Econometrics PS 4

Himani Pasricha (20201143)
Dimitrios Argyros (20201790)
Tatiana Bezdenezhnykh (20201138)
Matthew Abagna(20201136)
Manuel Rodriguez(20201144)

March 2021

1 Problem Set 4: RDD

1. Following a similar approach as [Nekoei and Weber, 2017] and/or [Caliendo, 2012], employing a Regression Discontinuity (RD) design that allows us to properly analyze the age cut-off for eligibility for a nine-week extension of the potential unemployment benefit in addition to the base of 30 weeks, and thus its causal effect on wages/job quality after re-employment. The setting that allows for variation of the benefit duration entitlement is based on the age that a worker enters unemployment, in essence allows the researcher to compare people just below the threshold and just above the the threshold. In particular, upon achieving the age of 40 by an unemployed person, the duration of unemployment benefits increases from 30 to 39 weeks).

   In order to undertake this analysis, one has to assume a random assignment to treatment for both treated and control group. This assumption is crucial for the control group to be a good counterfactual to the treatment group since we never observe either in both states. Henceforth, it can be assumed that the assignment to this specific treatment, i.e. the discontinuous jump in benefit entitlement, is a deterministic function of the age itself. E$[y_{0i}|D_i = 1]$ - E$[y_{0i}|D_i = 0] = 0$ holds.

   As such, the main identification assumptions that being entitled to extended benefits (assignment to treatment) is only determined by age and is orthogonal to all other heterogeneity. There is no way to test that assumption in its entirety, what a researcher can do is to show by using balancing tests that individuals above and below cut-off are homogeneous with respect to observables (something we examine in question 3). A RD analysis as employed in our example, suffers from two potential selection issues, identified by [Caliendo, 2012]. Firstly, manipulation of the treatment status may occur through unobservable and unmeasurable factors such as motivation or dishonesty. In essence, if recipients misreport their age to get a higher benefit duration, then RD analysis is invalided and does not capture the causal effect of interest. Following the strategy of [Nekoei and Weber, 2017], we examine in question 3 the distribution of lay-offs around the age of 40. If individuals or their employers respond strategically to the policy we will see bunching of layoffs above the threshold. That will provide us with an indication of presence or absence of manipulation around the cut-off. Secondly, even in the case that assignment to treatment at the beginning of unemployment is based on this sharp discontinuity, there might be a case of selection in the resulting sample of the re-employed, as we observe only individuals that are re-employed, and that selection is based on observed and unobserved characteristics [Ham, 1996]. This problem can be addressed according to [Caliendo, 2012] using a bivariate discrete-time hazard rate model jointly with wages and allowing for potentially correlated unobserved heterogeneity.

Besides, a Regression Discontinuity setting allows us to capture the causal effect by distinguishing the non-linear and discontinuous function, $1(x_i \geq x_0)$, where x_i is equal to the age of each individual and x_0 is the cutoff level, from the smooth and (in this specific case) linear function, x_i.

Moreover, another assumption needed for a RD framework and which allow us to employ this technique is the Continuity Assumption. It implies that in the absence of treatment, there would be no jump in the expected potential outcomes. They would be smooth functions of X. In this specific setting, we know that an unemployed person receives unemployment benefits for 30 weeks. Therefore, continuous relationship between the duration of unemployment benefits and the quality of the job after starting to work again can be assumed.

The mechanisms behind the effects of differences in the duration of benefits on the quality of the job after finding employment comes from differences in optimal search behaviour that in turn lead to different distributions of job-match quality [Ehrenberg and Oaxaca, 1976]. As unemployed workers become less selective when they are approaching benefit expiration, differences in the duration of benefits affect the quality of the jobs accepted. Also, individuals with the same level of unemployment and same benefits who are over the cut-off age and thus have a longer remaining benefit entitlement duration have the option to wait for jobs that are of a better quality [Caliendo, 2012]. At the same time, a higher benefit entitlement duration might lead to a reduction in the intensity of the search effort leading to longer duration of non-employment and lower quality of jobs, as well as wages after re-employment [Schmieder and Bender, 2016].

