Stefan Pohl

PS 7

Question 1

b) The dist. of birth weights is not very smooth. In particular, the following two characteristics stand out:

i) From the 1-gram-bin hist., it's clear that alot of obs. are concentrated on some specific gram-values. As it turns out, these focal points are gram-quantities that are multiples of 10, or gram-quantities (rounded) that correspond to measurements in ounces.

ii) From the 25-gram-bin-histogram, it's clear that there are "surpisingly" few obs. in the bin with range 1475-1499 grams.

c) We would have to be concerned that our running variable is manipulable, i.e., that there is selection towards one side of the cutoff. In the present case, it would be meaningful that "ambitious"/caring parents would try to downreport the weight of their child in order to get extra care. If there is bunching in the data, then it would be on the left side of the 1500-gram-cutoff.

However, and more plausible, it seems that nurses and doctors intentionally misreport around the cutoff by reporting 1500+ weights when the child in fact is below 1500g, in order not to have to give the extra treatment.

Related to manipulation is local independence, where we would test independence of baseline characteristics, which would have to hold for the RDD to be valid. This will be done in e).

Stefan Pohl

d)

From the Stata code, I get the following jumps at the cutoff (jump from the left to the right)

	bwidth150	bwidth100	bwidth50
bin 1gr	447,1754	740,1982	1611,6
bin 10gr	4279 233	7335,79	11093,7
bin 25gr	13216,73	20013	38534,5

none is significant

Why robust SE?

The dataset is such that higher birthweights are more frequent. Thus, it is natural to assume that there is heteroskedasticity, since frequencies are non-negative (compared to the wages-expl., where we have heterosk. by construction).

Even though there seems to be a jump in all specifications, the jumps are not significant, thus I would say we cannot reject that it's not smooth (meaning that implementing an RDD seems legit).

2)

I used the rd command throughout the whole exercise, in i)-iii), rectangular kernel is added as an option, whereas for iv)-vi) it's the default option (triangular kernel).

Table & Results "jump" at cutoff

		mom_white	mom_ed1
Kernel: rectangular	bw=90	0.0005634 (0.0115876)	0.0035626 (0.0068183)
	bw=60	0.0056396 (0.0151777)	0.0075472 (0.008805)
	bw=30	0.0124496 (0.01987)	-0.0066281 (0.034935)
kernel: triangular	bw=90	0.0048244 (0.0132902)	0.0047352 (0.0079835)
	bw=60	0.0084344 (0.0152047)	0.0019931 (0.009288)
	bw=30	0.0071293 (0.0200868)	-0.0009516 (0.0122499)

Stefan Pohl

The estimates are all very small in magnitude, and not significant.
I would conclude that the RD design <u>passes</u> the <u>balanced covariates</u> test.

f)

When considering the outcome dummy white yes/no, most ounce heaps as well as the 1500-gr-heaps feature significantly different outcomes.

With mom_ed1, i.e. dummy "less than highschool" yes/no, this is less pronounced; it seems to be mainly the 1500-gr-heaps that have a significant impact on mom_ed1.

g) I repeat all of the $\underbrace{3}_{\substack{\text{bwidth} \\ 30,60,90}} \cdot \underbrace{2}_{\substack{\text{kernel:} \\ \text{rect./triang.}}} = 6$ specifications from e), for both robust SEs / cluster SEs.

⇒ Effect on 1-yr.-mortality:
While all estimates are significant ("positive" jump from left to right side of cutoff, i.e. positive effect of "extra care"-treatment), they differ substantially, from 0.0074272 (for bw = 90, kernel = rect.) to 0.026285 (for bw = 30, kernel = triangular).
This might be evidence that est. is quite sensitive to specification.

Furthermore, SE are considerably bigger when clustering on grams. This suggests that there are indeed clusters, and that one <u>should</u> use clustered SEs (on grams).

