

# optrdd2

*haiyan yu*

*1/18/2020*

```
library("rdd")
```

```
## Loading required package: sandwich
```

```
## Loading required package: lmtest
```

```
## Loading required package: zoo
```

```
##
```

```
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      as.Date, as.Date.numeric
```

```
## Loading required package: AER
```

```
## Loading required package: car
```

```
## Loading required package: carData
```

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

```
library("rddtools")
```

```
## Loading required package: np
```

```
## Nonparametric Kernel Methods for Mixed Datatypes (version 0.60-9)
```

```
## [vignette("np_faq",package="np") provides answers to frequently asked questions]
```

```
## [vignette("np",package="np") an overview]
```

```
## [vignette("entropy_np",package="np") an overview of entropy-based methods]
```

```
library(readxl)
```

```
#award <- read_excel("~/Downloads/Rpkg/optrdd2/data/sum_avg_award_paper.xlsx",sheet = "df")
```

```
award <- read_excel("~/Downloads/Rpkg/optrdd2/data/data_avg_treat.xlsx")
```

```
#head(award)
```

```
#award<-award[1:4152,]
```

```
#award<-data_avg_treat
```

```
head(award[award$D==0,])
```

```
## # A tibble: 6 x 9
##   period   DRR Thank Gift Contri Article Patient Views   D
##   <dbl>   <dbl> <dbl> <dbl> <dbl>   <dbl>   <dbl>   <dbl> <dbl>
## 1     1 -0.007   4.2  14    1515   -2.02     87 143432     0
## 2     2 -0.005   4.1 10.9   1471    0.61     79 141268     0
## 3     3 -0.004   3.3 14.8   1331    0.52     65 107423     0
## 4     4 -0.001   2.8 10.3   1436    0.45     74 101340     0
## 5     5 -0.009    3    9.2   1475    0.55     72 151260     0
## 6     6 -0.003    3    9.6   1475    0.47     70 153008     0
```

```
award1<-award[award$D==0,]
award2<-award[award$D==1,]
##
cut.off<-6
#award<-sum_avg_award_paper
rdd_dat_indh <- rdd_data(#y=Patient,
                        y=Views,
                        x=period,
                        data=award1,
                        cutpoint=cut.off)
#award<-sum_avg_award_paper
rdd_dat_indh2 <- rdd_data(#y=Patient,
                         y=Views,
                         x=period,
                         data=award2,
                         cutpoint=cut.off)
str(rdd_dat_indh)
```

```
## Classes 'rdd_data' and 'data.frame': 11 obs. of 2 variables:
## $ x: num 1 2 3 4 5 6 7 8 9 10 ...
## $ y: num 143432 141268 107423 101340 151260 ...
## - attr(*, "hasCovar")= logi FALSE
## - attr(*, "labels")= list()
## - attr(*, "cutpoint")= num 6
## - attr(*, "type")= chr "Sharp"
```

```
str(rdd_dat_indh2)
```

```
## Classes 'rdd_data' and 'data.frame': 11 obs. of 2 variables:
## $ x: num 1 2 3 4 5 6 7 8 9 10 ...
## $ y: num 173977 184098 145770 128537 180300 ...
## - attr(*, "hasCovar")= logi FALSE
## - attr(*, "labels")= list()
## - attr(*, "cutpoint")= num 6
## - attr(*, "type")= chr "Sharp"
```

```
#plot(y=award$patient1, x=award$article, data=award)
#plot(y=award$patient1, x=award$period, data=award)
summary(rdd_dat_indh)
```