2. In a Sharp RD design, treatment is a deterministic function of the running variable x_i. In other words, if we know x_i, we know D_i. In this scenario, the running variable age is the co-factor which strictly determines the unemployment benefit duration, since the potential benefit duration is 9 weeks longer for agents older than 40, therefore a Sharp RD design could be employed. Using that framework we estimate the reduced form and implicitly assume perfect compliance to the policy that causes the discontinuity.

   Following [Angrist and Pischke, 2009], a Sharp RD framework estimation can be stated as the following:

   $\mathrm{Y}_i = \alpha + \beta_{01}\tilde{x}_i + \beta_{02}\tilde{x}_i^2 + ... + \beta_{0p}\tilde{x}_i^p + \rho D_i + \beta_1^* D_i\tilde{x}_i + \beta_2^* D_i\tilde{x}_i^2 + ... + \beta_p^* D_i\tilde{x}_i^p + \eta_i$, with $\tilde{x}_i \equiv \mathrm{x}_i - x_0$.

   In this specific framework, our estimation will thus be the following:

   $\mathrm{wage}_i = \alpha + \beta_0 * age_i + \beta_1 * age_i^2 + \rho * 1(age_i \geq 40) + \gamma_0 * 1(age_i \geq 40) * age_i + \gamma_1 * 1(age_i \geq 40) * age_i^2 + \epsilon_i$.

   ρ is the treatment effect parameter, which represents the causal effect of a increase in benefit duration on earnings for individual i once it becomes employed, while $1(age_i \geq 40)$ is the binary indicator for the running variable age_i equal to one if person i is at least 40 years old. Besides, centered regressors are employed by subtracting the cut-off level 40 from the age_i polynomials. Since the error term is assumed to not change discontinuously at the age of 40, ρ, thus, provides an unbiased estimate of the causal effect of an unemployment benefit extension even in the absence of controls for other observable factors. This implies that other co-factors which may affect the outcome variable wage do not change discontinuously at the cutoff.

   As stated before, the Continuity Assumption is needed in order to implement this framework. The main reason to argue that this assumption is satisfied is that respondents cannot manipulate the cutoff point and self-select into treatment, for example by lying about their age. Henceforth, the local randomization assumption is not undermined and differences across the cutoff are only related to the treatment itself.

   However, note that this estimation will give us the average causal effect of the treatment as running variable approaches the cutoff in the limit, for it is only in the limit that we have overlap = LATE. This implies that we only identify an average causal effect for those units at the cutoff, implying no common support and relying on extrapolation for the estimation.

3. The idea behind these graphs is to graphically show if there is a presence of any discontinuity in the outcome variable at the cutoff for those individuals who are losing their job prior to the age of 40, and examine as well whether predicted outcomes evolve smoothly with respect to age. If our RD design is correct, we expect no discontinuities around the wage cutoff. Moreover, one may be concerned about the presence of strategic timing of layoff.

Therefore, looking for evidence of bunching in the frequency of layoffs around the 40-year age threshold in order to detect a sign of strategic timing of layoff is needed to validate the identification assumption.
Figure 1 and Figure 2 thus provide two of several RD validity tests. First, note that the dashed vertical line denotes the cutoff for unemployment benefit eligibility extension from 30 to 39 weeks at the cutoff, while the solid lines represent quadratic fits on each side of the age threshold. Figure 1 plots the density of age at layoff with bin size 4 months. As we can see in the figure, the distribution evolves smoothly over the threshold, thus not pointing to the presence of a strategic layoff behavior. Figure 2 show how the natural logarithm of previous wage against the age at layoff. It can be inferred that predetermined observables evolve smoothly around the 40-year threshold, thus supporting, as well as the results in Figure 1, the validity of the RD design.