→ Hospitals in poorer areas measure less precisely, are more inclined to round to certain numbers (as multiples of 10), or simply have a preference for, say, reporting multiples of 10.
If this is correlated with general health conditions of the children (would expect the children to be less healthy in less privileged areas), and that there are many obs. from the same hospitals, then these are not indep. but have a related error structure. → Should [illegible]

Stefan Pohl

h) Code repeats all steps of g), dropping 1500gr.

→ When dropping the 1500s, the RD estimates are considerably lower for <u>all</u> specifications.

No, RD should never be sensitive to dropping obs. exactly at the cutoff, since this would suggest that these is bunching (non-random sorting). Since the weight scale is "finely granulated", the RD estimate for 1499 ↔ 1500 should hardly be different from the RD estimate, 1499 ↔ 1501, but yet, this seems to be the case here.

When dropping the obs. measured in ounces, the RD estimates are much larger, in fact as large as 0.0473915 (from left to right of cutoff) when using a 30-bandwidth, and triangular kernel.

Note: About the kernel, triangular seems to be better for minimizing mean squared error, and it is intuitively appealing (more weight for closer obs.). Further, I would cluster on bweight.

⇒ Main Spec: bandwidth 60, triangular, drop if bweight = 1500, cluster (bweight), see regression fct. for specification.

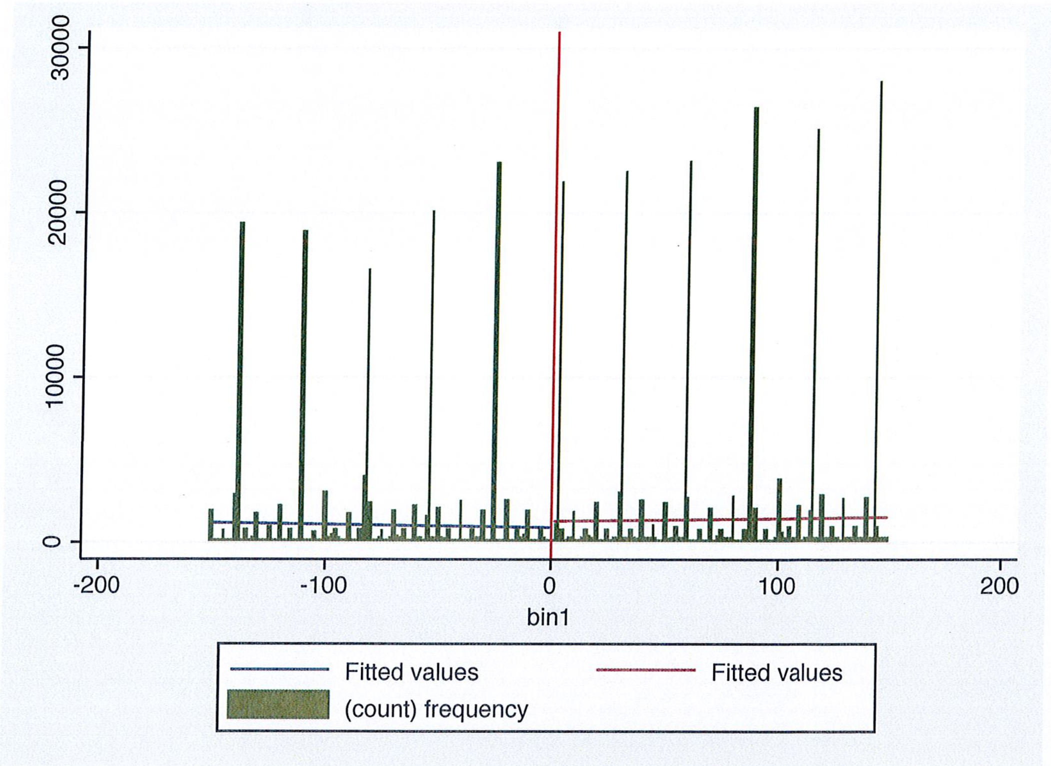
I conclude that the effect of very low bweight classification has a decreasing / negative effect on 1-year-mortality of ~1.012 pp. (which is a relatively large effect).