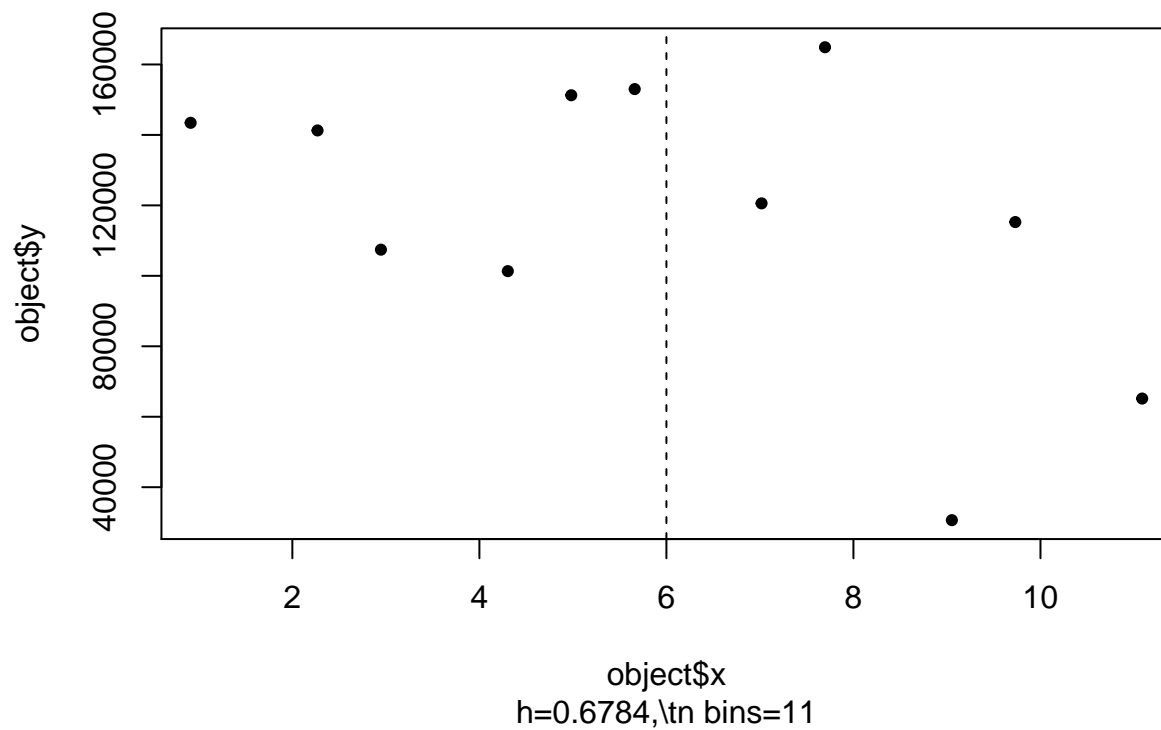
```
## ### rdd_data object ###
##
```

```
## Cutpoint: 6
## Sample size:
## -Full : 11
## -Left : 5
## -Right: 6
## Covariates: no
```

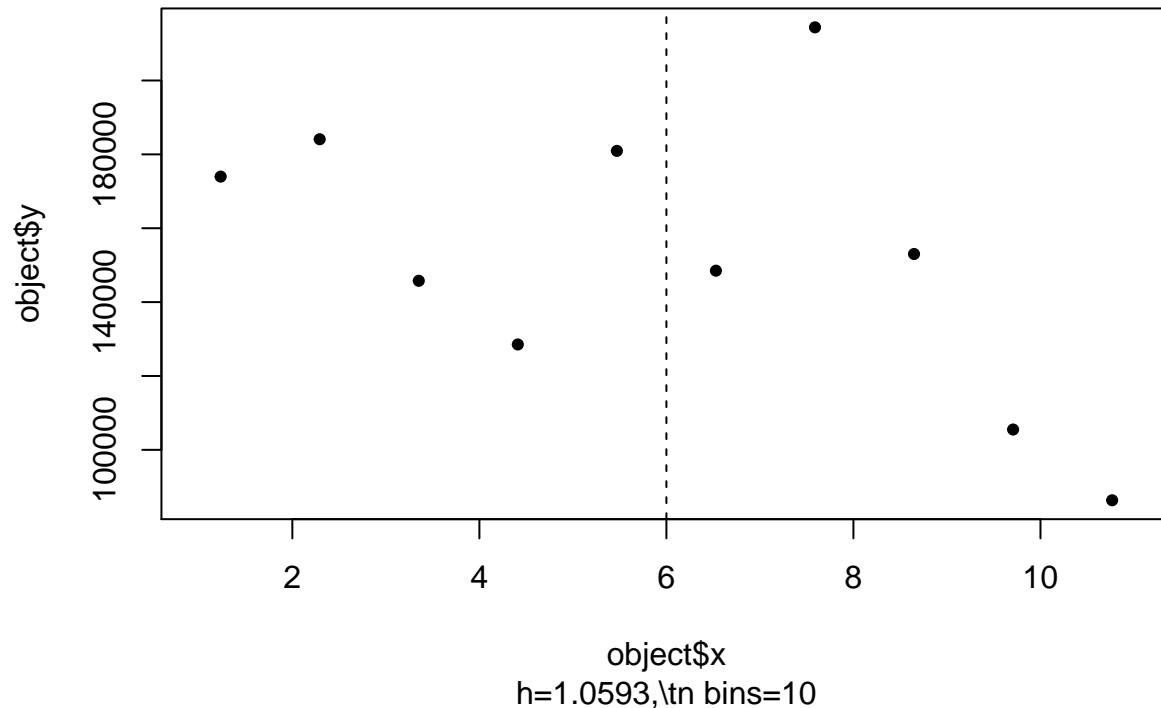
```
head(rdd_dat_indh)
```

```
##   x      y
## 1 1 143432
## 2 2 141268
## 3 3 107423
## 4 4 101340
## 5 5 151260
## 6 6 153008
```

```
#df<-data.farme(award)
#award_rdd <- rdd_data(y=award$patient1, x=award$article, cutpoint=700)
plot(rdd_dat_indh)
```



```
plot(rdd_dat_indh2)
```



```
(reg_para <- rdd_reg_lm(rdd_dat_indh, order=1))
```

```
## ### RDD regression: parametric ###
## Polynomial order: 1
## Slopes: separate
## Number of obs: 11 (left: 5, right: 6)
##
## Coefficient:
## Estimate Std. Error t value Pr(>|t|)
## D      28694      49008  0.5855  0.5766
```

```
(reg_para2 <- rdd_reg_lm(rdd_dat_indh2, order=1))
```

