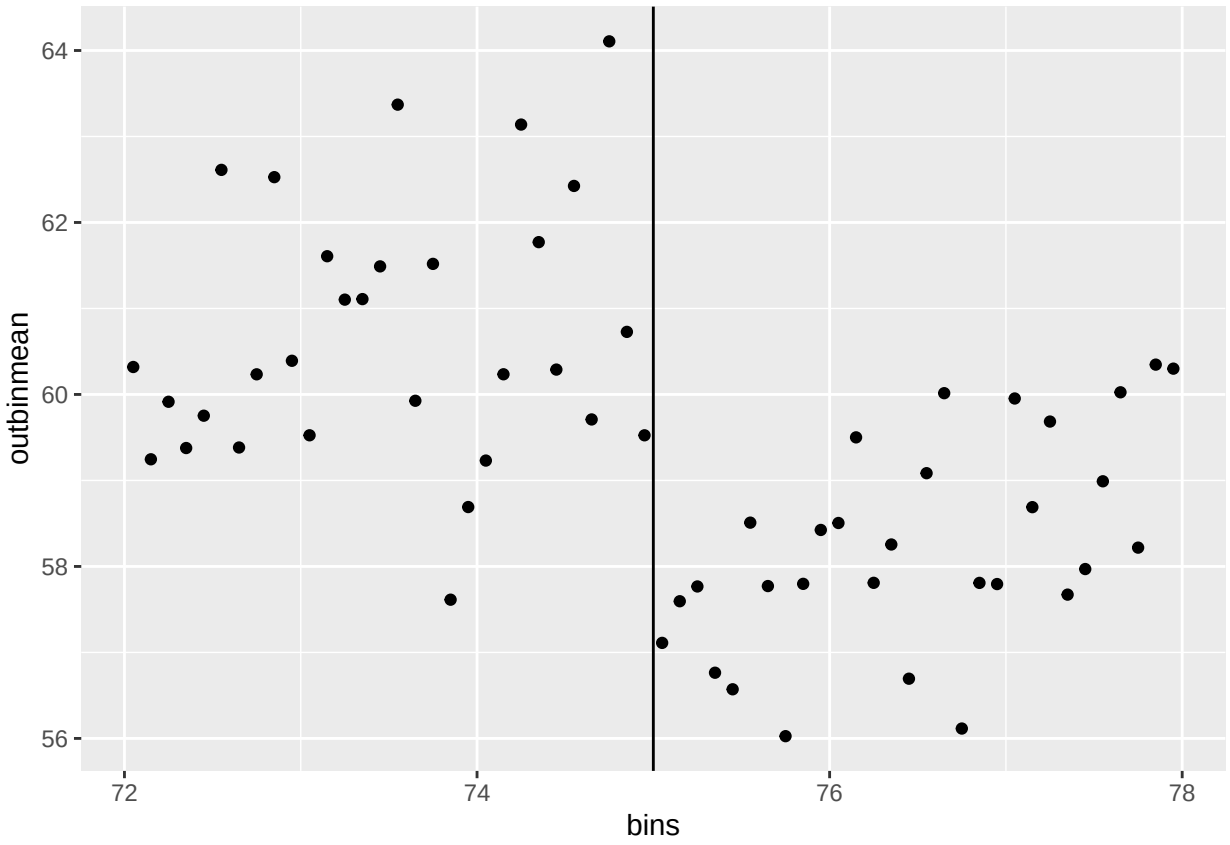
A visual break at c implies that crossing the threshold has a significant effect on the outcome, which in turn implies (under our assumptions) that the treatment has a significant effect on the outcome. This graph is the equivalent of the reduced form in an IV analysis.

In addition to inspecting the threshold for a discontinuity, you should also inspect whether there are any other discontinuities of similar (or greater) magnitude at other values of the running variable. If there are, and if there is not a clear a priori reason to expect these discontinuities, then the research design is called into question - effectively we have detected a violation of Assumption 1 (smoothness in expected potential outcomes).

```
plot2shp<-ggplot(sharp_mean, aes(x=bins, y=outbinmean))+
         geom_point()+
         geom_vline(xintercept = 75)
plot2shp
```




```
plot2fuz<-ggplot(fuzzy_mean, aes(x=bins, y=outbinmean))+
         geom_point()+
         geom_vline(xintercept = 75)
plot2fuz
```



