CA4aVdVeen

Ties van der Veen

21-9-2019

Rebecca Costello - Ties van der Veen - Niels van Opstal

I. Introduction to the assignment

Packages

```
options(repos="https://cran.rstudio.com")
install.packages("rdrobust")
```

```
## package 'rdrobust' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\tiess\AppData\Local\Temp\RtmpwRwxQO\downloaded_packages
```

```
install.packages("lubridate")
```

```
## package 'lubridate' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\tiess\AppData\Local\Temp\RtmpwRwxQO\downloaded_packages
```

```
library(foreign)
library(tidyverse)
library(ggdag)
library(dplyr)
library(tinytex)
library(jtools)
library(huxtable)
library(summarytools)
library(ggstance)
library(pwr)
library(knitr)
library(lemon)
library(AER)
library(rdrobust)
library(lubridate)
```

```
knit_print.data.frame <- lemon_print
```

```
st_options(plain.ascii = FALSE, style = "rmarkdown")
st_css()
```

```
## <style type="text/css">
##  img {   background-color: transparent;   border: 0; }  .st-table td, .st-table th {   padding: 8px;
```

Data

```
theUrl_ca4a <- "https://surfdrive.surf.nl/files/index.php/s/QR62KtWGsAlKiE8/download"
parents <- read.dta (file = theUrl_ca4a)
```

II. Potential outcomes

(a) Y(0,i) = more likely to return to work and not have a 2nd child Y(1,i) = less likely to return to work and more likely to have a 2nd child

(b) The new policy allowed women to have longer parental leave if their baby was born after July 1st, 1990. This also allowed them to chain together parental leave for the 2nd (and more) children if this child was born within a certain time period. So some mothers chose to get a second child faster, while they had this parental leave they could use.

III. Descriptive statistics

(a)

```
parents$bd <- as.Date(parents$bd)
min(parents$bd, na.rm=TRUE)
```

```
## [1] "1990-06-01"
```

```
max(parents$bd, na.rm=TRUE)
```

```
## [1] "1990-07-30"
```

(b)

```
NROW(parents$bd)
```

```
## [1] 6180
```

(c)

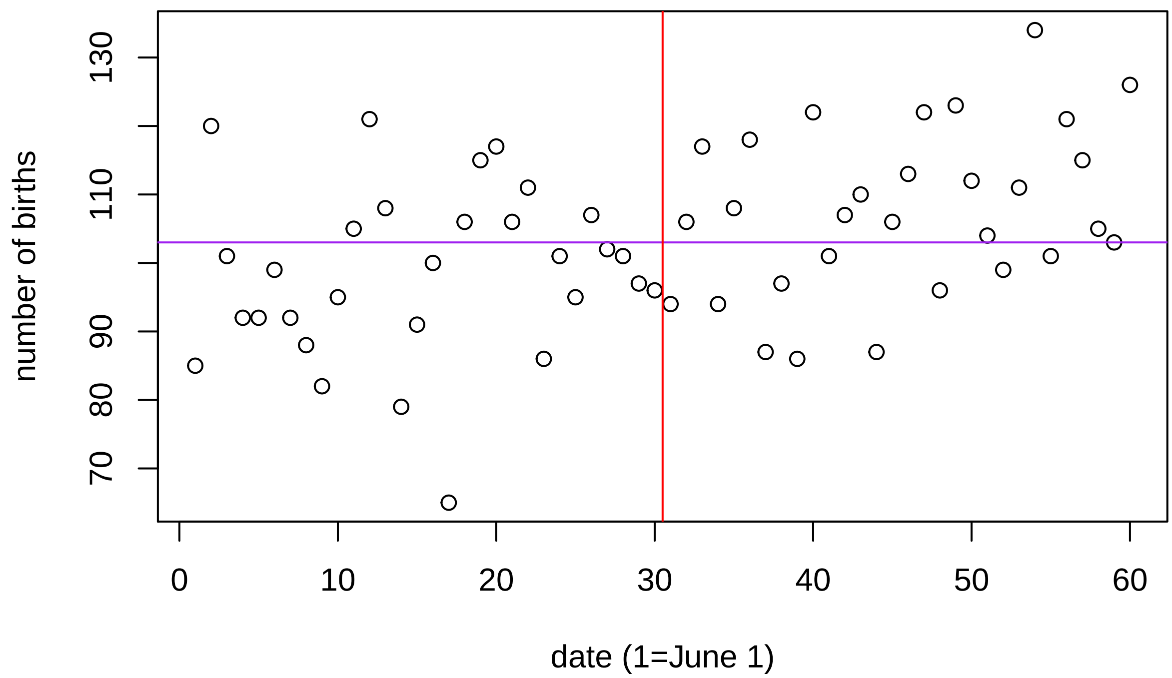
```
obsperday <- parents %>% group_by(month(parents$bd), mday(parents$bd)) %>% tally()
mean(obsperday$n)
```

```
## [1] 103
```