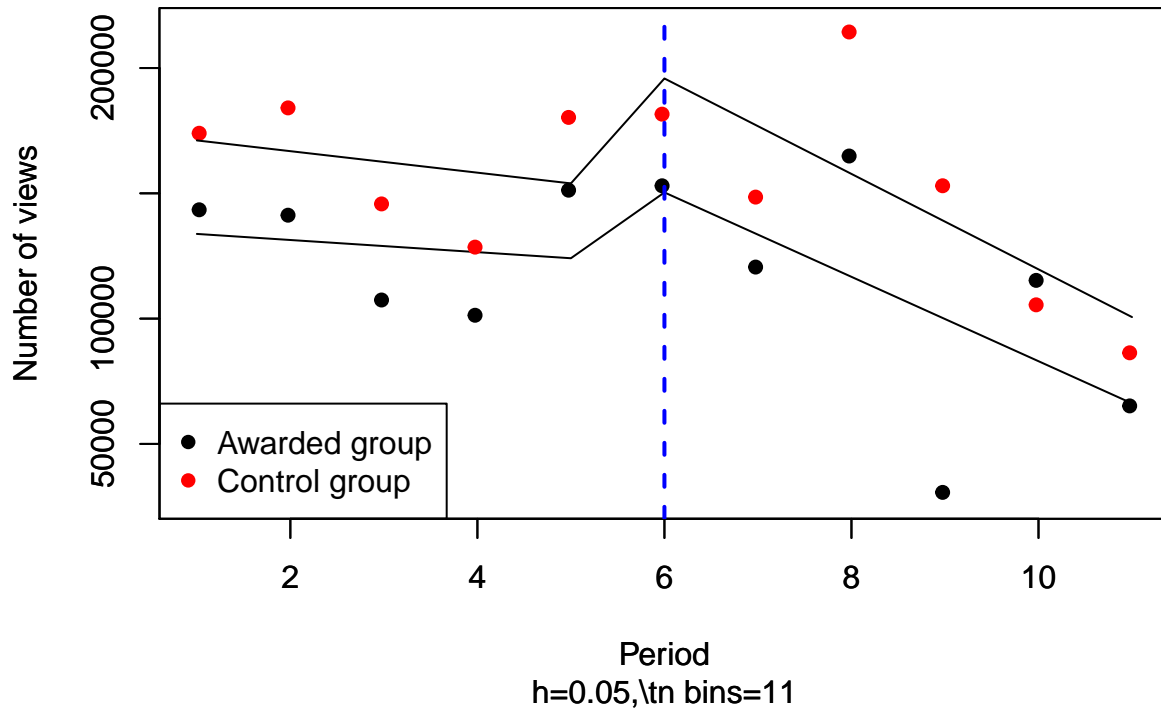
```
## ### RDD regression: parametric ###
## Polynomial order: 1
## Slopes: separate
## Number of obs: 11 (left: 5, right: 6)
##
## Coefficient:
## Estimate Std. Error t value Pr(>|t|)
## D      46180      40251  1.1473  0.289
```

```
rmin<-min(award$Views)-3000
rmax<-max(award$Views)+2000
plot(reg_para,xlab=c("Period"),ylab=c("Number of views"),ylim=range(rmin,rmax))
#use par and plot on the same graph but different axis.
```

```

par(new=TRUE)
plot(reg_para2,xlab=c("Period"),ylab=c("Number of views"),ylim=range(rmin,rmax), col = "red")
legend("bottomleft",legend=c("Awarded group","Control group"), pch=c(16,16),col = c("black","red"),horiz=TRUE)
abline(v = cut.off, col="blue", lwd=2, lty=2)

```



```

#(reg_para <- rdd_reg_lm(rdd_dat_indh, order=0))
#plot(reg_para,xlab=c("Period"),ylab=c("Number of patients"))

#(reg_para <- rdd_reg_lm(rdd_dat_indh, order=4))
#plot(reg_para,xlab=c("Period"),ylab=c("Number of patients"))

```

```

bw_ik <- rdd_bw_ik(rdd_dat_indh)
bw_ik2 <- rdd_bw_ik(rdd_dat_indh2)
(reg_nonpara <- rdd_reg_np(rdd_object=rdd_dat_indh, bw=bw_ik))

