EC 587 HW5

Erik Andersen

November 2023

Question 1

a)

The results from the regression are reported below.

Variable	Estimate	SE
Intercept	0.330	0.0263
Test	0.267	0.0688
Treatment	0.165	0.0377
Test:Treatment	0.0151	0.0967

b)

First, not the outcome variable is completion of secondary school. The intercept is the probability of completing secondary school if you get an average score on the test. Zero is normalized in this data set to the average score. The test coefficient is the extra probability of completing secondary school for each standard deviation above the mean on the test. The treatment coefficient is the extra probability of completing secondary school for someone who scored above average on the exam. The interaction term is the difference is slope on the two sides of the discontinuity. So for someone above the mean value, they get an extra boost of about one percent per standard deviation above the mean of a chance to complete secondary school.

The treatment effect is identified by the coefficient on treatment. The other coefficient identify the effects described above.

c)

The point estimates from each of the bandwidths are reported below. First, we will look at the standard errors. As we should expect, as the bandwidth shrinks the standard errors increase. We are trading off precision for unbiasedness. The point estimates increase as we shrink the bandwidth which suggests that with a large bandwidth we are underestimating the true effect.

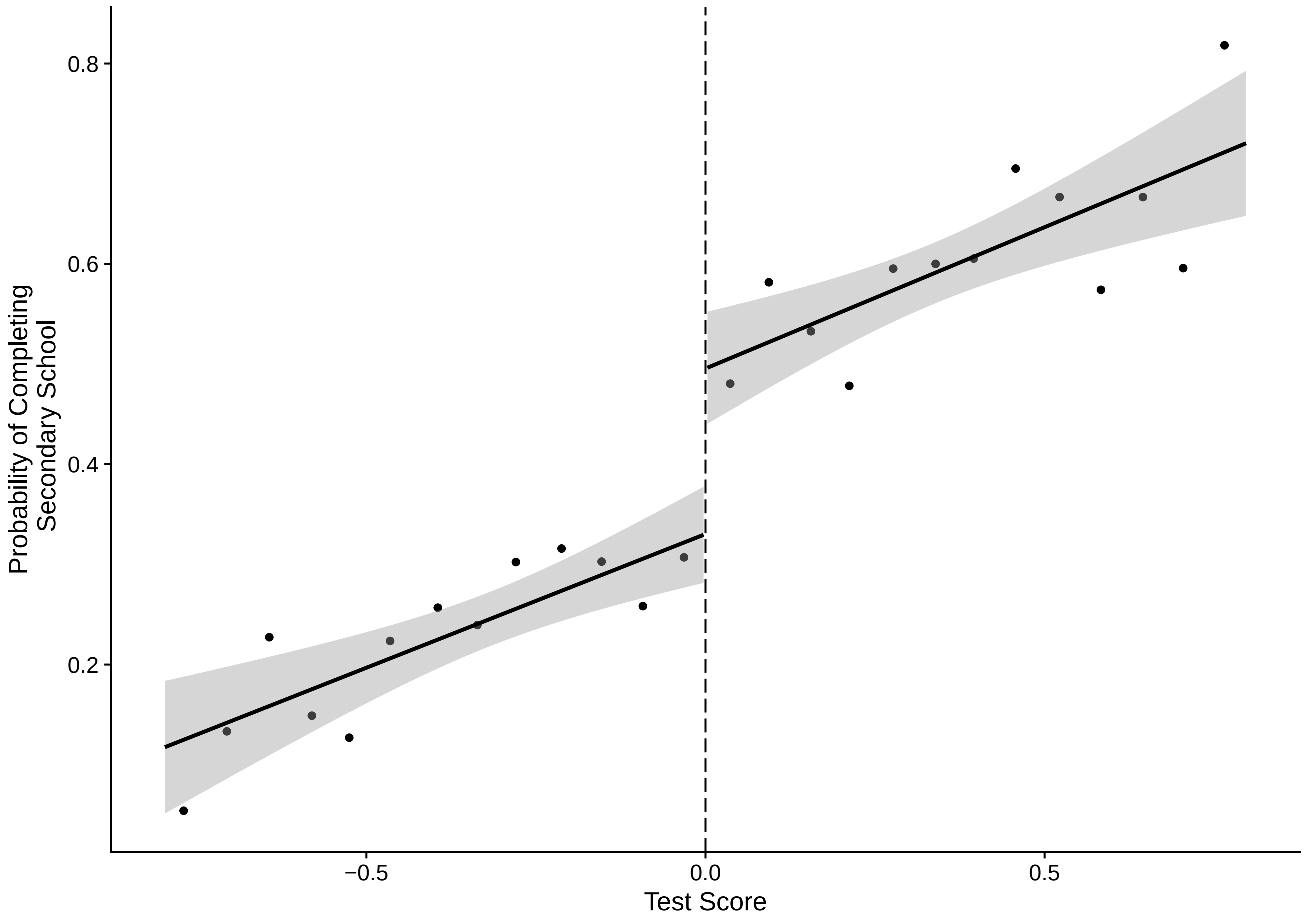
Bandwidth	Treatment Effect	SE
0.8	0.165	0.0377
0.4	0.186	0.0521
0.2	0.237	0.0752
0.1	0.224	0.114

d)