(d)

```
obsperday$daycounter <- seq.int(nrow(obsperday))
```

```r
plot(obsperday$daycounter, obsperday$n, ylab = "number of births", xlab = "date (1=June 1)", xlim = (c(
abline(h=mean(obsperday$n), col="purple")
abline(v=30.5, col="red")
```



(e)

```r
summary(parents$uncb3)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0000  0.0000  0.0000  0.3448  1.0000  1.0000
```

So around 34% had an additional birth within three years.

(f)

```r
summary(parents$uncj3)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0000  0.0000  1.0000  0.5639  1.0000  1.0000
```

So around 56% returned to work within three years.

(g)

```r
summary(parents$age3034)
```

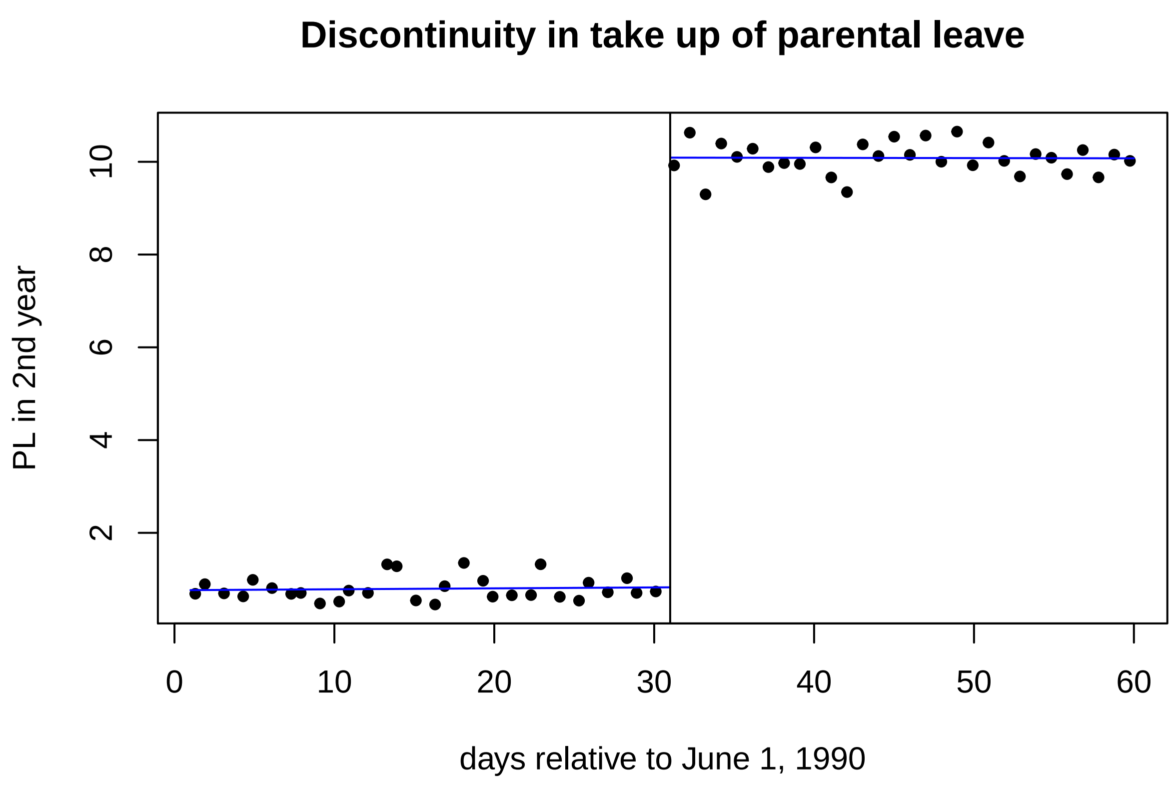
```
##     Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.00000 0.00000 0.00000 0.09822 0.00000 1.00000
```

So around 10% is aged between 30 and 34.