Figure 1: Distribution of Ages

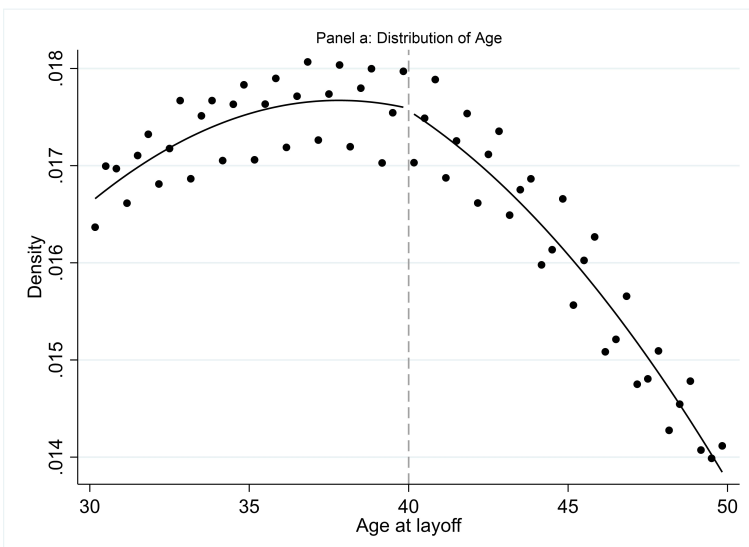
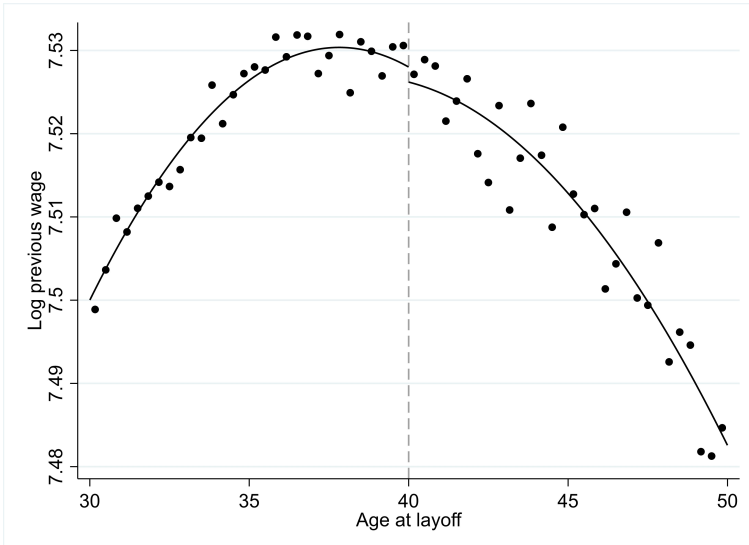


Figure 2: Log(Previous Wage) around threshold



4. Figure 3 plots the average non-employment duration (time to next job) against age at layoff. The two lines shown in the graph represent quadratic fits. It can be observed that this duration is generally increasing in age. At the age threshold of 40, it can be seen an approximate increase and jump of 2 days in the average non-employment duration, in response to the unemployment benefit extension.
Figure 4 plots the probability of finding a job within 39 weeks of layoff for each age, which is generally decreasing in the age at layoff. We observe a downward jump at the threshold, thus, showing how the benefit extension decreases the probability of finding a job within the first 39 weeks after layoff. Note that while the dashed vertical line denotes the cutoff for unemployment benefit eligibility, extension from 30 to 39 weeks at the threshold, the solid lines represent quadratic fits.
In Figure 5 it is showed the graphical results for the effect on new job wages and plots average natural logarithm wages at these new jobs within bins against the age at layoff. There is an upwards jump at the threshold showing a positive effect on new wages derived by the unemployment benefit extension.

