

## Code Replication #9

Andres Perez

2022-12-01

### 1) Load Data and libraries

```
suppressMessages(library(AER))

load('E:/CalPoly SLO/Fall Quarter 2022/Advanced Econometrics 1/Working
directory (datafiles too)/births.RData')

df = births
```

*a) In one to two sentences, state in your own words, what research question, specifically what causal effect, is this paper trying to address.*

This research journal studies the impact of a universal child benefit on fertility and maternal labor supply. We see a decrease in the time a kid spent in child care as the mother had more paid time off to take care of her kids, and we also saw an increase in Spain's fertility due to the benefits with most of these increases being to a decrease in abortions.

*b) How many observations are in the data set, and how many variables are in the data set?*

```
n = nrow(df)
k = ncol(df)
```

There are 120 observations and 4 variables

*c) Rather than replicate summary statistics in Table 1, instead calculate the total number of conceptions in 2005?*

```
x = subset(df, year==2005)
xx = x$conceptions
xx = sum(xx)
```

There were 471196 of conceptions during 2005

*d) Table 2 conducts the regression discontinuity analysis of the effect of the policy on both conceptions and abortions (each row corresponds to a separate analysis). Replicate the estimated causal effect on conceptions for the specification in column 2. Note, if the bandwidth is 6, you should include 6 month before July 2007, July 2007, and the 5 months after July 2007, so you would have 12 observations total. What is the estimate of the treatment effect on the outcome? Interpret your coefficient estimate of the main effect. Is this result statistically significant? (note: the robust standard errors are calculated in the table)*

```
df$newvar = 1:n #Basically replicating index
df$newvar = df$newvar - 91 #Centering
df$lnconceptions = log(df$conceptions)
```

```

df$treat = df$newvar >= 0 #treat
df$treat = ifelse(df$treat == "TRUE", 1, 0)

df$newvar2 = df$newvar^2

df2 = subset(df, newvar >= -30 & newvar <= 29) # #Bandwith

model1 = lm(lnconceptions ~ treat + daysinmonth + I(newvar*treat) +
I(newvar2*treat) + newvar + newvar2, data = df2)
round(coeftest(model1, vcov = vcovHC(model1, type="HC1")), 4)

##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    10.6400    0.1120  95.0308   <2e-16 ***
## treat           0.0695    0.0248   2.7961   0.0072 **
## daysinmonth    -0.0005    0.0038  -0.1252   0.9009
## I(newvar * treat) -0.0101    0.0040  -2.5313   0.0144 *
## I(newvar2 * treat) 0.0002    0.0001   1.2218   0.2272
## newvar          0.0017    0.0027   0.6516   0.5174
## newvar2         0.0000    0.0001  -0.1356   0.8926
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

The treatment effect I got for Table 2, column 2 is .0695 with a standard error of .0278. This means that for every one unit increase in the treatment effect variable, conceptions in Spain will increase by 6.95%.

**e) In a paragraph, summarize your conclusion from the analysis above. What was the purpose of the exercise and what were the findings from this analysis. Comment. (note: you have just replicated part of the paper's analysis, I only want you to comment on the findings of the analysis you did, not the other analysis that was in the paper that were not part of the replication.)**

The purpose of code replication 9 was for us to do some regression discontinuity exercises and some R practice too (question c). With the exercises in this analysis we found that there were 495,699 conceptions in 2005 and that the treatment effect in table 2 #>column 2 #>Like stated before, the research found that the 2,500 euros given to soon-to-be mothers increase fertility in Spain and we see a decrease in the time a kid spent in child care as the mother had more paid time off to take care of her kids. The increase in Spain's fertility is mostly due to a decrease in abortions thanks to the benefits provided by the government.