Iv. Graphical evidence of a treatment effect

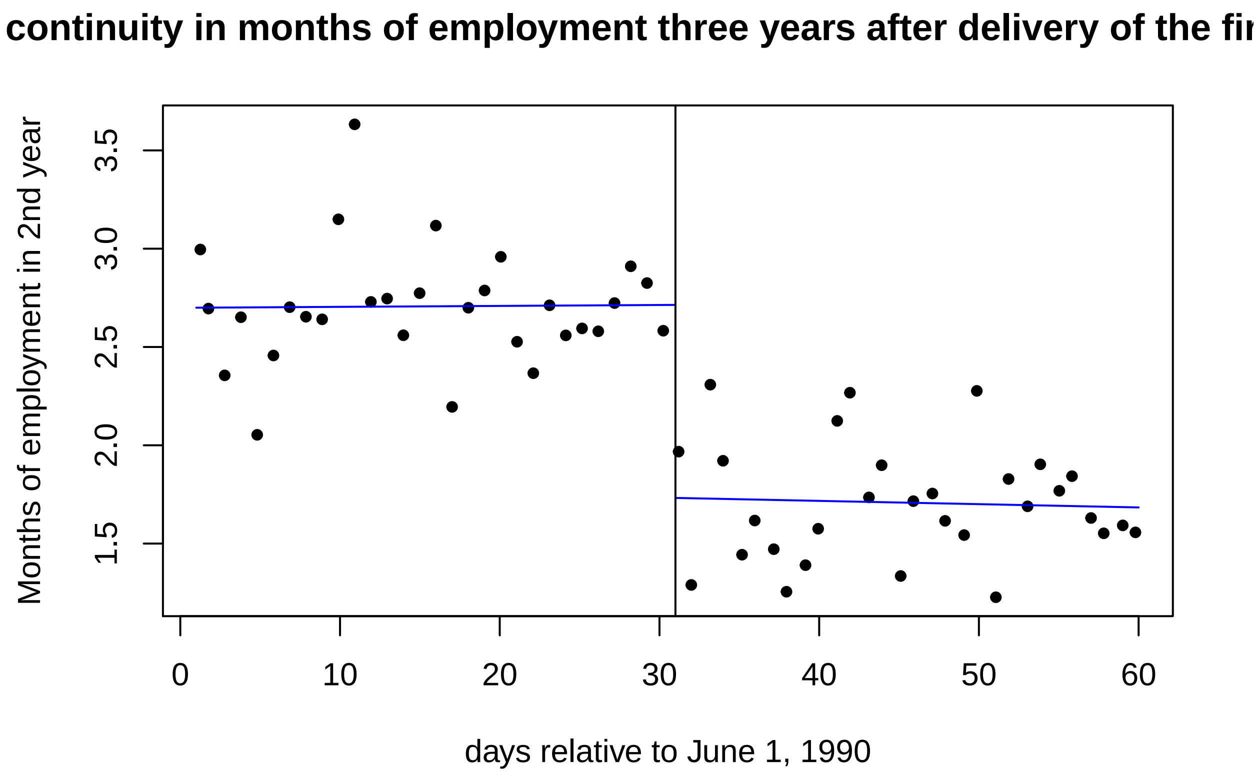
(a)

```r
with(parents, rdplot(y=ikar4, p=1, x=yday(bd)-151, c=31,
                     title="Discontinuity in take up of parental leave",
                     x.label="days relative to June 1, 1990",
                     y.label="PL in 2nd year"))
```



(b)