1.3.3 Covariates

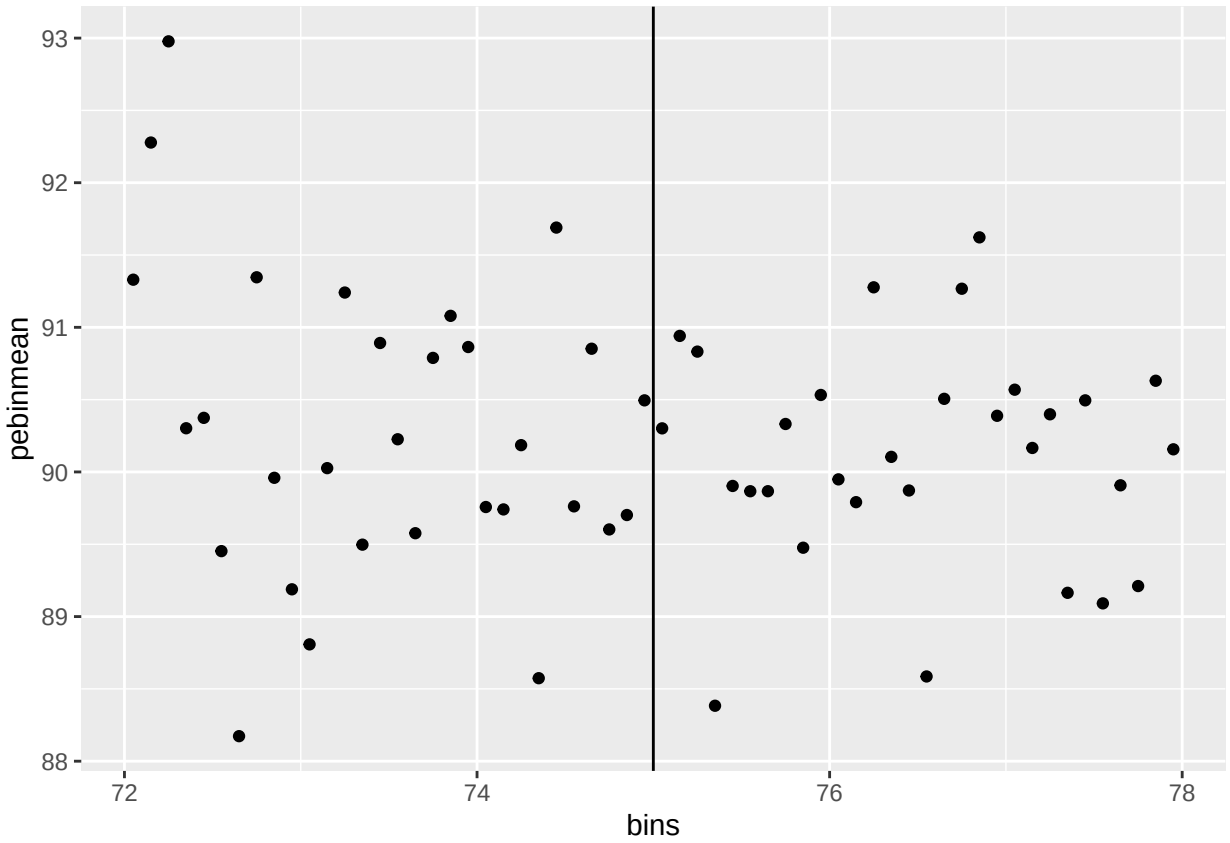
Using the same methodology as above, it is common to plot the average value of certain covariates that may be related to the outcome but should not be affected by the treatment. As above, we calculate $\bar{Z}_i$ where

