

Ectonometrics 2 CA1b

Ties van der Veen

2-2-2020

Rebecca Costello, Ties van der Veen, Matei van der Meer

I. Preliminaries

```
#Clearing the workspace
```

```
rm(list=ls())
```

```
#Calling packages in library
```

```
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(lubridate)
library(ggplot2)
library(interflex)
library(plm)
library(margins)
```

```
#Improving layout of summarise command
```

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

```
#Telling summarytools we're in R markdown
```

```
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;
```

II. Introduction to the computer assignment

#Obtaining the data

```
theUrl_calb_ectrics2 <- "https://surfdrive.surf.nl/files/index.php/s/aaEVKXoZ3VTU bq3/download"
waste2 <- read.dta (file = theUrl_calb_ectrics2)
```

#Generating a dichotomous variable from a continuous variable (pre-treatment level of unseperated waste

```
high_sorting <- as.numeric(waste2$sorting >= 61)
waste2$high_sorting=high_sorting
waste2$low_sorting=(1-waste2$high_sorting)
```

III. What to submit

A. Binary intercation variable

(a)

#Regression for the fully interacted model

```
ate3 <- plm(residual_weight ~ treatment*high_sorting, data=waste2, effect = "twoways",
            model = "within", index = c("route", "week"))
coefest(ate3, vcov=vcovHC(ate3, cluster= "group"))
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## treatment          -1.015812   0.090979 -11.165  <2e-16 ***
## treatment:high_sorting -0.040319   0.103127  -0.391   0.6958
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(b)

High_sorting being 0 would give a CATE of just the treatment variable, so -1.015812.

(c)

High_sorting being 1 would give a CATE of -0.040319

(d)

Yes, if people have low sorting then we expect to see a bigger effect from the treatment than if they already sort well (and thus are high sorting).

(e)

The treatment variable itself is significant, yet the treatment*interaction term coefficient is not. So it seems to be unlikely that this would be statistically significant.

B. Continuous interaction variable

```
#Re-importing the data to get the continuous variable 'sorting' back
```

```
theUrl_ca1b_ectrics2 <- "https://surfdrive.surf.nl/files/index.php/s/aaEVKXoZ3VTUbq3/download"
waste2 <- read.dta (file = theUrl_ca1b_ectrics2)
```

(f)

```
ate4 <- plm(residual_weight ~ treatment*sorting, data=waste2, effect = "twoways",
            model = "within", index = c("route", "week"))
coeftest(ate4, vcov=vcovHC(ate4, cluster= "group"))
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## treatment      -0.0989032  0.4281904 -0.2310  0.81735
## treatment:sorting -0.0156784  0.0069609 -2.2524  0.02437 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(g)

```
# print(-0.0989032(mean(waste2$sorting)*(-0.0156784)))
# tried this function and it didn't work, not sure why.
```

(h)

The interaction term is significant, showing that the treatment and sorting level have a statistically significant effect together.

(i)

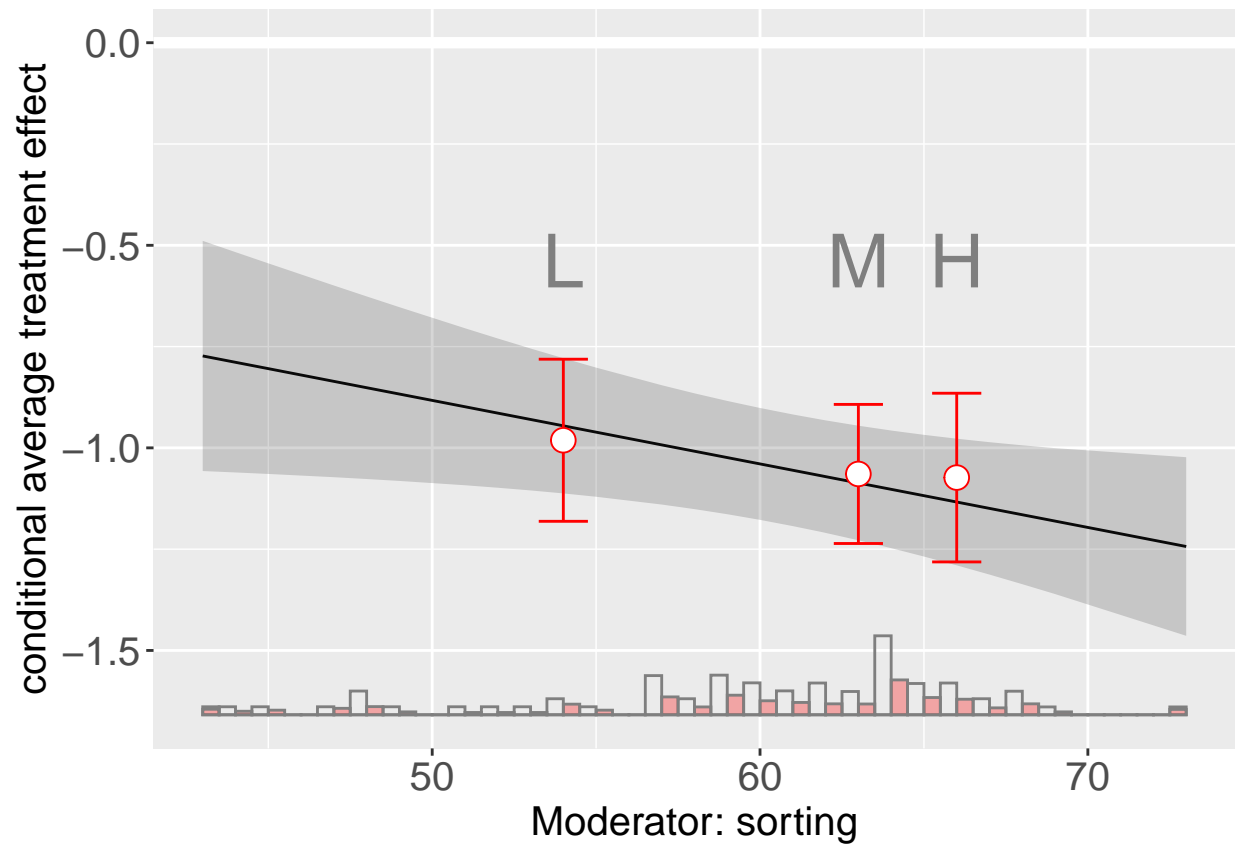
```
out2 <- inter.binning(Y = "residual_weight", D = "treatment", X = "sorting",
                     Z = NULL, FE=c("route", "week"), data=waste2, na.rm=TRUE,
                     vartype="cluster", cl="route", main=NULL,
                     wald=TRUE, ylab="conditional average treatment effect")
```

```
## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either
## rank-deficient or indefinite
```

```
## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either
## rank-deficient or indefinite
```

```
## Warning in chol.default(mat, pivot = TRUE, tol = tol): the matrix is either
## rank-deficient or indefinite
```

out2\$graph



(j)

The low value bin has a higher CATE than the other two, though the interval has a lot of common area between the three. So the difference is actually not very big.