```

```

## ### RDD regression: nonparametric local linear###
## Bandwidth: 5.950942
## Number of obs: 11 (left: 5, right: 6)
##
## Coefficient:
## Estimate Std. Error z value Pr(>|z|)
## D 16979 52675 0.3223 0.7472

```

```

(reg_nonpara2 <- rdd_reg_np(rdd_object=rdd_dat_indh2, bw=bw_ik2))

```

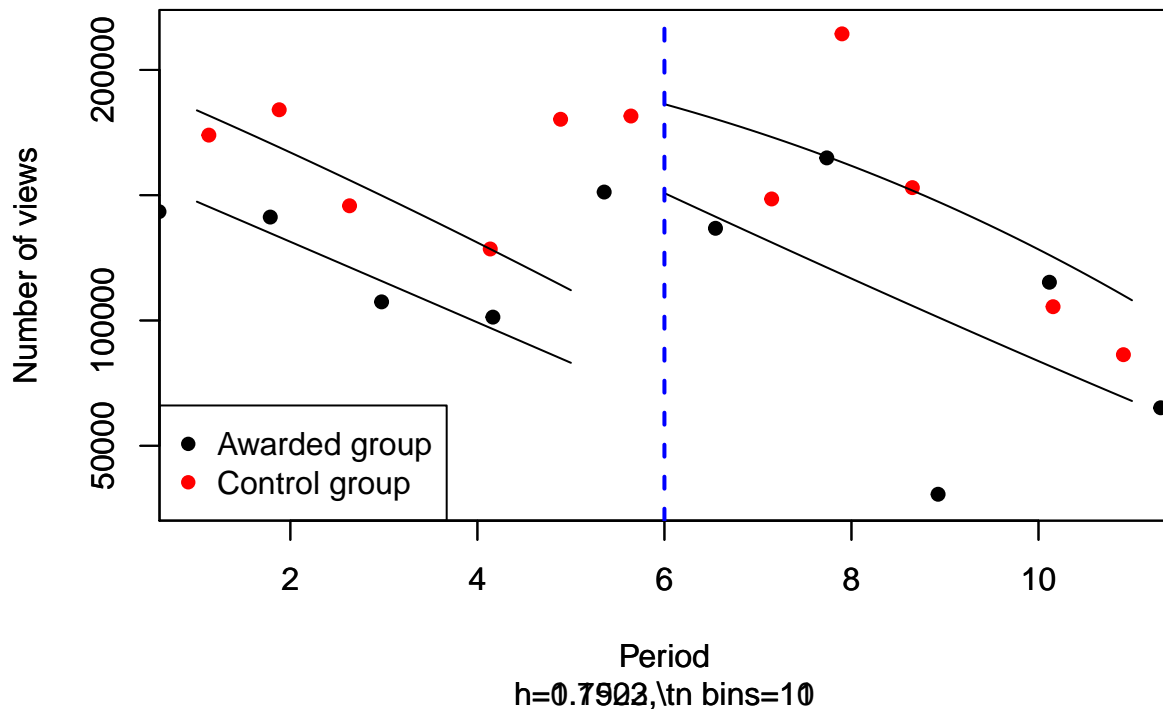
```

## ### RDD regression: nonparametric local linear###
## Bandwidth: 3.761362
## Number of obs: 7 (left: 3, right: 4)

```

```
##
## Coefficient:
## Estimate Std. Error z value Pr(>|z|)
## D -25330 47149 -0.5372 0.5911

rmin<-min(award$Views)-3000
rmax<-max(award$Views)+2000
plot(reg_nonpara,xlab=c("Period"),ylab=c("Number of views"),ylim=range(rmin,rmax))
#use par and plot on the same graph but different axis.
par(new=TRUE)
plot(reg_nonpara2,xlab=c("Period"),ylab=c("Number of views"),ylim=range(rmin,rmax), col = "red")
legend("bottomleft",legend=c("Awarded group","Control group"), pch=c(16,16),col = c("black","red"),horiz=TRUE)
abline(v = cut.off, col="blue", lwd=2, lty=2)
```



```
##
cut.off<-6
#award<-sum_avg_award_paper
rdd_dat_indh3 <- rdd_data(y=Patient,
                        x=period,
                        data=award1,
                        cutpoint=cut.off)
#award<-sum_avg_award_paper
rdd_dat_indh4 <- rdd_data(y=Patient,
                        x=period,
                        data=award2,
                        cutpoint=cut.off)
str(rdd_dat_indh3)
```

```
## Classes 'rdd_data' and 'data.frame': 11 obs. of 2 variables:
## $ x: num 1 2 3 4 5 6 7 8 9 10 ...
```

```
## $ y: num 87 79 65 74 72 70 54 83 43 77 ...
## - attr(*, "hasCovar")= logi FALSE
## - attr(*, "labels")= list()
## - attr(*, "cutpoint")= num 6
## - attr(*, "type")= chr "Sharp"
```

```
str(rdd_dat_indh4)
```

```
## Classes 'rdd_data' and 'data.frame': 11 obs. of 2 variables:
## $ x: num 1 2 3 4 5 6 7 8 9 10 ...
## $ y: num 186 173 143 149 150 151 123 182 164 157 ...
## - attr(*, "hasCovar")= logi FALSE
## - attr(*, "labels")= list()
## - attr(*, "cutpoint")= num 6
## - attr(*, "type")= chr "Sharp"
```