Figure 3: Unemployment Benefit Extension and Nonemployment Duration

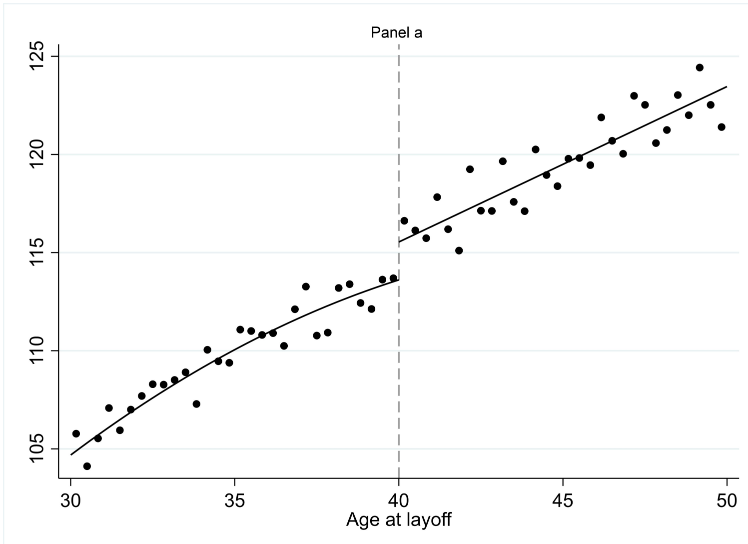


Figure 4: Unemployment Benefit Extension and Probability of Finding a Job within 39 Weeks

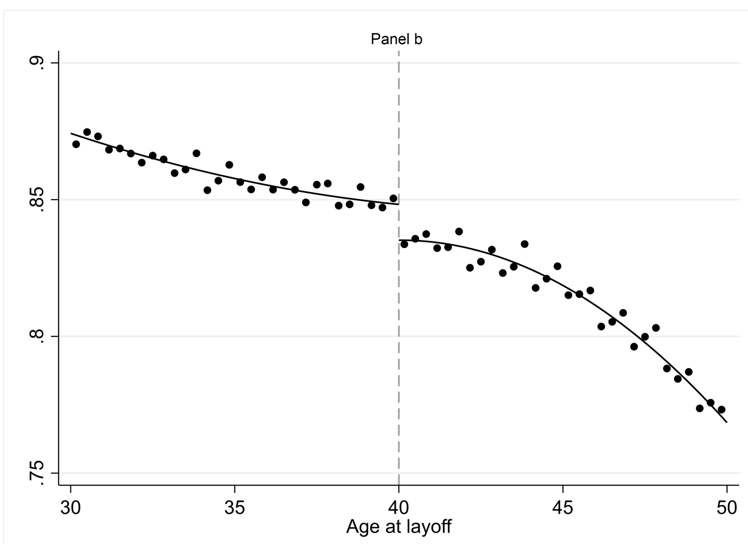
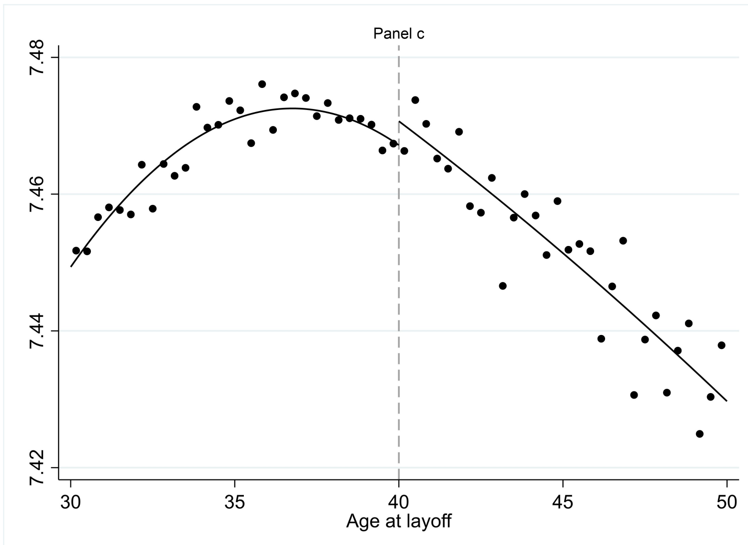


Figure 5: Unemployment Benefit Extension and Log(New Wage)



5. The Table 1 represents the effects of Unemployment Benefit Extension from 30 weeks to 39 weeks.

   Panel A depicts the reduced form effect at the discontinuity.Column 1 shows that average unemployment duration is positively related to age and thus, is an increasing function of age. This suggest that at threshold of 40 there is an increase in unemployment duration by 1.976 units. Column 2 shows that probability of getting new job within 39 weeks is decreasing function of age.This suggest that at threshold the probability of finding a new job reduces by 0.013 units. Column 3 shows the statistically significant positive relationship between new job wages and age. This shows that at threshold the wages for new jobs increases by 0.004 units.