$$\bar{Z}_k = \frac{1}{N_k} \sum_{i=1}^{N} Z_i * 1(b_k < X_i \leq b_{k+1})$$

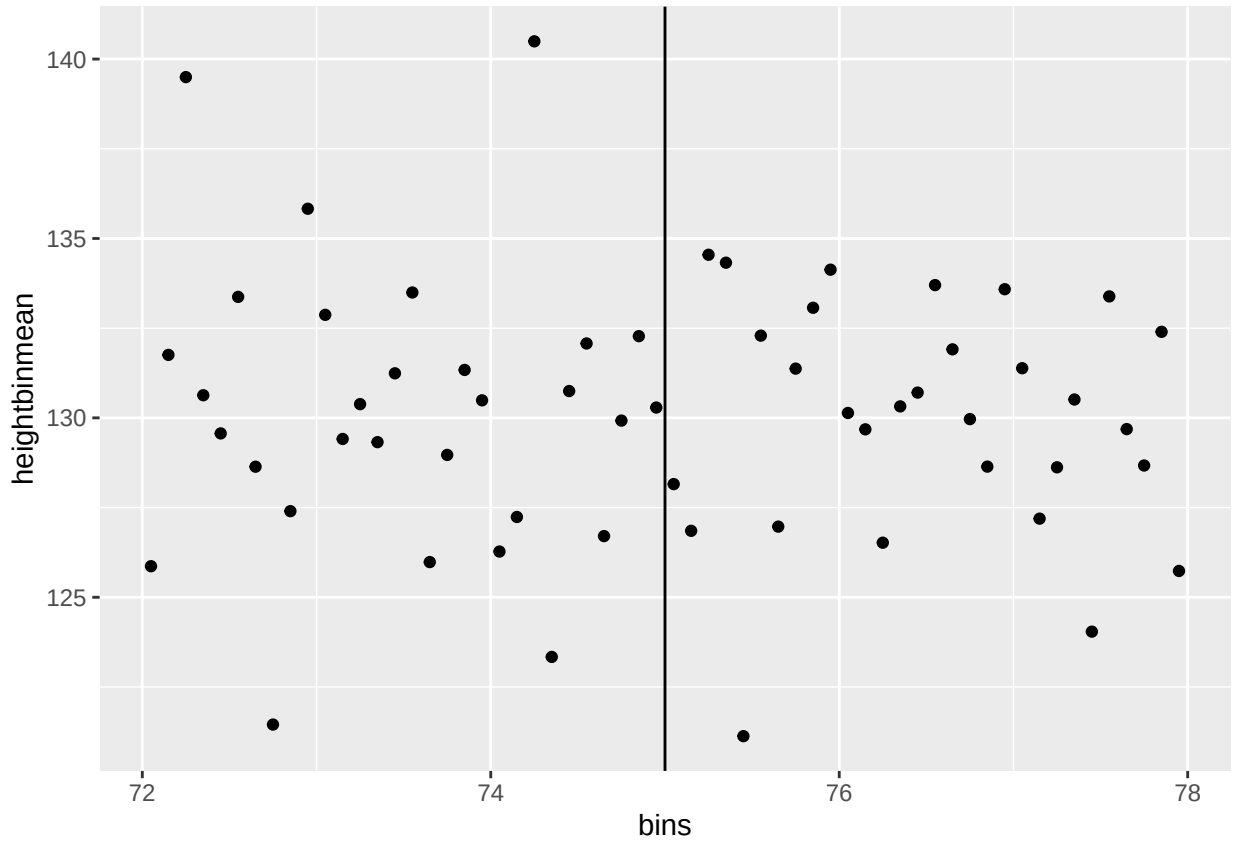
is plotted against the midpoint of each bin.

For the covariate graphs, if the research design is valid there should not be any discontinuity in $\bar{Z}_k$ as the running variable crosses the threshold c. This plot allows us to determine whether the covariate is balances across the threshold- the equivalent of showing covariates are balanced in an RCT. As in an RCT, you are essentially checking that the treated and un-treated groups are similar along covariates which validates the argument that treatment is as good as randomly assigned within the bandwidth being studied.

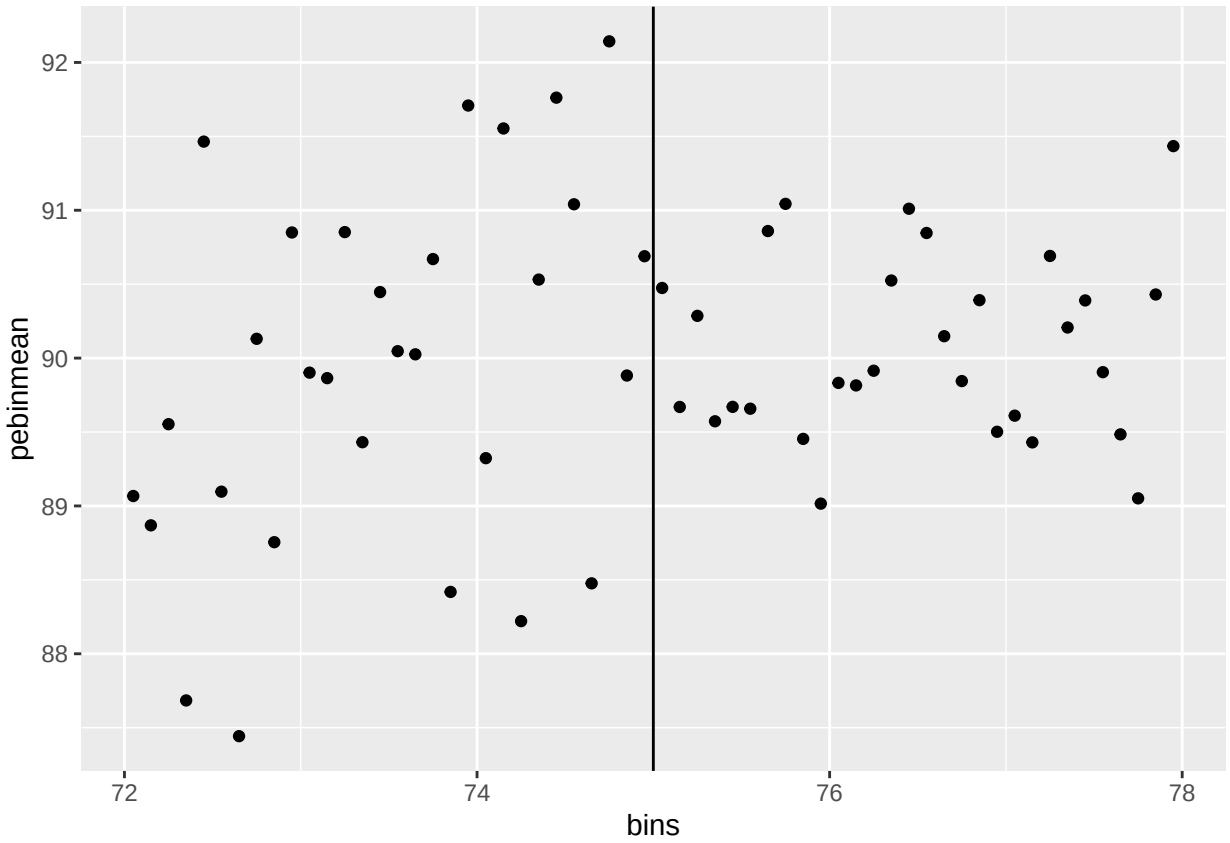
```
plot3shp<-ggplot(sharp_mean, aes(x=bins, y=pebinmean))+
        geom_point()+
        geom_vline(xintercept = 75)
plot3shp
```