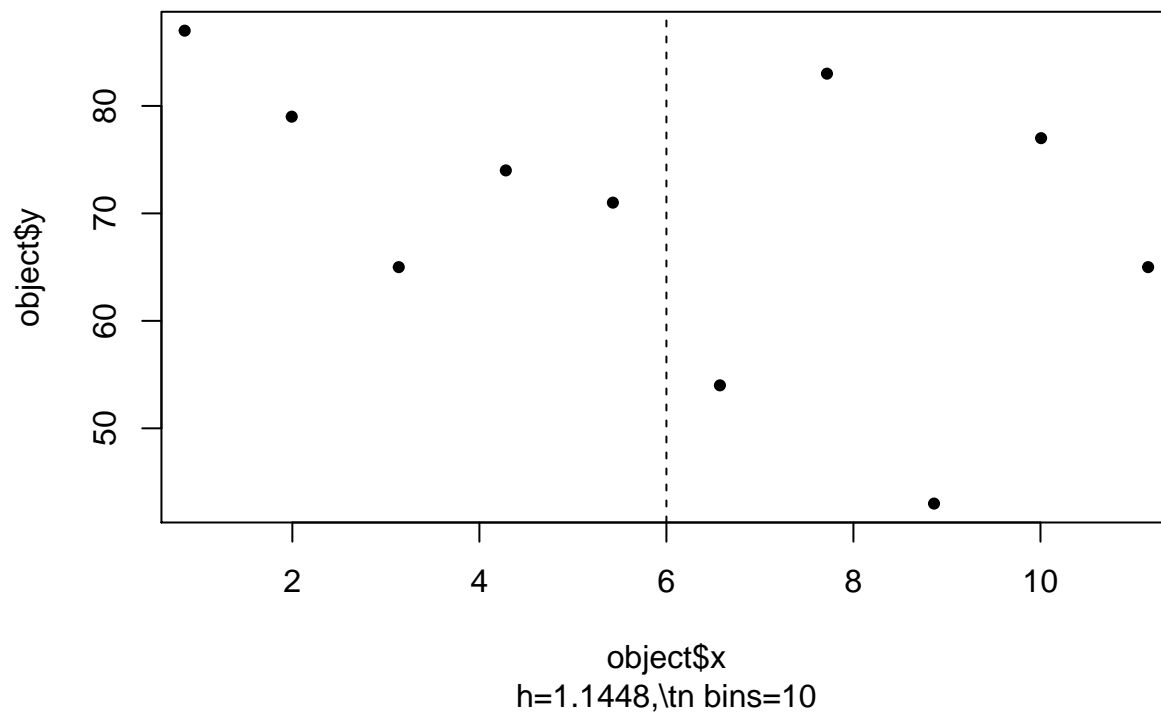
```
#plot(y=award$patient1, x=award$article, data=award)
#plot(y=award$patient1, x=award$period, data=award)
summary(rdd_dat_indh3)
```

```
## ### rdd_data object ###
##
## Cutpoint: 6
## Sample size:
## -Full : 11
## -Left : 5
## -Right: 6
## Covariates: no
```