The estimated treatment effect from the default options of the rdd package estimation is 0.181. This is similar to the estimate from part c when the bandwidth was 0.4. The function also reports the LATE with half and double the optimal bandwidth. These are respectively 0.227 and 0.171. These are very close to our estimates in part c with bandwidth of 0.2 and 0.8 respectively.

e)

The rdd graph is below. The graph is showing the treatment effect we've been estimating in the previous parts. The running variable is the test score, and the outcome is the probability of finishing high school. The effect we've estimated is the jump of the regression line at the zero point for test score.



f)

The treatment effect changing the kernel to rectangular, and changing the bandwidth to 0.8 yields a LATE of 0.164. This is basically exactly the same as our estimate from part c with the restriction of 0.8.

g)

The results of the two stage least square regression are reported below. These results are close to the results reported in the paper. The estimates for the interaction, and the female coefficient are almost exact matches. The estimates for the other two are farther away from their counterparts in the paper, but still qualitatively the same.

Variable	Estimate	SE
Secondary	0.726	0.311
Test	0.612	0.159
Test:Treatment	-0.311	0.133
Female	-0.176	0.0483

h)

When we run the two stage regression in the RDD framework and the default options, the estimated LATE is 1.47. This isn't particularly close to the estimated effect from the 2SLS regression in part g. It is about double.

i)

To generate a similar result to part g, we use a rectangular kernel and a bandwidth of 0.6. We get an estimate of 0.744 which is qualitatively the same as in part g.

Question 2

a)

The kernel densities are shown below. Reported wealth's 1 and 2 are very similar in all bandwidth specifications. They are smooth around the zero point. The red line represents reported wealth 2. It is slightly bumpier than the red line, but smooth around the true values. This fits that data generating process of the true value plus some random noise. The reported wealth 3 is the blue line. It matches the true values on the edges, but not around the middle. The data generating process has people just above zero report their wealth just below 0. This is shown on the density plot with a bump to the left of 0, and a dip to the right. Reported wealth 4 is the green line. In that case, some people just randomly report their wealth as too low. The distribution is shifted to the left representing this.

b)

Histograms of the four scenarios are shown below. This gives us an ocular test of the smoothness of the running variable around the cutoff of zero. I did this for all scenarios and bandwidths, but for brevity I only report the largest bandwidth ones below. We see the running varirable is smooth in scenarios 1 and 2, and fairly smooth in scenario 4. There is a sharp discontinuity in scenario 3. Scenario 4 doesn't have a cutoff so much as a slow decline as we'd expect from the data generating process.

c)

The p-values from the McCrary, Kovak test for the four scenarios are reported in the table below. The results are as expected. The first two reports shouldn't have a discontinuity around 0, and we have no evidence of this. The data generating process should generate the strongest discontinuity for scenario 3, which this test shows. We should see some discontinuity from scenario 4, but not as strong as in scenario 3, which is exactly what we see. We have just enough evidence to support a discontinuity at the 5% level.

Scenario	P-Value
Report 1	0.9792
Report 2	0.3191
Report 3	8.155e-56
Report 4	0.04972

d)

The new wealth's are summarized by the scatter plots below. I graph the various reported wealth's against the new generated wealth's on the y axis. This shows the discontinuity generated by the policy under the four scenarios of reporting.