Considering the other estimates of the alternative specifications, this seems almost a conservative estimate - by the arguments highlighted before, I would trust it the most.

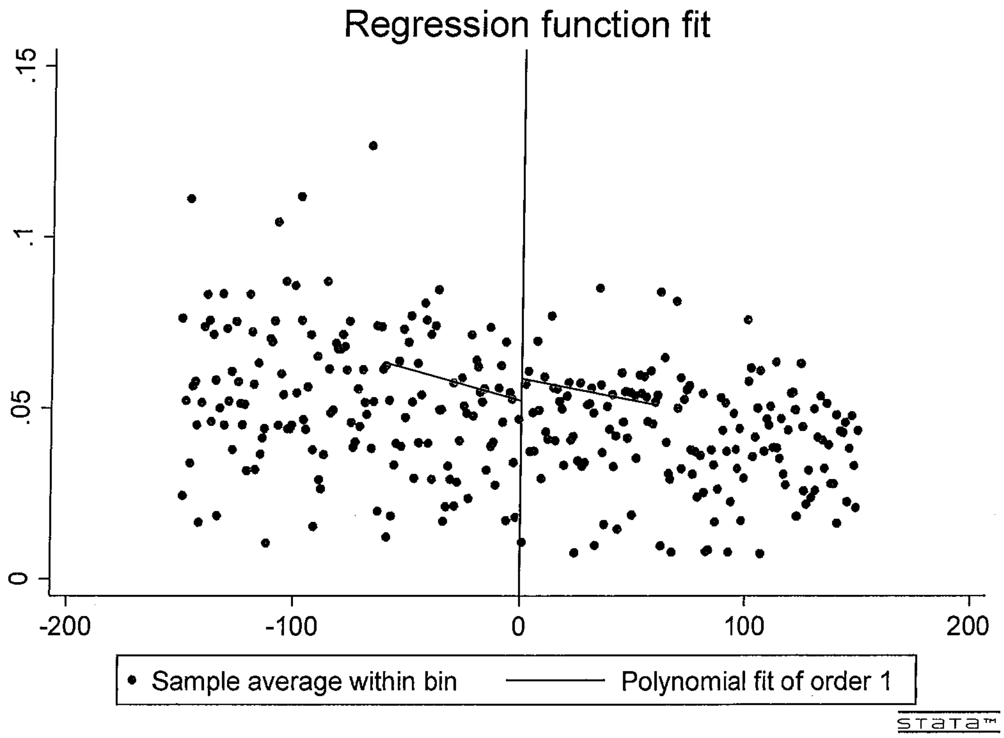
In conclusion however, I would say that there is not very robust evidence for large effects.

i) I'm most inclined to believe the results when dropping the obs. at 1500gr, using triangular kernel. Further, since we have so many obs., I would tend to use a smaller bandwidth, since that seems to minimize the MSE (by limiting the bias, variance is relatively low, anyway)

7.1) d) 150gr bandwidth, 1gr-bins (other graphs "similar", see do-file)