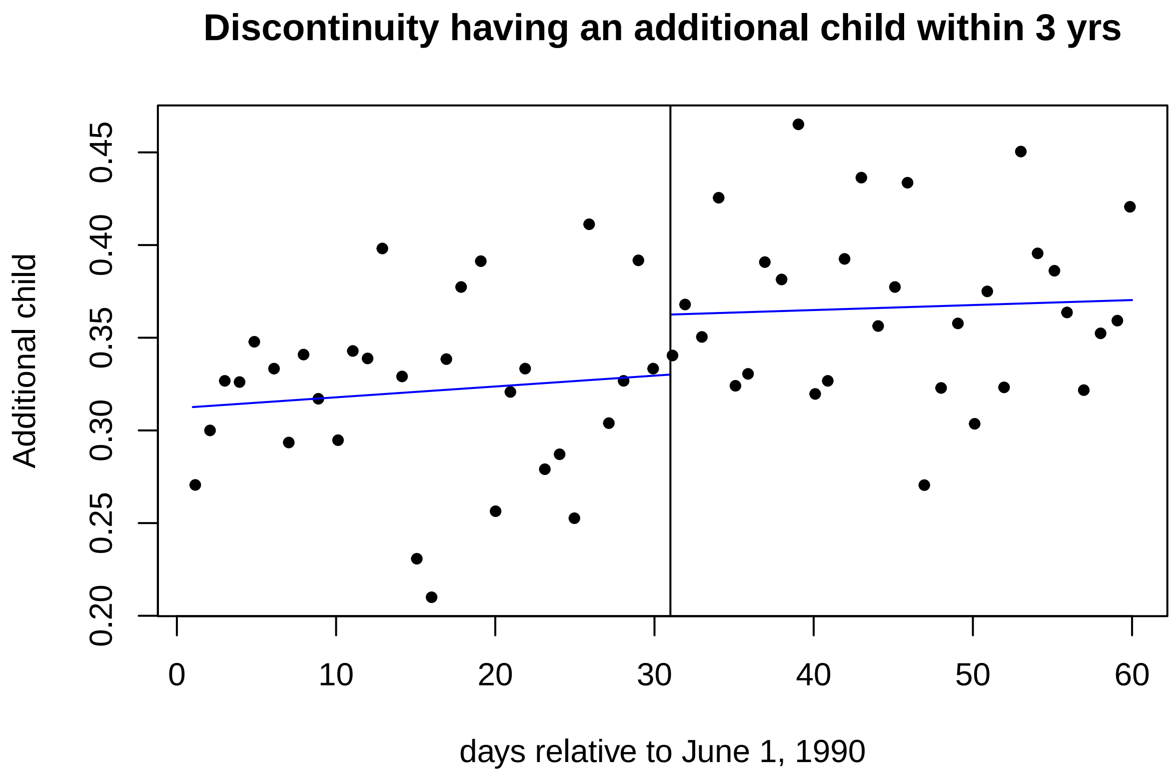
```r
with(parents, rdplot(y=pbexp3, p=1, x=yday(bd)-151, c=31,
                     title="Discontinuity in months of employment three years after delivery of the fir
                     x.label="days relative to June 1, 1990",
                     y.label="Months of employment in 2nd year"))
```



This graph shows us that the months of employment go down after the policy change. It appears to go down with a full month, as is also shown in the statistics in Table V on page 1389 of the paper.

V. Estimation of the treatment effect

```
with(parents, rdplot(y = uncb3, p=1, x = yday(bd)-151, c=31,
title="Discontinuity having an additional child within 3 yrs",
x.label="days relative to June 1, 1990", y.label="Additional child"))
```



(a)

Y = a + Bx + e

(b)

```
reg4a1 <- lm(uncb3 ~ july, data=parents)
summ(reg4a1)
```

```
## MODEL INFO:
## Observations: 6180
## Dependent Variable: uncb3
## Type: OLS linear regression
##
## MODEL FIT:
## F(1,6178) = 14.07, p = 0.00
## R² = 0.00
## Adj. R² = 0.00
##
## Standard errors: OLS
## ------------------------------------------------
##                      Est.   S.E.   t val.      p
## ----------------- ------ ------ -------- ------
## (Intercept)          0.32   0.01    36.76   0.00
## july                 0.05   0.01     3.75   0.00
## ------------------------------------------------
```

This effect is very small in size (0.05), yet significant. It is the same as in Table III, just rounded to two decimals.