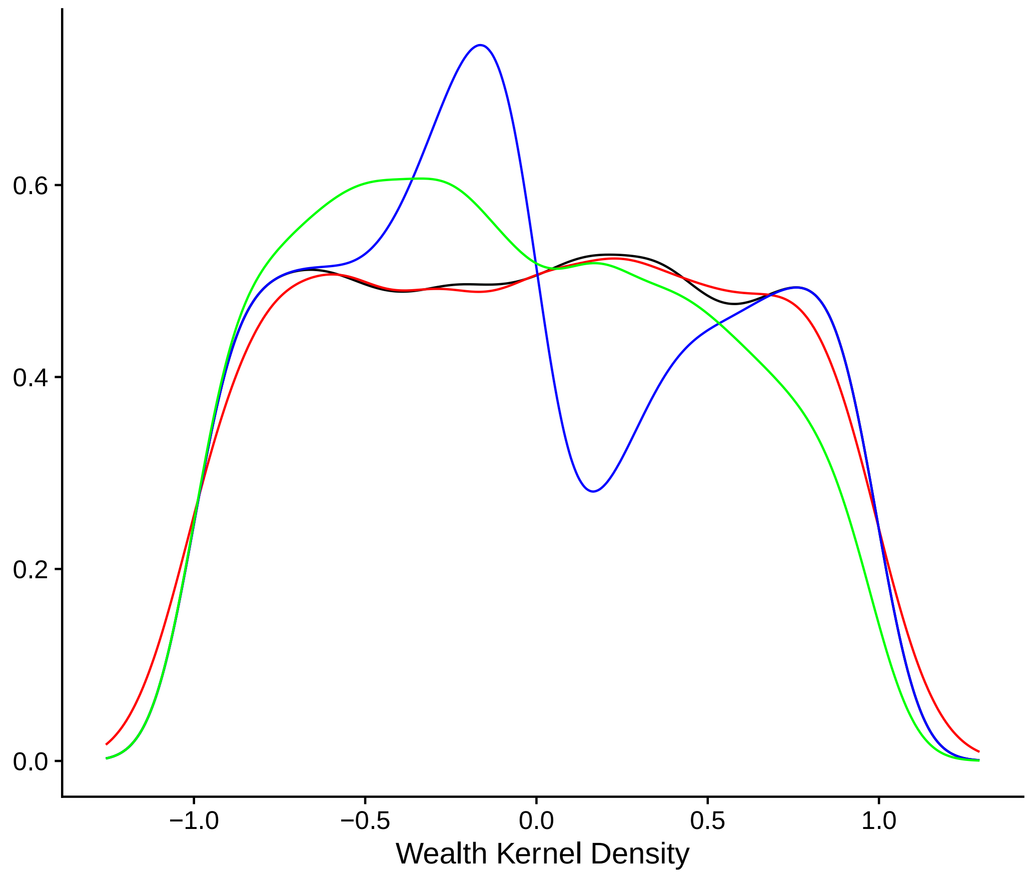
The discontinuity generated in scenario 1 generates the cleanest discontinuity. The only noise is from the program itself, so we should get an accurate estimate from the discontinuity. The second scenario seems to have a smaller discontinuity. There is classical measurement error in reported wealth, so the effect on new wealth is smaller. In the third scenario, we have a very strange discontinuity. In this scenario, people just above 0 report just below 0, so we have bunching below 0. If we estimated an rdd, we would estimate a much larger effect than is correct. The fourth scenario generates a very noisy graph. We have a bit of bunching below the cutoff so again we violate the continuity assumption, so we estimate the wrong effect.

e)

The treatment effects from the simple regression of treatment on new wealth controlling for the running variable of reported wealth are reported below. The results are expected. In scenario two, the treatment effect is attenuated because of the classical measurement error. In scenario 3, we estimate a huge effect, but it is not identifying anything causal because of violating continuity. In scenario 4, we estimate something close to the true effect, but I'm not sure what this identifies. These results are consistent for a variety of bandwidths.

Scenario	Treatment Effect
Report 1	0.200550
Report 2	0.16551
Report 3	0.32850
Report 4	0.18619

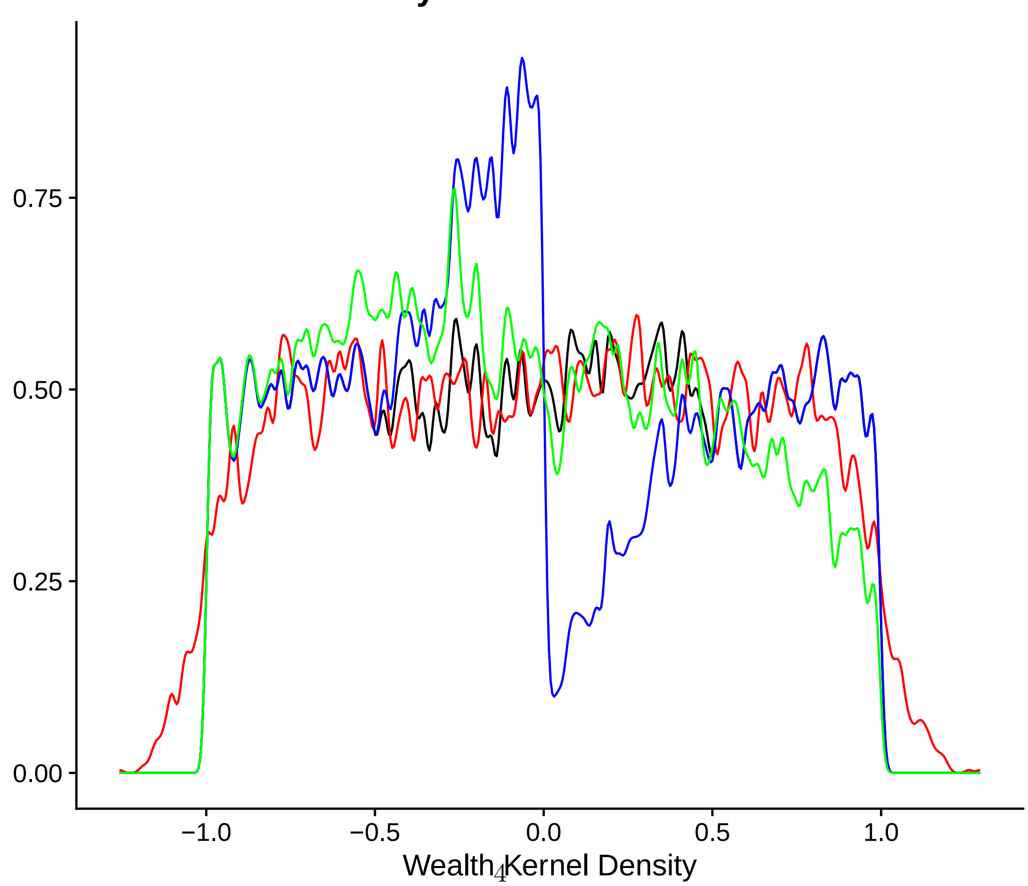
Wealth Kernel Density with Bandwidth = 0.1