7.1) i) Main Specification: Bw = 60, drop 1500gr, triangular, cluster bweight



```
use "/Users/stefan/Desktop/FS19/Econometrics II -
PhD/PS7/runandjump_sample1500g.dta"

*a)
gen bwtcent=bweight-1500

*b)
gen bin1=bwtcent+0.5
hist bin1, width(1) start(-150) freq

gen modulo10=mod(bwtcent,10)
gen bin2=bwtcent+(5-modulo10)
hist bin2, width(10) start(-150) freq

gen modulo25=mod(bwtcent,25)
gen bin3=bwtcent+(12.5-modulo25)
hist bin3, width(25) start(-150) freq

*d)
*************************
******150 bandwidth******
*************************
***150 bandwidth, 1gr bin
preserve
gen frequency=1
collapse (count) frequency, by(bin1)
*interact bweigth with above cutoff
gen bin1_interact=0
replace bin1_interact=bin1 if bin1<0
gen treatment=0
replace treatment=1 if bin1<0
drop if bin1<-150
drop if bin1>150
***regression***
reg frequency bin1 bin1_interact treatment, robust
predict hat
twoway (line hat bin1 if bin1<0) (line hat bin1 if bin1>=0) (bar
frequency bin1, xline(0, lwidth(vthin)))
restore

***150 bandwidth, 10gr bin
preserve
gen frequency=1
collapse (count) frequency, by(bin2)
*interact bweigth with above cutoff
gen bin2_interact=0
replace bin2_interact=bin2 if bin2<0
gen treatment=0
replace treatment=1 if bin2<0
drop if bin2<-150
drop if bin2>150
***regression
```

```
reg frequency bin2 bin2_interact treatment, robust
predict hat
twoway (line hat bin2 if bin2<0) (line hat bin2 if bin2>=0) (bar
frequency bin2, xline(0, lwidth(vthin)))
restore

***150 bandwidth, 25gr bin
preserve
gen frequency=1
collapse (count) frequency, by(bin3)
*interact bweigth with above cutoff
gen bin3_interact=0
replace bin3_interact=bin3 if bin3<0
gen treatment=0
replace treatment=1 if bin3<0
drop if bin3<-150
drop if bin3>150
***regression
reg frequency bin3 bin3_interact treatment, robust
predict hat
twoway (line hat bin3 if bin3<0) (line hat bin3 if bin3>=0) (bar
frequency bin3, xline(0, lwidth(vthin)))
restore

*************************
******100 bandwidth******
*************************
***100 bandwidth, 1gr bin
preserve
gen frequency=1
collapse (count) frequency, by(bin1)
*interact bweigth with above cutoff
gen bin1_interact=0
replace bin1_interact=bin1 if bin1<0
gen treatment=0
replace treatment=1 if bin1<0
drop if bin1<-100
drop if bin1>100
***regression***
reg frequency bin1 bin1_interact treatment, robust
predict hat
twoway (line hat bin1 if bin1<0) (line hat bin1 if bin1>=0) (bar
frequency bin1, xline(0, lwidth(vthin)))
restore

***100 bandwidth, 10gr bin
preserve
gen frequency=1
collapse (count) frequency, by(bin2)
*interact bweigth with above cutoff
gen bin2_interact=0
replace bin2_interact=bin2 if bin2<0
```

```
gen treatment=0
replace treatment=1 if bin2<0
drop if bin2<-100
drop if bin2>100
***regression***
reg frequency bin2 bin2_interact treatment, robust
predict hat
twoway (line hat bin2 if bin2<0) (line hat bin2 if bin2>=0) (bar
frequency bin2, xline(0, lwidth(vthin)))
restore

***100 bandwidth, 25gr bin
preserve
gen frequency=1
collapse (count) frequency, by(bin3)
*interact bweigth with above cutoff
gen bin3_interact=0
replace bin3_interact=bin3 if bin3<0
gen treatment=0
replace treatment=1 if bin3<0
drop if bin3<-100
drop if bin3>100
***regression
reg frequency bin3 bin3_interact treatment, robust
predict hat
twoway (line hat bin3 if bin3<0) (line hat bin3 if bin3>=0) (bar
frequency bin3, xline(0, lwidth(vthin)))
restore

************************
******50 bandwidth******
************************
***50 bandwidth, 1gr bin
preserve
gen frequency=1
collapse (count) frequency, by(bin1)
*interact bweigth with above cutoff
gen bin1_interact=0
replace bin1_interact=bin1 if bin1<0
gen treatment=0
replace treatment=1 if bin1<0
drop if bin1<-50
drop if bin1>50
***regression***
reg frequency bin1 bin1_interact treatment, robust
predict hat
twoway (line hat bin1 if bin1<0) (line hat bin1 if bin1>=0) (bar
frequency bin1, xline(0, lwidth(vthin)))
restore