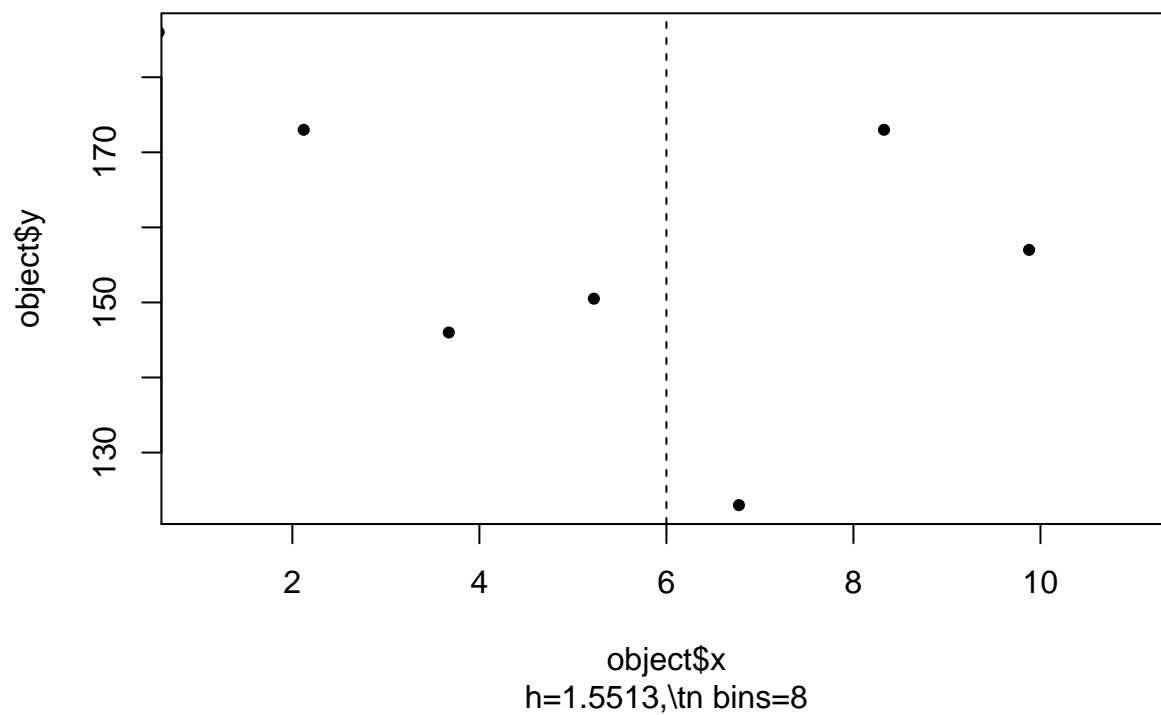
```
head(rdd_dat_indh3)
```

```
## x y
## 1 1 87
## 2 2 79
## 3 3 65
## 4 4 74
## 5 5 72
## 6 6 70
```

```
#df<-data.farme(award)
#award_rdd <- rdd_data(y=award$patient1, x=award$article, cutpoint=700)
plot(rdd_dat_indh3)
```



```
plot(rdd_dat_indh4)
```



```
(reg_para3 <- rdd_reg_lm(rdd_dat_indh3, order=1))
```

```
## ### RDD regression: parametric ###
## Polynomial order: 1
```

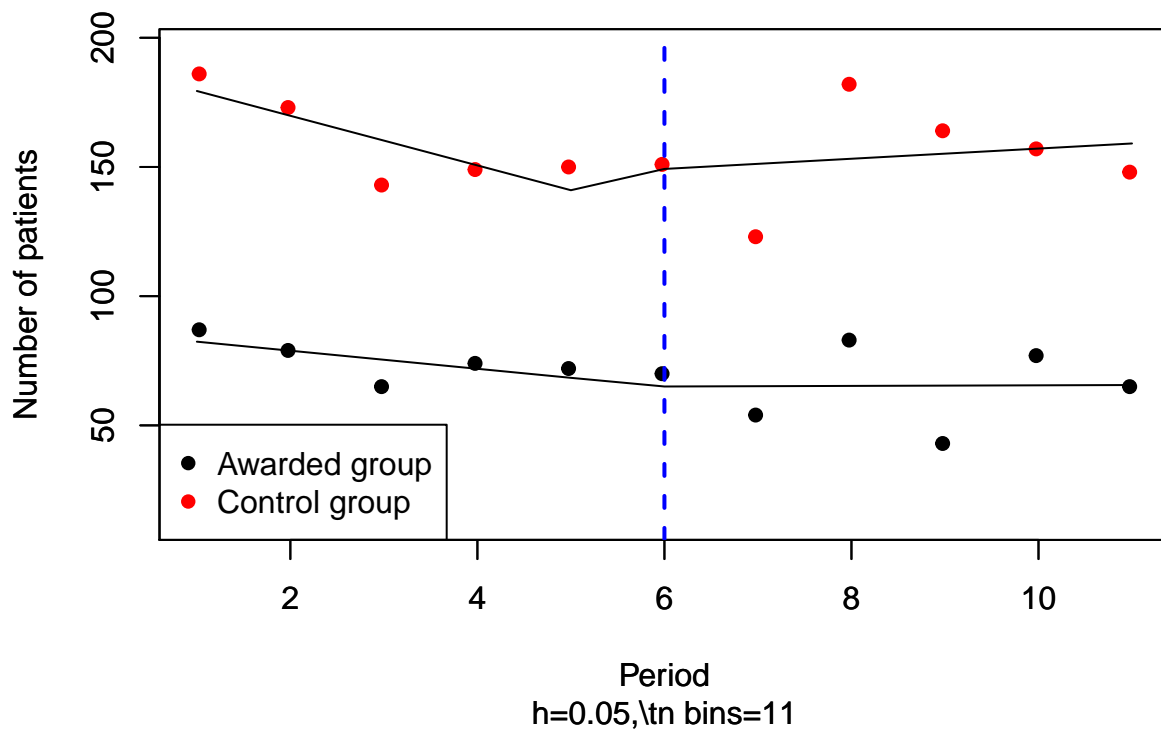


```
## Slopes: separate
## Number of obs: 11 (left: 5, right: 6)
##
## Coefficient:
## Estimate Std. Error t value Pr(>|t|)
## D 0.14762 16.98614 0.0087 0.9933
```

```
(reg_para4 <- rdd_reg_lm(rdd_dat_indh4, order=1))
```

```
## ### RDD regression: parametric ###
## Polynomial order: 1
## Slopes: separate
## Number of obs: 11 (left: 5, right: 6)
##
## Coefficient:
## Estimate Std. Error t value Pr(>|t|)
## D 17.838 22.917 0.7784 0.4618
```

```
rmin<-min(award$Patient)-30
rmax<-max(award$Patient)+10
plot(reg_para3,xlab=c("Period"),ylab=c("Number of patients"),ylim=range(rmin,rmax), col = "black")
par(new=TRUE)
plot(reg_para4,xlab=c("Period"),ylab=c("Number of patients"),ylim=range(rmin,rmax), col = "red")
legend("bottomleft",legend=c("Awarded group","Control group"), pch=c(16,16),col = c("black","red"),horiz=TRUE)
abline(v = cut.off, col="blue", lwd=2, lty=2)
```