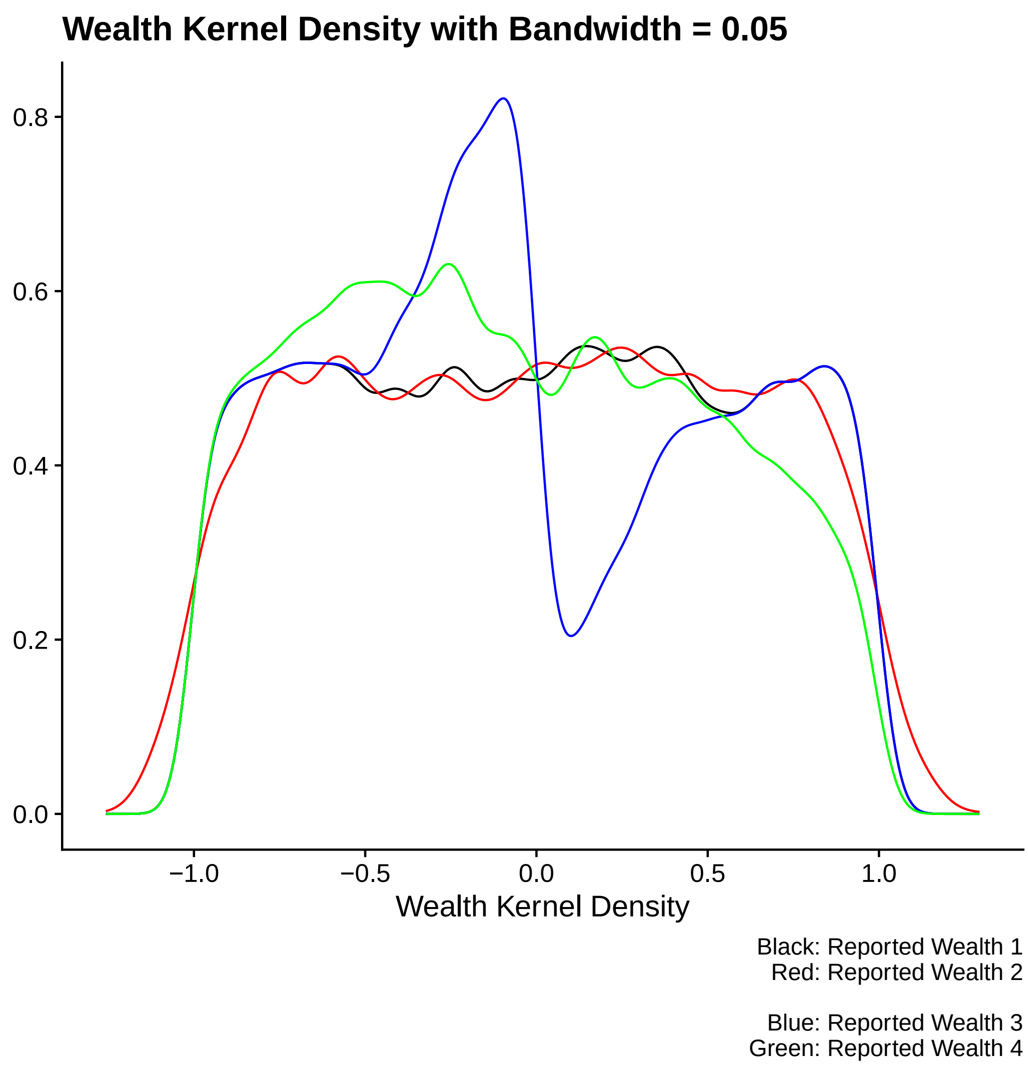


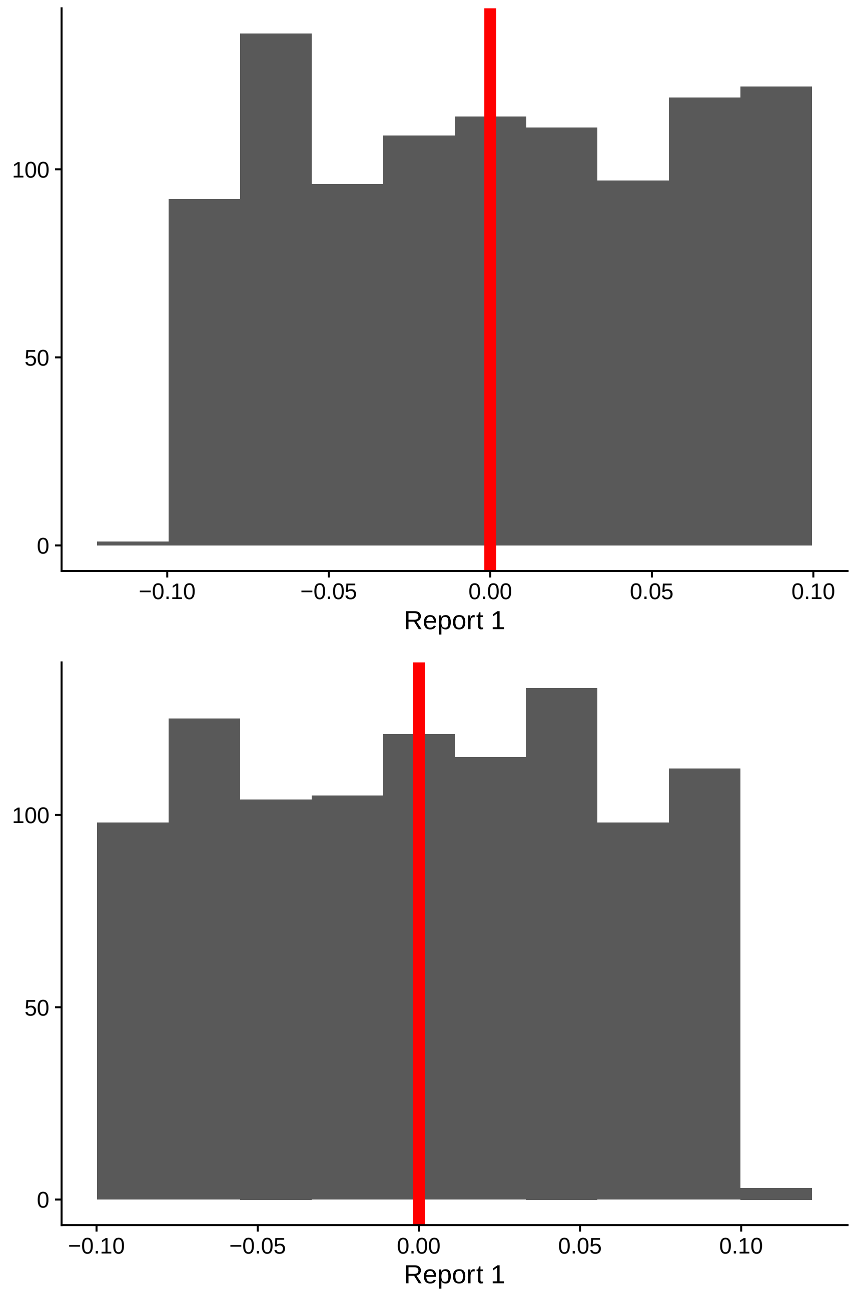
Black: Reported Wealth 1