   Panel B depicts the reduced form effects at the discontinuity with a bandwidth ± 5 in year at layoff not the full sample. In this panel the estimates have consistent signs as of Panel A. Column 1 suggest that the average unemployment duration increases by 1.750 units with age at the threshold, which is less than increase in Panel A. Column 2 shows that probability of getting new job within 39 weeks is decreasing function of age. But the decrease in probability is higher than the decrease in Panel A. Column 3 shows that new job wages are increasing function of age at the threshold and increases by 0.007 units which is higher than increase in Panel A.

Panel C depicts the reduced form effect at the discontinuity in the full sample with linear controls for the running variable i.e age, allowing for different slopes above and below the threshold. The sign of estimate of unemployment duration and probability of getting job after 39weeks is consistent with Panel A and Panel B. Column 1 suggest that average unemployment duration increases by 1.352 units with age at threshold ,but this increase is smaller than the increase in Panel A and Panel B. Column 2 suggest the statistically insignificant estimate of probability of getting new job after 39 weeks.Column3 shows that new job wages are decreasing function of age at the threshold. Because of adding linear controls of age this relationship changed its sign in contrast to Panel A and Panel B.

Panel D depicts the reduced form effect at the discontinuity in the full sample with controlling for running variable with a fourth order polynomial, and allowing for different slopes above and below the threshold. The signs of all the estimates are consistent with the Panel A and B.

Panel E depicts the reduced form effects at the discontinuity in the full sample using optimal bandwidth with rdrobust. The estimates of this panel is has consistent signs as of Panel A, B and D. Column 1 shows that the average unemployment duration increases by 1.465 units with the age at threshold, this increase is lower than increase in Panel A, B, D but higher than increase in Panel C. The estimate of probability of getting job after 39 weeks is consistent with above panels. Column 3 also shows the consistent and statistically significant positive relationship between new job wages and age at the threshold.

All the Panel shows the robustness of the estimates in different specifications.

Table 1: Effects of Unemployment Benefit Extension from 30 Weeks to 39 Weeks

	Dependent variable:		
	Non-Employment	New Job 39 Weeks	Post-Wages
	(1)	(2)	(3)
Panel A: Second order polynomals			
Treatment	1.976***	−0.013***	0.004
	(0.526)	(0.002)	(0.002)
Bandwith/Full Sample	Full Sample	Full Sample	Full Sample
Polynomial Degree	2	2	2
Observations	1,589,178	1,738,787	1,189,446
Panel B: Second order polynomals with Bandwith			
Treatment	1.750**	-0.010***	0.007**
	(0.742)	(0.002)	(0.003)
Bandwith/Full Sample	5	5	5
Polynomial Degree	2	2	2
Observations	828,019	901,123	619,039
Panel C: Linear Controls			
Treatment	1.352***	-0.0001	-0.004***
	(0.352)	(0.001)	(0.002)
Bandwith/Full Sample	Full Sample	Full Sample	Full Sample
Polynomial Degree	1	1	1
Observations	1,589,178	1,738,787	1,189,446
Panel D: Fourth order polynomial			
Treatment	1.583*	-0.009***	0.007*
	(0.873)	(0.003)	(0.004)
Bandwith/Full Sample	Full Sample	Full Sample	Full Sample
Polynomial Degree	4	4	4
Observations	1,589,178	1,738,787	1,189,446
Panel E: Optimal Bandwidth with rdrobust			
Treatment	1.465***	-0.009***	0.005*
	(0.708)	(0.002)	(0.003)
Estimated Bandwith	2.833	3.437	3.382
Polynomial Degree	1	1	1
Observations	1,589,178	1,738,787	1,189,446
Note:			*p<0.1; **p<0.05; ***p<0.01