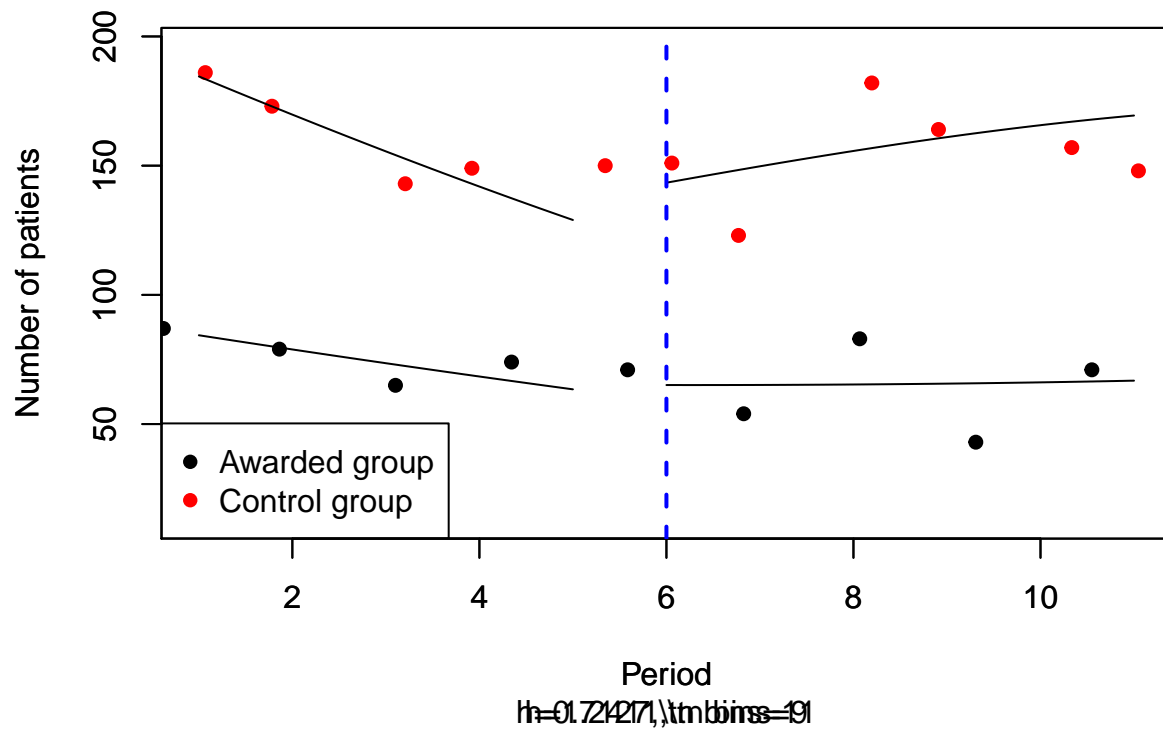
```
bw_ik3 <- rdd_bw_ik(rdd_dat_indh3)
bw_ik4 <- rdd_bw_ik(rdd_dat_indh4)
(reg_nonpara3 <- rdd_reg_np(rdd_object=rdd_dat_indh3, bw=bw_ik3))
```

```
## ### RDD regression: nonparametric local linear###
## Bandwidth: 6.205503
## Number of obs: 11 (left: 5, right: 6)
##
## Coefficient:
## Estimate Std. Error z value Pr(>|z|)
## D -2.1794 15.8199 -0.1378 0.8904
```

```
(reg_nonpara4 <- rdd_reg_np(rdd_object=rdd_dat_indh4, bw=bw_ik4))
```

```
## ### RDD regression: nonparametric local linear###
## Bandwidth: 3.563635
## Number of obs: 7 (left: 3, right: 4)
##
## Coefficient:
## Estimate Std. Error z value Pr(>|z|)
## D -11.653 29.417 -0.3961 0.692
```

```
rmin<-min(award$Patient)-30
rmax<-max(award$Patient)+10
plot(reg_nonpara3,xlab=c("Period"),ylab=c("Number of patients"),ylim=range(rmin,rmax))
par(new=TRUE)
plot(reg_nonpara4,xlab=c("Period"),ylab=c("Number of patients"),ylim=range(rmin,rmax), col = "red")
legend("bottomleft",legend=c("Awarded group","Control group"), pch=c(16,16),col = c("black","red"),horiz=TRUE)
abline(v = cut.off, col="blue", lwd=2, lty=2)
```