Red: Reported Wealth 2

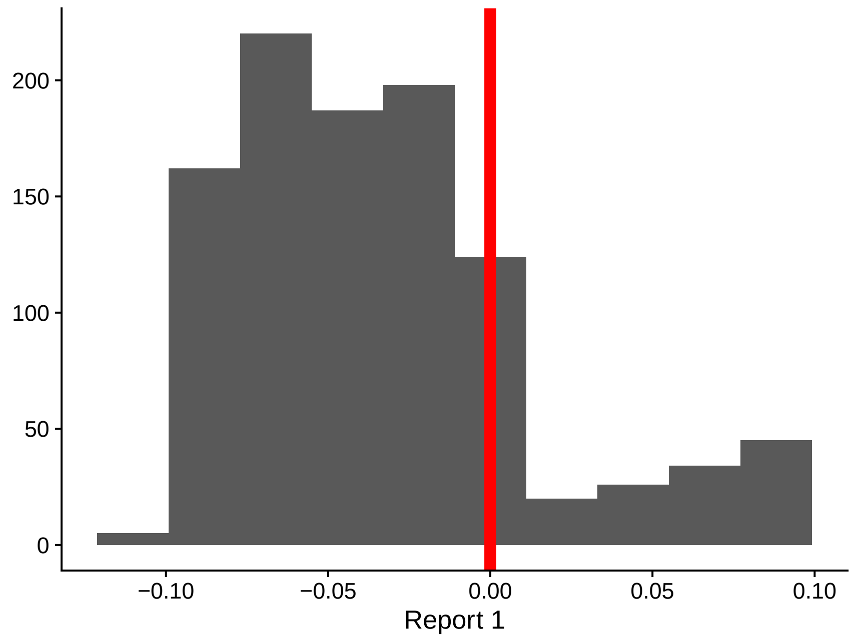
Blue: Reported Wealth 3
Green: Reported Wealth 4