Source Code

```
[language=Stata,breaklines=true]
local ii = 3
 gen agecluster = (int(age*‘ii’)/‘ii’)+1/(2*‘ii’)
 bysort agecluster : gen id = _n
 gen total = _N
 bys agecluster : gen nec = _N
 replace nec = nec / total

*question 2 Density//
twoway ///
(scatter nec   agecluster   if id==1, mcolor(black) msize(small)) ///
(qfit    nec   agecluster   if id==1 & age>40, lcolor(black)) ///
(qfit    nec   agecluster   if id==1 & age<40, lcolor(black)) ///
, xline(40, lp(dash) lcolor(gs10)) ///
xt("Age at layoff") yt("Density") leg(off) graphregion(fcolor(white)) ///
subtitle("Panel a: Distribution of Age",size(small)) name(F1A, replace)

*question 2 logwage
bysort agecluster : egen av_prevwage = mean( lwage0 )
twoway ///
(scatter av_prevwage   agec if   id==1, mcolor(black) msize(small)) ///
(qfit    lwage0    age   if   age>=40, lcolor(black)) ///
(qfit    lwage0    age   if   age<40 , lcolor(black)) ///
, xline(40, lp(dash) lcolor(gs10)) graphregion(fcolor(white)) legend(off)  ///
ytitle("Log previous wage") xt("Age at layoff")

*question 4 non-employment duration
bysort agecluster : egen av_nonemp = mean(nonemp)
twoway ///
(scatter av_nonemp    agecluster if   id==1, mcolor(black) msize(small)) ///
(qfit    nonemp    age   if   age>=40, lcolor(black) ) ///
(qfit    nonemp    age   if   age<40 , lcolor(black) ) ///
, xline(40, lp(dash) lcolor(gs10)) graphregion(fcolor(white)) legend(off)  ytitle("‘tit’") xtitle("Age at layoff") subtitle("Panel a ",size(small) ) name(F3a, replace)



*question4 probability of finding a job
bysort agecluster : egen avjobfind = mean(jobfind)
```

```
twoway ///
(scatter avjobfind   agecluster if   id==1, mcolor(black) msize(small)) ///
(qfit    jobfind    age   if   age>=40, lcolor(black) ) ///
(qfit    jobfind   age   if   age<40 , lcolor(black) ) ///
, xline(40, lp(dash) lcolor(gs10)) graphregion(fcolor(white)) legend(off) ///
  ytitle("‘Probability of finding job’")   xtitle("Age at layoff") subtitle("Panel b ",size(small))

*question4 logwage
bysort agecluster : egen avernewwagenewwage = mean( lwage1 )
twoway ///
(scatter avernewwagenewwage   agecluster if   id==1, mcolor(black) msize(small)) ///
(qfit    lwage1    age   if   age>=40, lcolor(black) ) ///
(qfit    lwage1    age   if   age<40 , lcolor(black) ) ///
, xline(40, lp(dash) lcolor(gs10)) graphregion(fcolor(white)) legend(off)  ytitle("‘tit’") xtitle("Age at layoff") subtitle("Panel c",size(small) ) name(F3c, replace)
```

```
[language=R,breaklines=true]
library(tidyverse)
library(haven)
library(dplyr)
library(stargazer)
library(rdrobust)

dataemp <- read_dta(file = "C:\\Users\\btpta\\Desktop\\Metrics 2\\PS 4\\dataset_ps4.dta")
attach(dataemp)

#generating bins
i <- 3
dataemp <- dataemp  %>% add_column(agegroup = (ceiling(age*i)/i)+1/(2*i)) %>%
  group_by(agegroup) %>%
  #mutate(id = cur_group_id())%>%
  mutate(gn = n()) %>%
  mutate(dens = gn/nrow(dataemp))

#Figure 1 Density of age
rdplot( dataemp$dens, dataemp$age, c=40, p=2, masspoints="off", title="Density of Age at Layoff", x.label = "Age", y.label ="Density")

#Figure 2 Previous wage at the threshold
dataemp <- dataemp %>%
  group_by(agegroup) %>%
```

```
  mutate(avwage0 = mean(as.numeric(lwage0),na.rm=T))
rdplot( dataemp$avwage0, dataemp$agegroup, c=40, p=2,
    masspoints="off", title="Previous wage at the threshold",
         x.label = "Age", y.label ="Log Previous Wage")

#Figure 3 Nonemployment Duration