***50 bandwidth, 10gr bin
preserve
```

```
gen frequency=1
collapse (count) frequency, by(bin2)
*interact bweigth with above cutoff
gen bin2_interact=0
replace bin2_interact=bin2 if bin2<0
gen treatment=0
replace treatment=1 if bin2<0
drop if bin2<-50
drop if bin2>50
***regression***
reg frequency bin2 bin2_interact treatment, robust
predict hat
twoway (line hat bin2 if bin2<0) (line hat bin2 if bin2>=0) (bar 
frequency bin2, xline(0, lwidth(vthin)))
restore

***50 bandwidth, 25gr bin
preserve
gen frequency=1
collapse (count) frequency, by(bin3)
*interact bweigth with above cutoff
gen bin3_interact=0
replace bin3_interact=bin3 if bin3<0
gen treatment=0
replace treatment=1 if bin3<0
drop if bin3<-50
drop if bin3>50
***regression
reg frequency bin3 bin3_interact treatment, robust
predict hat
twoway (line hat bin3 if bin3<0) (line hat bin3 if bin3>=0) (bar 
frequency bin3, xline(0, lwidth(vthin)))
restore


*e(discontinutiy in covariates?)
********************************
***Mom white******
***Mom educated***
***i, ii, iii*****
******************

ssc install rd, replace

gen mom_white=0
replace mom_white=1 if mom_race==1
gen treatment=bwtcent<0

rd mom_white bwtcent, z0(0) bwidth(90) kernel(rectangle) cluster(
bwtcent)
rd mom_white bwtcent, z0(0) bwidth(60) kernel(rectangle) cluster(
bwtcent)
```

```
rd mom_white bwtcent, z0(0) bwidth(30) kernel(rectangle) cluster(
bwtcent)

rd mom_ed1 bwtcent, z0(0) bwidth(90) kernel(rectangle) cluster(
bwtcent)
rd mom_ed1 bwtcent, z0(0) bwidth(60) kernel(rectangle) cluster(
bwtcent)
rd mom_ed1 bwtcent, z0(0) bwidth(30) kernel(rectangle) cluster(
bwtcent)


***iv, v, vi***********
***triangular kernel***
***(default)***********

rd mom_white bwtcent, z0(0) bwidth(90) cluster(bwtcent)
rd mom_white bwtcent, z0(0) bwidth(60) cluster(bwtcent)
rd mom_white bwtcent, z0(0) bwidth(30) cluster(bwtcent)

rd mom_ed1 bwtcent, z0(0) bwidth(90) cluster(bwtcent)
rd mom_ed1 bwtcent, z0(0) bwidth(60) cluster(bwtcent)
rd mom_ed1 bwtcent, z0(0) bwidth(30) cluster(bwtcent)


*f)
gen ounce_multiple=0
replace ounce_multiple=51 if floor(51*28.3495231)==bweight
replace ounce_multiple=52 if floor(52*28.3495231)==bweight
replace ounce_multiple=53 if floor(53*28.3495231)==bweight
replace ounce_multiple=54 if floor(54*28.3495231)==bweight

replace ounce_multiple=51 if ceil(51*28.3495231)==bweight
replace ounce_multiple=52 if ceil(52*28.3495231)==bweight
replace ounce_multiple=53 if ceil(53*28.3495231)==bweight
replace ounce_multiple=54 if ceil(54*28.3495231)==bweight

gen ounce51_dummy=ounce_multiple==51
gen ounce52_dummy=ounce_multiple==52
gen ounce53_dummy=ounce_multiple==53
gen ounce54_dummy=ounce_multiple==54

***look at jump off trend line at 51, 52, 53, 54 ounces, and
1500gr
***each with bandwidth of 25 and 100, respectively, for
mom_white and mom_ed1

****************
***mom_white****
****************

***Bandwidth 25
preserve
```

```
drop if bweight<51*28.3495231-25
drop if bweight>51*28.3495231+25
*generate ounce51_interact=0
*replace ounce51_interact=bwtcent if bweight>51*28.3495231
reg mom_white ounce51_dummy bwtcent

restore
preserve

drop if bweight<52*28.3495231-25
drop if bweight>52*28.3495231+25
reg mom_white ounce52_dummy bwtcent

restore
preserve

drop if bweight<53*28.3495231-25
drop if bweight>53*28.3495231+25
drop if bweight==1500
reg mom_white ounce53_dummy bwtcent

restore
preserve

drop if bweight<54*28.3495231-25
drop if bweight>54*28.3495231+25
reg mom_white ounce54_dummy bwtcent

restore
preserve

drop if bweight<1500-25
drop if bweight<1500+25
drop if ounce53_dummy==1
drop if ounce52_dummy==1
gen gr1500_dummy=bweight==1500
reg mom_white gr1500_dummy bwtcent

restore

***Bandwidth 100
***dropping other ounce heaps, and 1500gr observations
preserve
drop if bweight==1500|ounce52_dummy==1|ounce53_dummy==1|
ounce54_dummy==1
drop if bweight<51*28.3495231-100
drop if bweight>51*28.3495231+100

reg mom_white ounce51_dummy bwtcent
restore