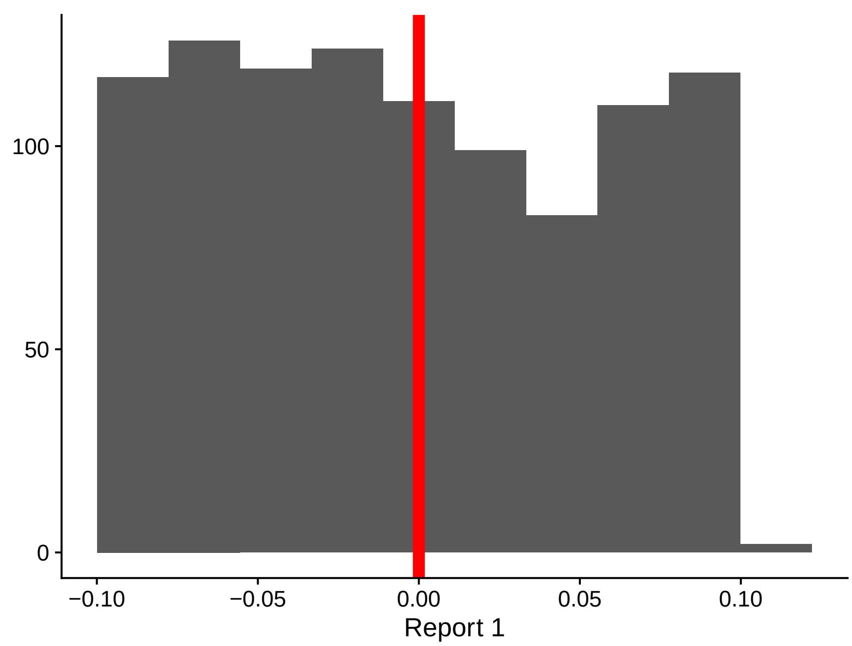
Wealth Kernel Density with Bandwidth = 0.01

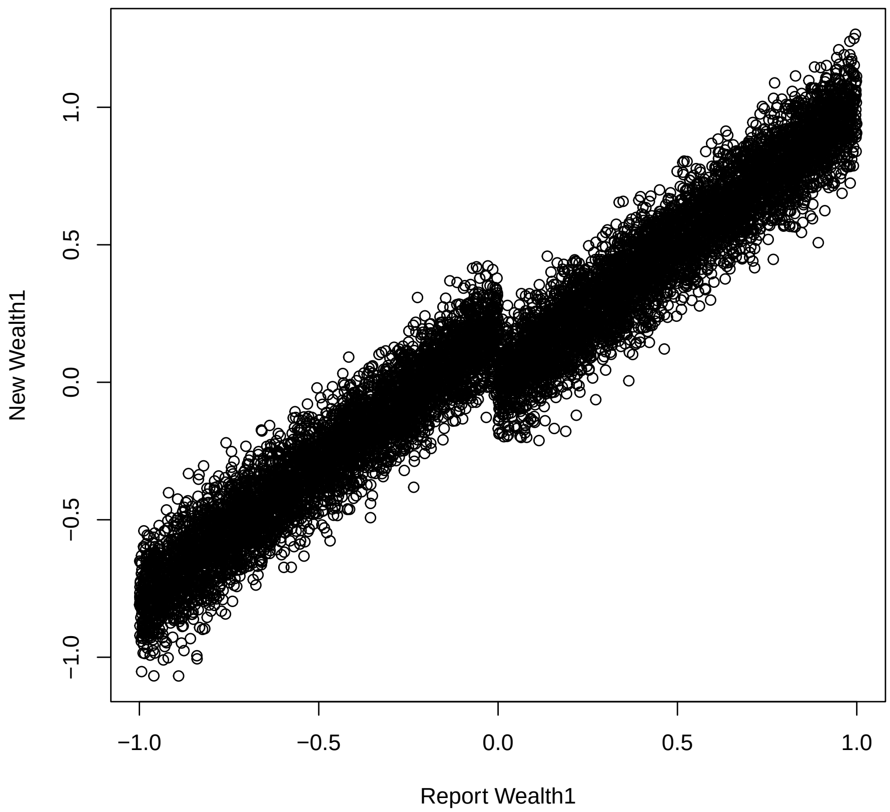



Wealth Kernel Density with Bandwidth = 0.05
0.8
0.6
0.4
0.2
0.0
−1.0
−0.5
0.0
0.5
1.0
Wealth Kernel Density
Black: Reported Wealth 1
Red: Reported Wealth 2
Blue: Reported Wealth 3
Green: Reported Wealth 4