dataemp <- dataemp %>%
  group_by(agegroup) %>%
  mutate(avnonem = mean(as.numeric(nonemp),na.rm=T))
rdplot( dataemp$avnonem, dataemp$agegroup, c=40, p=2,
    masspoints="off", title="Unemployment Benefit Extension
    and Nonemployment Duration",
         x.label = "Age", y.label ="Non-Emplyment Duration")

#Figure 4 Probability of Finding a Job after 39 Weeks

dataemp <- dataemp %>%
  group_by(agegroup) %>%
  mutate(avfindjob = mean(as.numeric(jobfind),na.rm=T))
rdplot(dataemp$avfindjob, dataemp$agegroup, c=40, p=2,
    masspoints="off", title="Finding a Job within 39 Weeks",
         x.label = "Age", y.label ="Probability of finding
             job")

#Figure 5 New wage around threshhold
dataemp <- dataemp %>%
  group_by(agegroup) %>%
  mutate(avwage1 = mean(as.numeric(lwage1),na.rm=T))
rdplot( dataemp$avwage1, dataemp$agegroup, c=40, p=2,
    masspoints="off", title="New wage at the threshold",
         x.label = "Age", y.label ="Log Previous Wage")


rdplot( dataemp$avwage1, dataemp$agegroup, c=40, p=2,
    masspoints="off",  title="New wage at the threshold",
         x.label = "Age", y.label ="Log Previous Wage")

##########Ex 5
cutoff=40
dataemp$age40=0
dataemp$age40[age>cutoff]=1
attach(dataemp)
#centered regressors
dataemp$age_p1=(age-cutoff)^1
dataemp$age_p2=(age-cutoff)^2
dataemp$age_p3=(age-cutoff)^3
dataemp$age_p4=(age-cutoff)^4

dataemp$age_q1=(age-cutoff)^1*age40
dataemp$age_q2=(age-cutoff)^2*age40
```

```
dataemp$age_q3=(age-cutoff)^3*age40
dataemp$age_q4=(age-cutoff)^4*age40

#Table 1
#Panel 1 - Second order polynomal
model1<-lm(nonemp ~ age40+age_p1+age_p2+age_q1+age_q2, data=dataemp)  #noneplyment duration
model2<-lm(jobfind~age40+age_p1+age_p2+age_q1+age_q2, data=dataemp) #Prob of finding a job within 39
model3<-lm(lwage1~ age40+age_p1+age_p2+age_q1+age_q2, data=dataemp) #re-employment wage

star.out.1 <- stargazer(model1, model2, model3,  align=F, dep.var.labels=c("Non-Employment", "New Job 39 Weeks", "Post-Wages"),
                        covariate.labels= c("Treatment"), omit=c("age_p1", "age_p2", "age_q1", "age_q2", "Constant"), keep.stat="n",
                        add.lines=list(c("Bandwith/Full Sample", "Full Sample", "Full Sample", "Full Sample"), c("Polynomial Degree", "2",  "2", "2")), column.sep.width = "1pt")
#Panel 2 - Second order polynomal with bandwith +/- 5 years
bandwith=5
datalim<-subset(dataemp, abs(age_p1)<=bandwith)

model4<-lm(nonemp ~ age40+age_p1+age_p2+age_q1+age_q2, data=datalim)  #noneplyment duration
model5<-lm(jobfind~age40+age_p1+age_p2+age_q1+age_q2, data=datalim) #Prob of finding a job within 39
model6<-lm(lwage1~ age40+age_p1+age_p2+age_q1+age_q2, data=datalim) #re-employment wage

star.out.2 <- stargazer(model4, model5, model6,  align=T, dep.var.labels=c("Non-Employment Duration", "New Job Within 39 Weeks", "Re-Employment Wages"),
                        covariate.labels= c("Treatment"), omit=c("age_p1", "age_p2", "age_q1", "age_q2","Constant"), keep.stat="n",
                        add.lines=list(c("Bandwith/Full Sample", "5", "5", "5"), c("Polynomial Degree", "2",  "2", "2"))
)

#Panel 3 - Linear Controls
model7<-lm(nonemp ~ age40+age_p1+age_q1, data=dataemp)  #
```

```r
noneplyment duration
model8<-lm(jobfind~age40+age_p1+age_q1, data=dataemp) #Prob of finding a job within 39
model9<-lm(lwage1~ age40+age_p1+age_q1, data=dataemp) #re-employment wage