(c)

```
reg4a2 <- lm(uncj3 ~ july, data=parents)
summ(reg4a2)
```

```
## MODEL INFO:
## Observations: 6180
## Dependent Variable: uncj3
## Type: OLS linear regression
##
## MODEL FIT:
## F(1,6178) = 74.97, p = 0.00
## R² = 0.01
## Adj. R² = 0.01
##
## Standard errors: OLS
## ------------------------------------------------
##                      Est.   S.E.   t val.      p
## ----------------- -------- ------ -------- ------
## (Intercept)          0.62   0.01    68.43   0.00
## july                -0.11   0.01    -8.66   0.00
## ------------------------------------------------
```

This effect is again small in size, but now points in the opposite direction (negative relationship). The effect is again significant. The result compares to Table V, rounded to two decimals.

(d)

```
reg4a3 <- lm(pbexp10 ~ july, data=parents)
reg4a4 <- lm(pbinc_tot10 ~ july, data=parents)
summ(reg4a3)
```

```
## MODEL INFO:
## Observations: 6180
## Dependent Variable: pbexp10
## Type: OLS linear regression
##
## MODEL FIT:
## F(1,6178) = 0.39, p = 0.53
## R² = 0.00
## Adj. R² = -0.00
##
## Standard errors: OLS
## -----------------------------------------------
##                     Est.   S.E.   t val.      p
## ----------------- ------ ------ -------- ------
## (Intercept)         5.10   0.08    63.33   0.00
## july                0.07   0.11     0.63   0.53
## -----------------------------------------------
```

```r
summm(reg4a4)
```

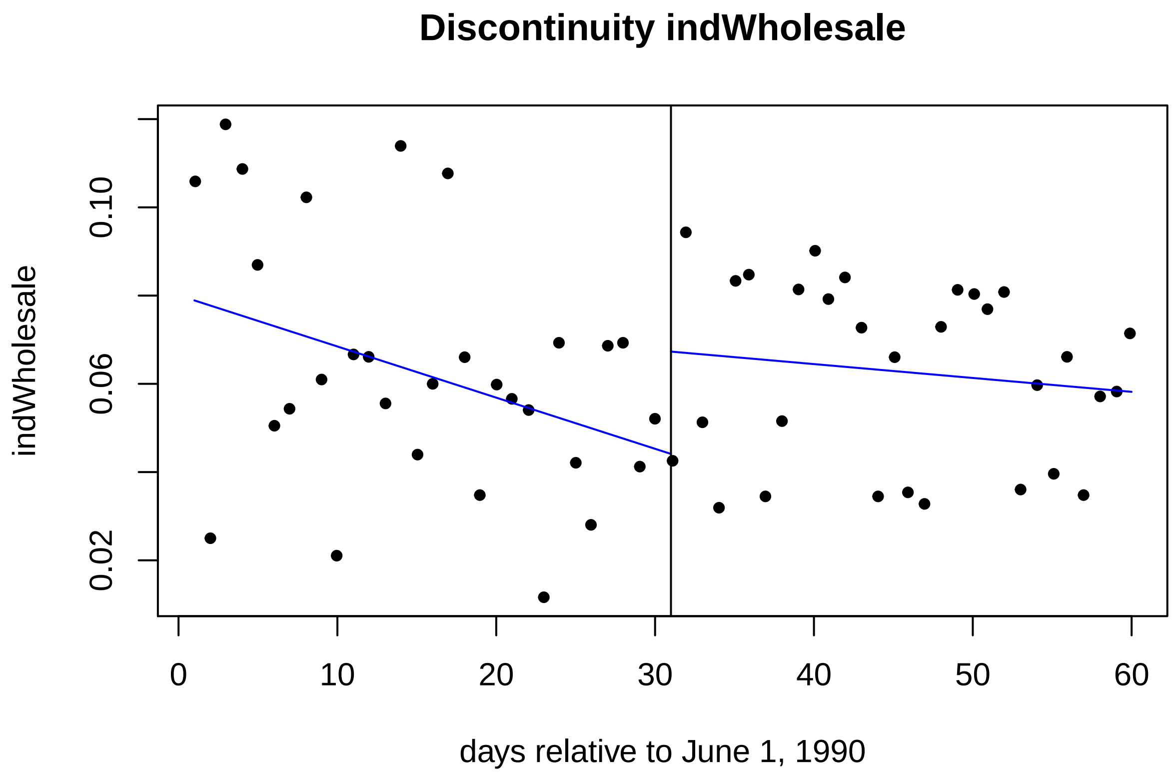
```
## MODEL INFO:
## Observations: 6180
## Dependent Variable: pbinc_tot10
## Type: OLS linear regression
##
## MODEL FIT:
## F(1,6178) = 2.28, p = 0.13
## R² = 0.00
## Adj. R² = 0.00
##
## Standard errors: OLS
## ------------------------------------------------
##                       Est.   S.E.   t val.      p
## ----------------- -------- ------ -------- ------
## (Intercept)          20.31   0.41    49.85   0.00
## july                  0.85   0.56     1.51   0.13
## ------------------------------------------------
```