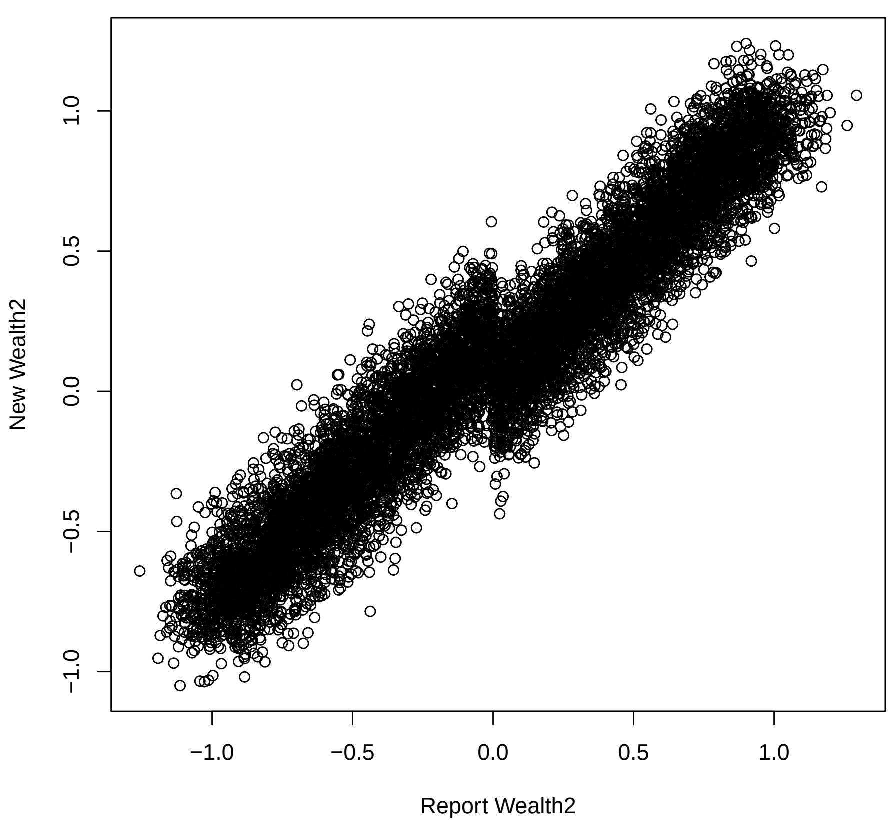

100
50
0
−0.10
−0.05
0.00
0.05
0.10
Report 1
100
50
0
−0.10
−0.05
0.00
0.05
0.10
Report 1