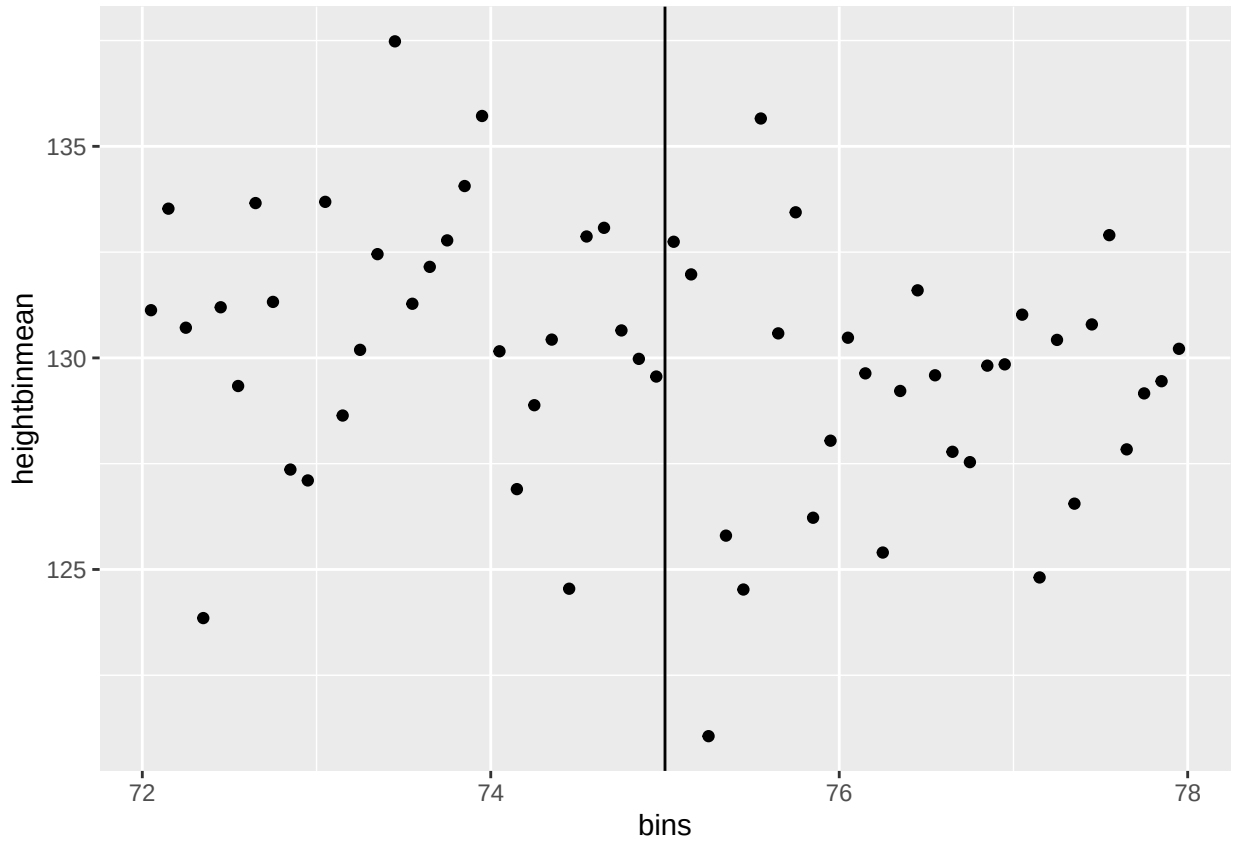

```
plot4shp<-ggplot(sharp_mean, aes(x=bins, y=heightbinmean))+
         geom_point()+
         geom_vline(xintercept = 75)
plot4shp
```




```
plot3fuz<-ggplot(fuzzy_mean, aes(x=bins, y=pebinmean))+
         geom_point()+
         geom_vline(xintercept = 75)
plot3fuz
```




```
plot4fuz<-ggplot(fuzzy_mean, aes(x=bins, y=heightbinmean))+
         geom_point()+
         geom_vline(xintercept = 75)
plot4fuz
```