star.out.3 <- stargazer(model7, model8, model9,  align=T, dep.var.labels=c("Non-Employment Duration", "New Job Within 39 Weeks", "Re-Employment Wages"),
                        covariate.labels= c("Treatment"), omit=c("age_p1", "age_q1", "Constant"), keep.stat="n",
                        add.lines=list(c("Bandwith/Full Sample", "Full Sample", "Full Sample", "Full Sample"), c("Polynomial Degree", "1",  "1", "1")))

#Panel 4 - Fourth order polynomial

model10<-lm(nonemp ~ age40+age_p1+age_p2+age_p3+age_p4+age_q1+age_q2+age_q3+age_q4, data=dataemp)  #noneplyment duration
model11<-lm(jobfind ~ age40+age_p1+age_p2+age_p3+age_p4+age_q1+age_q2+age_q3+age_q4, data=dataemp) #Prob of finding a job within 39
model12<-lm(lwage1 ~ age40+age_p1+age_p2+age_p3+age_p4+age_q1+age_q2+age_q3+age_q4, data=dataemp) #re-employment wage


star.out.4<- stargazer(model10, model11, model12,  align=T, dep.var.labels=c("Non-Employment Duration", "New Job Within 39 Weeks", "Re-Employment Wages"),
                       covariate.labels= c("Treatment"), omit=c("age_p1", "age_p2","age_p3", "age_p4", "age_q1", "age_q2","age_q3", "age_q4", "Constant"), keep.stat="n",
                       add.lines=list(c("Bandwith/Full Sample", "Full Sample", "Full Sample", "Full Sample"), c("Polynomial Degree", "4",  "4", "4")))

#Panel 5 - Nonparametric estimation

model13<- rdrobust(dataemp$nonemp, dataemp$age, c=40, all=T, masspoints="off")   #noneplyment duration
model14<- rdrobust(dataemp$jobfind, dataemp$age, c=40, all=T
```

```
, masspoints="off")  #Prob of finding a job within 39
model15<- rdrobust(dataemp$lwage1, dataemp$age, c=40, all=T, masspoints="off")  #re-employment wage

star.out.5<- stargazer(model13, model14, model15,  align=T, dep.var.labels=c("Non-Employment Duration",
             "New Job Within 39 Weeks",  "Re-Employment Wages"),
        covariate.labels= c("Treatment"), keep.stat="n",
        add.lines=list(c("Estimated Bandwith", "2.833", "3.437", "3.382"),  c("Polynomial Degree", "1",  "1", "1")))


Tablestar <- star_panel (star.out.1, star.out.2, star.out.3, star.out.4, same.summary.stats = FALSE,
                          panel.names = c("Second order polynomals", "Second order polynomals with Bandwith",
                                          "Linear Controls", "Fourth order polynomial", "Optimal Bandwidth with rdrobust" ))
star_tex_write(Tablestar, file = "my_tex_file2.tex", headers = TRUE)
```

References

[Angrist and Pischke, 2009] Angrist, J. D. and Pischke, J. (2009). Mostly harmless econometrics: An empiricist's companion.

[Caliendo, 2012] Caliendo, M., T. K. U. A. (2012). Benefit duration, unemployment duration and job match quality: A regression-discontinuity approach. *Journal of Applied Econometrics*, 28(4):604–627.

[Ehrenberg and Oaxaca, 1976] Ehrenberg, R. and Oaxaca, R. (1976). Unemployment insurance, duration of unemployment, and subsequent wage gain. *American Economic Review*, 66(5):754–766.

[Ham, 1996] Ham, J.C., L. R. (1996). The effect of sample selection and initial conditions in duration models: Evidence from experimental data on training. *Econometrica*, 64(1):175–205.

[Nekoei and Weber, 2017] Nekoei, A. and Weber, A. (2017). Does extending unemployment benefits improve job quality? *American economic review*, 107(2):527–561.

[Schmieder and Bender, 2016] Schmieder, J., v. W. T. and Bender, S. (2016). The effect of unemployment benefits and nonemployment durations on wages. *American Economic Review*, 106(3):739–777.