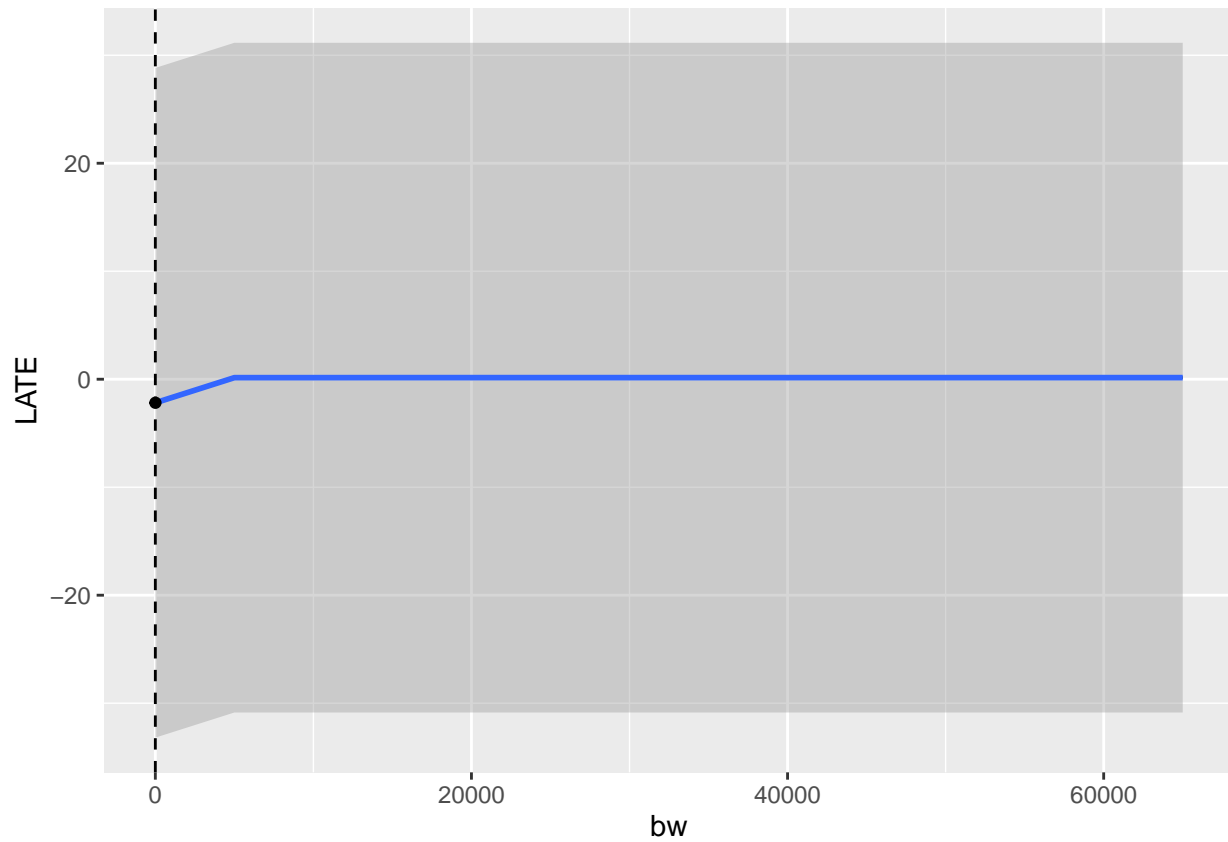
*#Before awards*

Sensitivity tests.

```
plotSensi(reg_nonpara3, from=1, to=70000, by=5000)
```

```
## Warning: Removed 1 rows containing missing values (geom_path).
```

```
## Warning: Removed 1 rows containing missing values (geom_smooth).
```



```
#plotPlacebo(reg_nonpara)  
#dens_test(reg_nonpara)
```