Months of employment is barely affected, and is not significant. Thus, it is likely that the treatment had no effect on future months of employment. Monthly earnings later on had a decently sized effect, however this is also not significant.

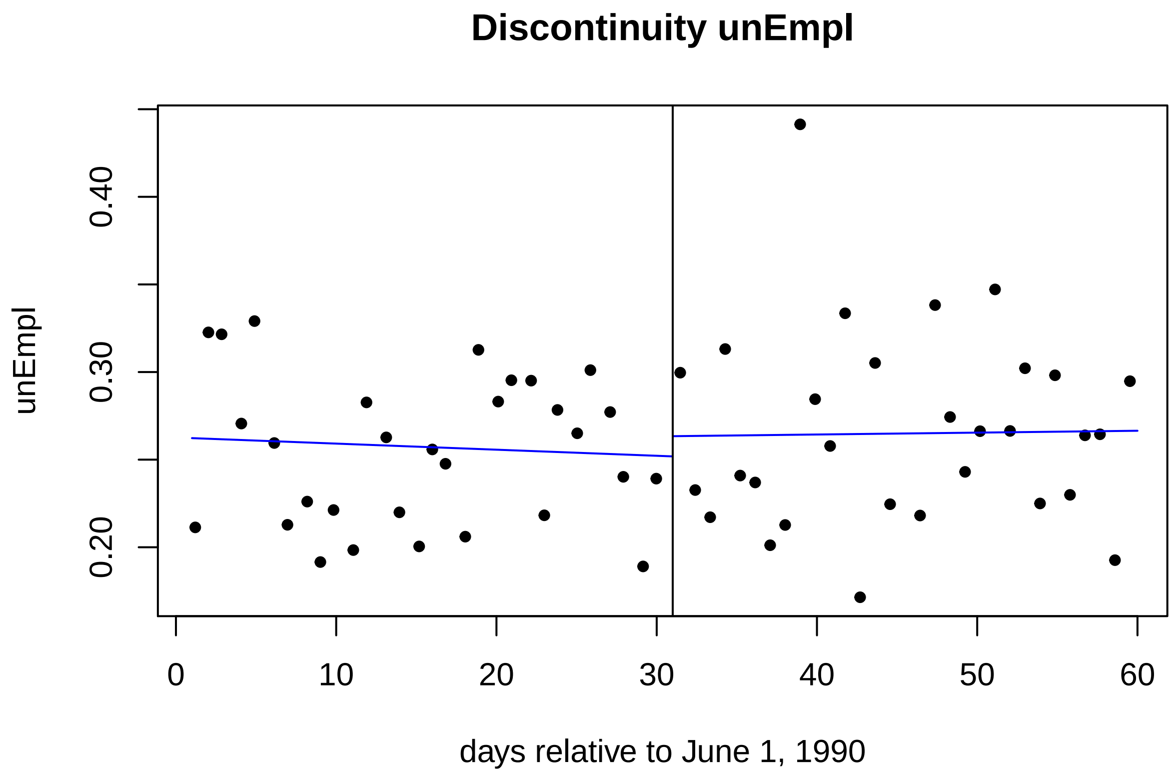
VI. Check on validity of the approach

(a)