1.3.4 Density of the Running Variable

Finally, it is also common to plot the density of the running variable. For each bin, you calculate

$$N_k = \sum_{i=1}^{N} 1(b_k < X_i \leq b_{k+1})$$

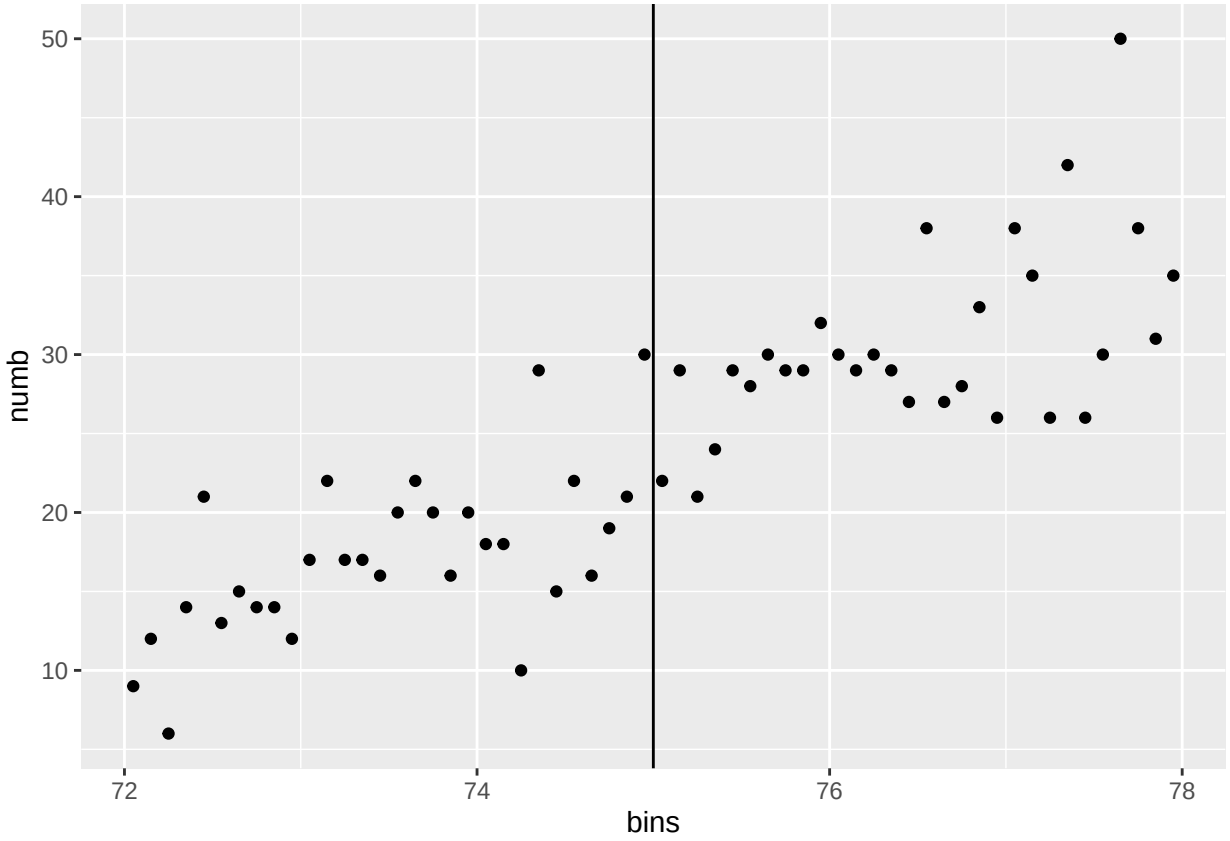
and plot these against the midpoint of the bin.