200
150
100
50
0
−0.10
−0.05
0.00
0.05
0.10
Report 1

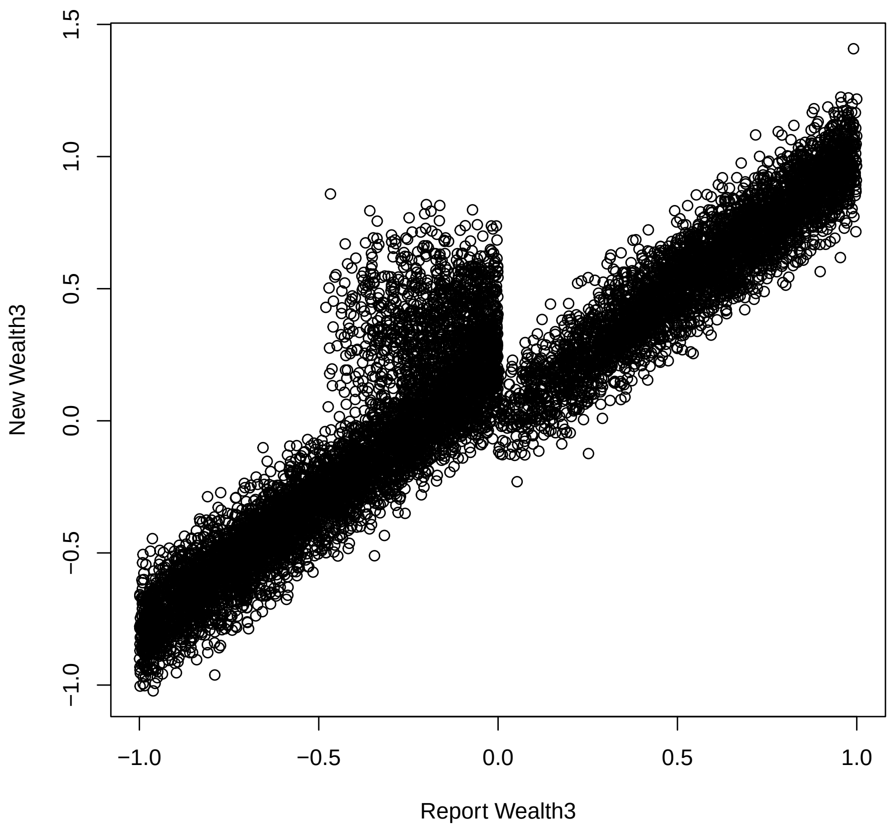

100
50
0
−0.10
−0.05
0.00
0.05
0.10
Report 1

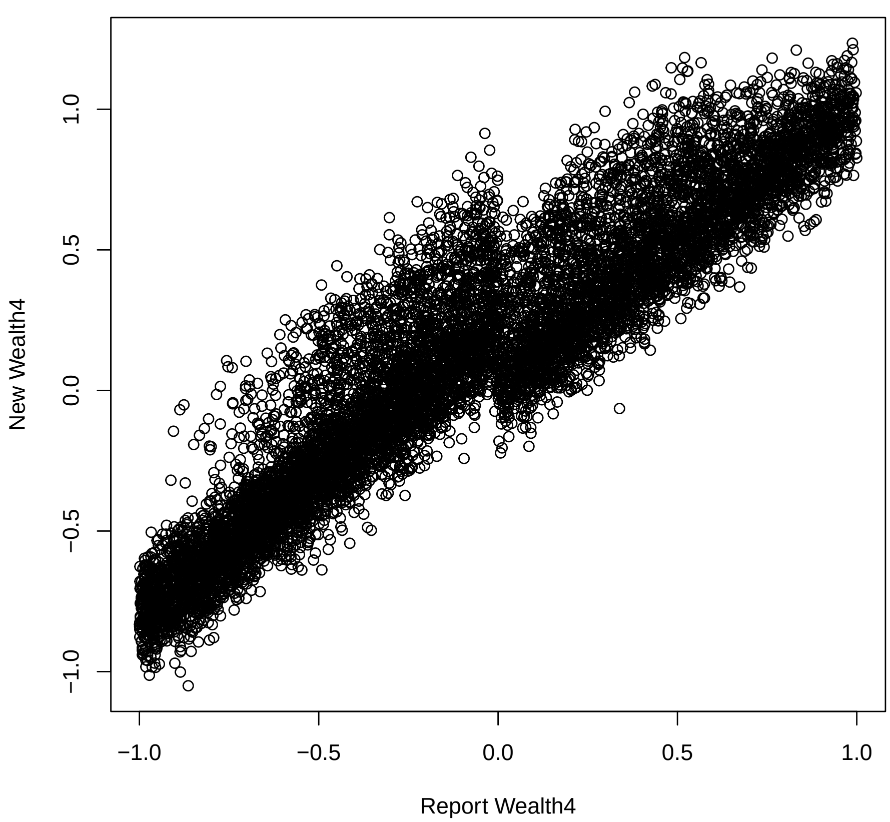


New Wealth1
Report Wealth1
−1.0
−0.5
0.0
0.5
1.0



New Wealth2
Report Wealth2
−1.0
−0.5
0.0
0.5
1.0


New Wealth3
Report Wealth3
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0


New Wealth4
Report Wealth4
1.0
0.5
0.0
−0.5
−1.0
−1.0
−0.5
0.0
0.5
1.0

f)

Below are the results from running the same regression but explicitly using the RDD function. The function seems to be doing something to correct the less pathological forms of measurement error. Scenarios 2 and 4 this time are corrected to the true effect. However, scenario 3 is still not identifying the causal effect we want because of bunching.

Scenario	Treatment Effect
Report 1	-0.201
Report 2	-0.196
Report 3	-0.284
Report 4	-0.208

g)

In part e, we only estimate the true effect in scenario 1. In scenario 2, there is classical measurement error, we we estimate an attenuated effect. In scenario 3, there is bunching around the cutoff, we we don't estimate the true effect at all. In scenario 4, we have non classical measurement error, but not really bunching, so we estimate an attenuated effect, but the attenuation isn't as much as in scenario 2.

In part f, we seem to be estimating the true effect in all scenarios but 3. Scenarios 1, 2, and 4 have no bunching, so the RDD design's continuity assumption holds. When this is true, RDD estimates the true effect. In scenario 3, there is bunching, so continuity fails and we don't esitmate the true effect.