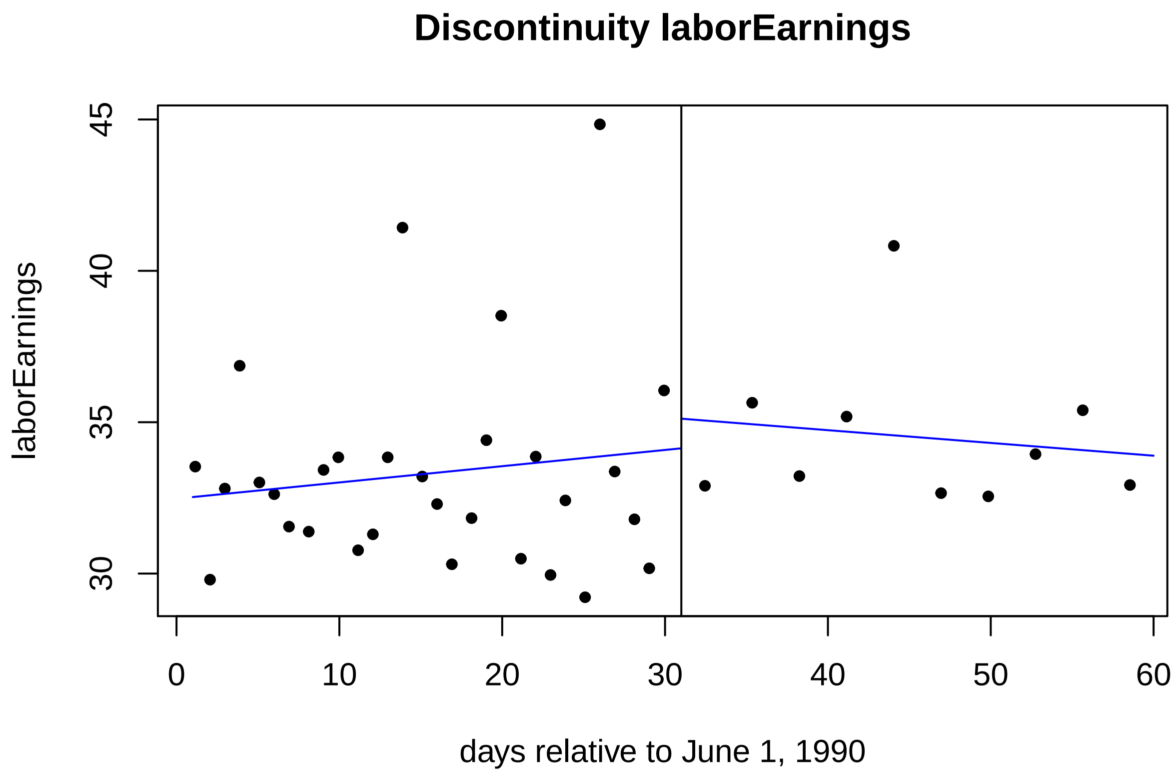
```r
with(parents, rdplot(y = indWholesale, p=1, x = yday(bd)-151, c=31,
title="Discontinuity indWholesale",
x.label="days relative to June 1, 1990", y.label="indWholesale"))
```




```
with(parents, rdplot(y = unEmpl, p=1, x = yday(bd)-151, c=31,
title="Discontinuity unEmpl",
x.label="days relative to June 1, 1990", y.label="unEmpl"))
```