A major concern in RD designs is that individuals may "game" the assignment rule. That is to say, if individuals understand the assignment mechanism and can manipulate their value of the running variable, then they may be able to place themselves just above (or just below) the threshold c. In that case, the individuals just above the threshold will disproportionately consist of those gaming the rule and they will not be directly comparable to the individuals lying just below the threshold. For instance, consider a scholarship that activates only when scores on a test fall above a certain threshold c. Shrewd students could retake the test many times until they pass the threshold. If a researcher uses an individuals maximum test score as the running variable, motivated individuals who retake the test many times are more likely to fall just above the threshold, then just below it. Thus this group of observations is selected and no longer directly comparable to the observations that fall directly below the threshold.

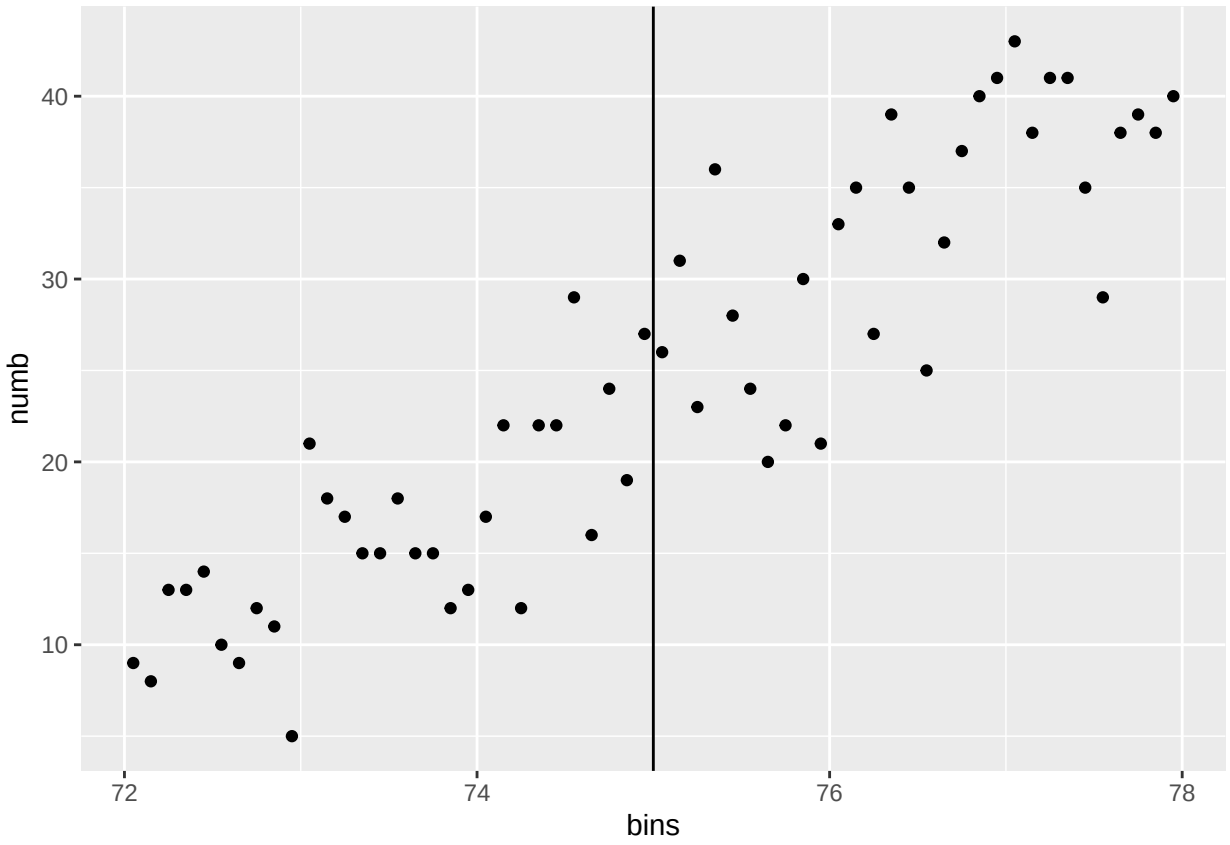
To address this issue, you can inspect the density of the running variable as it crosses c. If units are manipulating their values of the running variable, to fall just above or below c, then we should observe a discontinuity in the distribution of the running variable as it crosses c. If the distribution of the running variable is smooth as it crosses c, then it's unlikely that individuals are gaming the assignment mechanism.

```
plot5shp<-ggplot(sharp_mean, aes(x=bins, y=numb))+
        geom_point()+
```

```
        geom_vline(xintercept = 75)
plot5shp
```




```
plot5fuz<-ggplot(fuzzy_mean, aes(x=bins, y=numb))+
        geom_point()+
        geom_vline(xintercept = 75)
plot5fuz
```



1.4 Estimation

RD estimations can be done quite easily in a regression framework. The first step typically involves selecting a bandwidth, h, that will determine the regression sample on either side of the threshold point c.[1]

From here, estimation will depend on whether you are in sharp or fuzzy land.