```

```
preserve
drop if bweight==1500|ounce51_dummy==1|ounce53_dummy==1|
ounce54_dummy==1
drop if bweight<52*28.3495231-100
drop if bweight>52*28.3495231+100

reg mom_white ounce52_dummy bwtcent
restore

preserve
drop if bweight==1500|ounce52_dummy==1|ounce51_dummy==1|
ounce54_dummy==1
drop if bweight<53*28.3495231-100
drop if bweight>53*28.3495231+100

reg mom_white ounce53_dummy bwtcent
restore

preserve
drop if bweight==1500|ounce52_dummy==1|ounce51_dummy==1|
ounce53_dummy==1
drop if bweight<54*28.3495231-100
drop if bweight>54*28.3495231+100

reg mom_white ounce54_dummy bwtcent
restore

preserve
gen gr1500_dummy=bweight==1500
drop if ounce54_dummy==1|ounce52_dummy==1|ounce51_dummy==1|
ounce53_dummy==1
drop if bweight<1500-100|bweight>1500+100

reg mom_white gr1500_dummy bwtcent
restore


****************
***mom_ed1******
****************

***Bandwidth 25
preserve

drop if bweight<51*28.3495231-25
drop if bweight>51*28.3495231+25
*generate ounce51_interact=0
*replace ounce51_interact=bwtcent if bweight>51*28.3495231
reg mom_ed1 ounce51_dummy bwtcent

restore
preserve
```

```
drop if bweight<52*28.3495231-25
drop if bweight>52*28.3495231+25
reg mom_ed1 ounce52_dummy bwtcent

restore
preserve

drop if bweight<53*28.3495231-25
drop if bweight>53*28.3495231+25
drop if bweight==1500
reg mom_ed1 ounce53_dummy bwtcent

restore
preserve

drop if bweight<54*28.3495231-25
drop if bweight>54*28.3495231+25
reg mom_ed1 ounce54_dummy bwtcent

restore
preserve

drop if bweight<1500-25
drop if bweight<1500+25
drop if ounce53_dummy==1
drop if ounce52_dummy==1
gen gr1500_dummy=bweight==1500
reg mom_ed1 gr1500_dummy bwtcent

restore

***Bandwidth 100
***dropping other ounce heaps, and 1500gr observations

preserve
drop if bweight==1500|ounce52_dummy==1|ounce53_dummy==1|
ounce54_dummy==1
drop if bweight<51*28.3495231-100
drop if bweight>51*28.3495231+100

reg mom_ed1 ounce51_dummy bwtcent
restore

preserve
drop if bweight==1500|ounce51_dummy==1|ounce53_dummy==1|
ounce54_dummy==1
drop if bweight<52*28.3495231-100
drop if bweight>52*28.3495231+100

reg mom_ed1 ounce52_dummy bwtcent
restore
```

```
preserve
drop if bweight==1500|ounce52_dummy==1|ounce51_dummy==1|
ounce54_dummy==1
drop if bweight<53*28.3495231-100
drop if bweight>53*28.3495231+100

reg mom_ed1 ounce53_dummy bwtcent
restore

preserve
drop if bweight==1500|ounce52_dummy==1|ounce51_dummy==1|
ounce53_dummy==1
drop if bweight<54*28.3495231-100
drop if bweight>54*28.3495231+100

reg mom_ed1 ounce54_dummy bwtcent
restore

preserve
gen gr1500_dummy=bweight==1500
drop if ounce54_dummy==1|ounce52_dummy==1|ounce51_dummy==1|
ounce53_dummy==1
drop if bweight<1500-100|bweight>1500+100

reg mom_ed1 gr1500_dummy bwtcent
restore


*g)
******Robust*******
***same as in e)***
***rect. kernel****
rd agedth5 bwtcent, z0(0) bwidth(90) kernel(rectangle) robust
rd agedth5 bwtcent, z0(0) bwidth(60) kernel(rectangle) robust
rd agedth5 bwtcent, z0(0) bwidth(30) kernel(rectangle) robust

***triangular kernel
******(default)*****
rd agedth5 bwtcent, z0(0) bwidth(90) robust
rd agedth5 bwtcent, z0(0) bwidth(60) robust
rd agedth5 bwtcent, z0(0) bwidth(30) robust