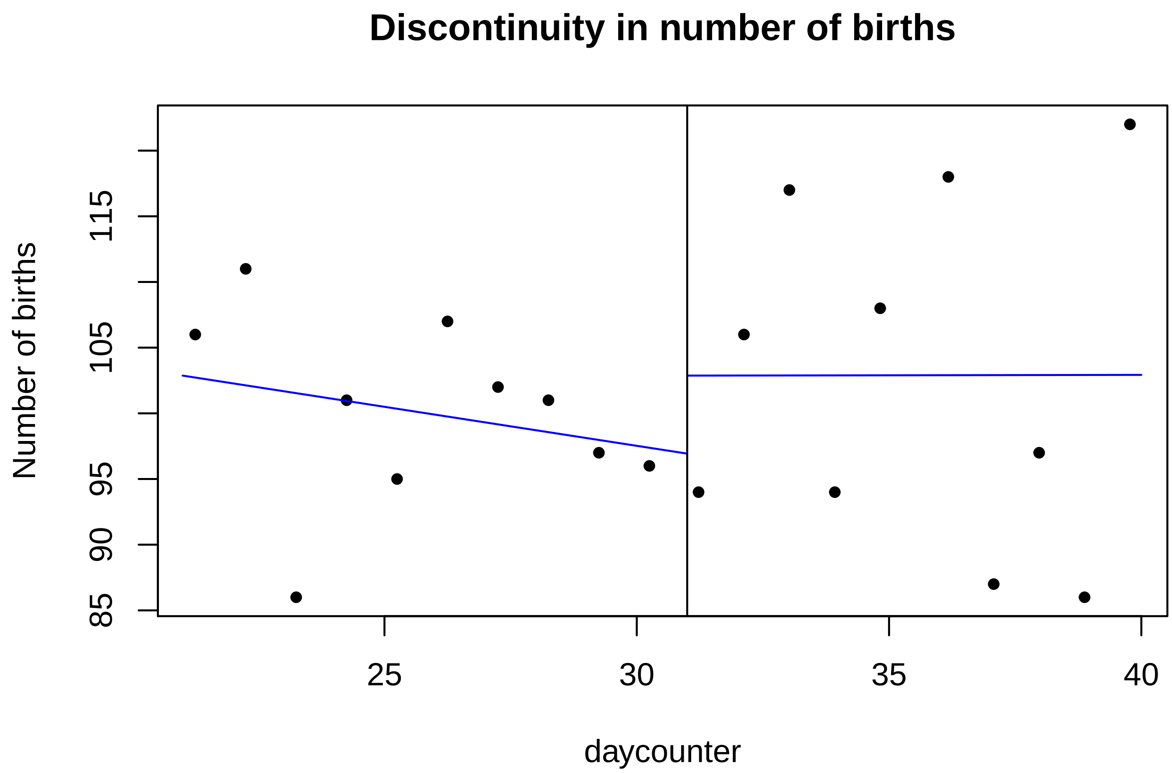



```
with(parents, rdplot(y = laborEarnings, p=1, x = yday(bd)-151, c=31,
title="Discontinuity laborEarnings",
x.label="days relative to June 1, 1990", y.label="laborEarnings"))
```



(b)

```
with(obsperday, rdplot(y=n, p=1, x=daycounter, c=31, nbins=20,
                       subset=obsperday$daycounter[21:40],
                       title="Discontinuity in number of births",
                       x.label="daycounter", y.label="Number of births"))
```



The graph suggests a downward trend in the number of births before July 1st, and a completely random number of births after this date. Thus, it is likely that there has been some self-selection. Some women knew of the planned policy change, and somehow manipulated the moment of birth to be after the 31st of June. This could create a small bias in the estimated treatment effect, as the women who were aware of the policy change may have been more in the know about politics, meaning they were higher educated (on average).

```
obsperday$july <- as.numeric(obsperday$daycounter>=31)
```

```
reg4a5 <- lm(n ~ july, data=obsperday, subset=(obsperday$daycounter[21:40]))
summ(reg4a5)
```

```
## MODEL INFO:
## Observations: 20
## Dependent Variable: n
## Type: OLS linear regression
##
## MODEL FIT:
## F(1,18) = 0.33, p = 0.58
## R² = 0.02
## Adj. R² = -0.04
##
## Standard errors: OLS
## ---------------------------------------------------
##                         Est.   S.E.   t val.      p
## ----------------- -------- ------ -------- ------
```

```
## (Intercept)         100.20   3.35    29.94   0.00
## july                  2.70   4.73     0.57   0.58
## --------------------------------------------------
```

Since p is 0.58, this effect does not seem to be statistically significant.