1.4.1 Sharp RD Estimation

We then fit a linear regression on either side of the threshold point for the samples with $X - i \in (c - h, c)$ and $X_i[c, c + h)]$. This can be done by estimating some version (you can include covariates) of the following specification,

$$Y_i = \alpha + \tau D - i + \beta(X_i - c) + \gamma(X_i - c) * D_i + u_i \text{ for } c - h < X_i \leq c + h.$$

With this estimation strategy, $\hat{\tau}_{SRD}$ will estimate the treatment effect for units right at the threshold.

```
sharp$runminc<-sharp$read3-75
shpestim<-felm(read4~treated+runminc+treated*runminc, sharp)

stargazer(shpestim, type="latex")
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Tue, Nov 03, 2020 - 3:31:08 PM

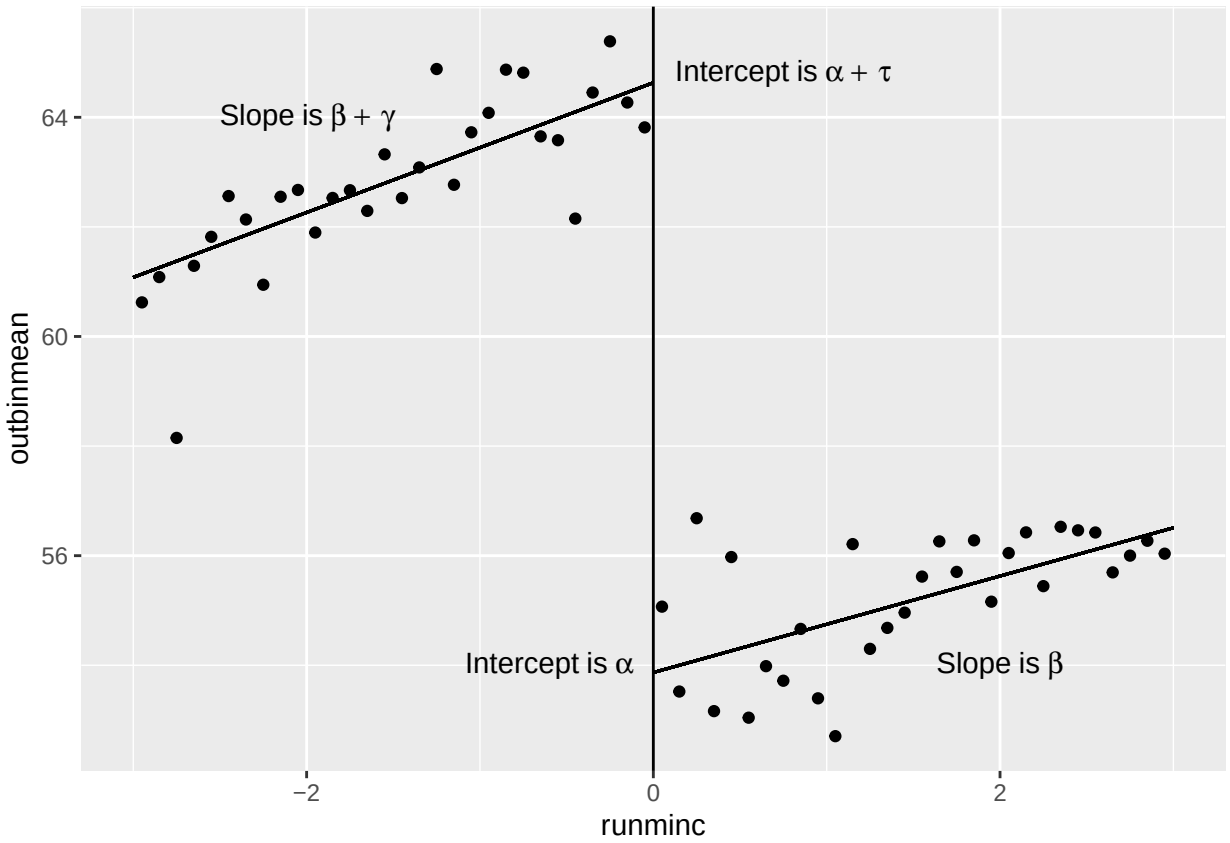
[1] While there is an econometric literature that discusses methods to choose the optimal bandwidth if you wish to get technical, in practice this choice is more art than science and many papers make this choice fairly arbitrarily. Whether you go the art or science route, it is generally good practice to check that results are not sensitive to the choice of bandwidth.

Table 1:

	Dependent variable:
	read4
treated	10.764***
	(0.555)
runminc	0.880***
	(0.197)
treated:runminc	0.304
	(0.333)
Constant	53.868***
	(0.356)
Observations	1,436
R^2	0.358
Adjusted R^2	0.356
Residual Std. Error	5.128 (df = 1432)
Note:	*p<0.1; **p<0.05; ***p<0.01

It is helpful to see how the coefficients estimated above translate to the RD graph.