******Cluster******
***same as in e)***
***rect. kernel****
rd agedth5 bwtcent, z0(0) bwidth(90) kernel(rectangle) robust
cluster(bwtcent)
rd agedth5 bwtcent, z0(0) bwidth(60) kernel(rectangle) robust
cluster(bwtcent)
rd agedth5 bwtcent, z0(0) bwidth(30) kernel(rectangle) robust
```

```
cluster(bwtcent)

***triangular kernel
******(default)*****
rd agedth5 bwtcent, z0(0) bwidth(90) robust cluster(bwtcent)
rd agedth5 bwtcent, z0(0) bwidth(60) robust cluster(bwtcent)
rd agedth5 bwtcent, z0(0) bwidth(30) robust cluster(bwtcent)


*h)
***drop 1500gr obs.
***repeat exercise g) (same specs.)
preserve
drop if bweight==1500

***Robust***
*same as in e)
*rect. kernel
rd agedth5 bwtcent, z0(0) bwidth(90) kernel(rectangle) robust
rd agedth5 bwtcent, z0(0) bwidth(60) kernel(rectangle) robust
rd agedth5 bwtcent, z0(0) bwidth(30) kernel(rectangle) robust

*triangualr kernel
rd agedth5 bwtcent, z0(0) bwidth(90) robust
rd agedth5 bwtcent, z0(0) bwidth(60) robust
rd agedth5 bwtcent, z0(0) bwidth(30) robust

***Cluster
*same as in e)
*rect. kernel
rd agedth5 bwtcent, z0(0) bwidth(90) kernel(rectangle) robust
cluster(bwtcent)
rd agedth5 bwtcent, z0(0) bwidth(60) kernel(rectangle) robust
cluster(bwtcent)
rd agedth5 bwtcent, z0(0) bwidth(30) kernel(rectangle) robust
cluster(bwtcent)

***triangular kernel
******(default)*****
rd agedth5 bwtcent, z0(0) bwidth(90) robust cluster(bwtcent)
rd agedth5 bwtcent, z0(0) bwidth(60) robust cluster(bwtcent)
rd agedth5 bwtcent, z0(0) bwidth(30) robust cluster(bwtcent)

restore

******drop obs. measured in ounces
***repeat exercise g) (same specs.)
preserve
drop if ounce==1

***Robust***
*same as in e)
```

```
*rect. kernel
rd agedth5 bwtcent, z0(0) bwidth(90) kernel(rectangle) robust
rd agedth5 bwtcent, z0(0) bwidth(60) kernel(rectangle) robust
rd agedth5 bwtcent, z0(0) bwidth(30) kernel(rectangle) robust

***triangular kernel
rd agedth5 bwtcent, z0(0) bwidth(90) robust
rd agedth5 bwtcent, z0(0) bwidth(60) robust
rd agedth5 bwtcent, z0(0) bwidth(30) robust

***Cluster
*same as in e)
*rect. kernel
rd agedth5 bwtcent, z0(0) bwidth(90) kernel(rectangle) cluster(bwtcent)
rd agedth5 bwtcent, z0(0) bwidth(60) kernel(rectangle) cluster(bwtcent)
rd agedth5 bwtcent, z0(0) bwidth(30) kernel(rectangle) cluster(bwtcent)

***triangular kernel
******(default)*****
rd agedth5 bwtcent, z0(0) bwidth(90) cluster(bwtcent)
rd agedth5 bwtcent, z0(0) bwidth(60) cluster(bwtcent)
rd agedth5 bwtcent, z0(0) bwidth(30) cluster(bwtcent)

*when clustering on bweight, SEs considerably larger. indicates that we should
*cluster on bweight

restore


*i)
*in light of exercise h), I would prefer a specification with clustering on
*bweight. Further, I would use the triangular kernel because whilst considering
*possibly more data points, the fact that it down-weights obs. far away means
*that the bias tend to be smaller.
*Next, I'll use the suggested bandwidth of 60 and compare the results of RD
*with triangular kernel, cluster (bwtcent)
*for both keeping/dropping the obs. of 1500gr

ssc install rdrobust
preserve
drop if bweight==1500
rd agedth5 bwtcent, z0(0) bwidth(60) cluster(bwtcent)
rdplot agedth5 bwtcent, p(1) z0(0) h(60) cluster(bwtcent)
restore
```

Stefan Pohl

Question 2

a) Cultural background is the treatment, i.e., whether the mother is Swedish or of foreign origin (in the 2nd step, could also differentiate between different countries of origin of the immigrants).

The outcome is birth-spacing, or more generally fertility. In a broader sense, also economic outcomes such as school performance of the child could be considered an outcome.

b) Unmeasured immigrant characteristics.
The immigrants are a non-representative subgroup of their countries of origin, since there was some sort of self-selection going on, which resulted in migrating to Sweden.
For example, I would expect that immigrants tend to be more liberal, ambitious, educated than the "average", representative resident in their country of origin.
This is a problem for identification, in particular for the epidemiological approach to culture, and also for the DiD. It affects moreover the external validity. Also, prior to the policy change, Swedish natives and immigrants possibly faced different incentives (possibly institutional differences).

c) Their "DiD" approach likely assumes that after the policy change, both natives / immigrants face the same circumstances and have access to the same level of "opportunities".
The crucial assumption is that any change, other than the policy change, that influences the outcome, affects both groups the same / equally.
This is the common / parallel trend assumption.

Stefan Pohl

d)

Although the parallel trends assumption cannot be tested, we can compare the pre-policy periods for internal validity.

→ Parallel trends, both flat and parallel for both groups, thus internally validity satisfied.

External validity: cannot tell whether or not satisfied.

e)

I think the general problem here is about external validity. Some immigrants might self-select into migrating to Sweden. This could make them fundamentally different.

Furthermore, and more generally, it might be that differences in birth-spacing between the two groups <u>prior</u> to the changes affect how "responsive" the birth-spacing-durations are to the fertility policy. Could divide the sample into immigrants closer to Sweden as country of origin vs. countries further away (Arabian Peninsula), and redo the DiD for Swedish vs. "close proximity to Sweden" immigrants.

Additionally, another concern for inference could also be that the composition of migrants over time changes, possibly induced by policy).