```
sharp_mean$runminc<-sharp_mean$bins-75

plot6shp<-ggplot(sharp_mean, aes(x=runminc, y=outbinmean))+
         geom_point()+
         geom_vline(xintercept = 0)+
         geom_segment(aes(x = 0, xend = 3, y = shpestim$coefficients[1], yend = shpestim$coefficients[1]
         geom_segment(aes(x = -3, xend = 0, y = shpestim$coefficients[1]+ shpestim$coefficients[2]+(-3*
  #adding some labeling for course notes:
  annotate("text", x = 0.75, y = 64.8, label = "Intercept~is~alpha~+~tau" ,parse = TRUE)+
  annotate("text", x = -0.6, y = 54, label = "Intercept~is~alpha" ,parse = TRUE)+
  annotate("text", x = -2, y = 64, label = "Slope~is~beta~+~gamma" ,parse = TRUE)+
  annotate("text", x = 2, y = 54, label = "Slope~is~beta" ,parse = TRUE)
plot6shp
```



1.4.2 Fuzzy RD Estimation

In the fuzzy RD design, we have two effects to estimate: the effect of crossing the threshold on the treatment (the "first stage") and the effect of crossing the threshold on the outcome (the "reduced form"). As you might expect, we apply the same methodology as in earlier IV's to estimate the effect of crossing the threshold on Y_i and the effect of crossing the threshold on D_i. For the sample with $c - h < X_i \leq c + h$ we run some version (you can include covariates) of the following regressions,

$$Y_i = \pi_0 + \pi_1 D_i + \pi_2(X_i - c) + \pi_3(X_i - c) * D_i + u_i$$

and

$$D_i = \gamma_0 + \gamma_1 Z_i + \gamma_2(X_i - c) + \gamma_3(X_i - c) * Z_i + v_i$$

where $Z_i = 1(X_i \geq c)$. The fuzzy RD estimator is then

$$\hat{\tau}_{FRD} = \frac{\hat{\pi}_1}{\hat{\gamma}_1}.$$

In other words, the FRD estimator is simply the ratio of the reduced form and the first stage estimates, i.e. the effect of crossing the discontinuity threshold on the outcome, scaled by the effect of crossing the discontinuity threshold on the probability of treatment.

```
fuzzy$runminc<-fuzzy$read3-75

#first stage
fuzzy$ittgroup<-0
```

```
fuzzy$ittgroup[fuzzy$read3<=75]<-1

fuzfs<-felm(treated~ittgroup+runminc+ittgroup*runminc,fuzzy)

#reduced form
fuzrf<-felm(read4~ittgroup+runminc+ittgroup*runminc,fuzzy)


fuzzy$interedog<-fuzzy$treated*fuzzy$runminc
fuzzy$interinst<-fuzzy$ittgroup*fuzzy$runminc
#IV
fuziv<-felm(read4~runminc|0|(treated|interedog~ittgroup+runminc+interinst),fuzzy)
```

```
## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either
## rank-deficient or indefinite
```

```
stargazer(fuzfs, fuzrf, fuziv, type="latex")
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Tue, Nov 03, 2020 - 3:31:08 PM

Table 2:

	Dependent variable:		
	treated	read4	
	(1)	(2)	(3)
ittgroup	0.464***	4.253***	
	(0.048)	(0.722)	
runminc	−0.004	0.821***	1.057***
	(0.017)	(0.251)	(0.332)
ittgroup:runminc	−0.013	−0.479	
	(0.029)	(0.441)	
‘treated(fit)‘			9.201***
			(1.162)
‘interedog(fit)‘			−0.685
			(0.631)
Constant	0.308***	57.010***	54.182***
	(0.030)	(0.460)	(0.640)
Observations	1,460	1,460	1,460
R^2	0.213	0.037	0.454
Adjusted R^2	0.211	0.035	0.453
Residual Std. Error (df = 1456)	0.443	6.728	5.065
Note:			*p<0.1; **p<0.05; ***p<0.01

In a fuzzy RD, you will not be able to plot and represent visually the “scaled” treatment effect as the graphs are limited to the graphical equivalent of the first stage and the reduced form estimates.