---

title: "Robustness Tests"

author: "Greta Lotz"

date: "2024-04-20"

output: html\_document

---

```{r setup, include=FALSE}

knitr::opts\_chunk$set(echo = TRUE)

```

Run Thesis.Rmd and SCM\_oecd.Rmd first

```{r}

library(Synth)

```

```{r}

# Assuming Norway's identifier is 10 and should not be included in the control group

control\_countries <- c(1:9, 11:15) # Corrected to exclude Norway's identifier

# Initialize a dataframe to store the results

results <- data.frame(Year = 1960:2005) # Adjust years as necessary

# Define the names of the countries corresponding to identifiers for clarity in naming columns

country\_names <- c("Australia", "Belgium", "Canada", "Denmark", "France", "Greece",

"Iceland", "Japan", "New Zealand", "Norway", "Poland", "Portugal",

"Spain", "Switzerland", "United States")

names\_to\_use <- country\_names[control\_countries] # Filter names to only include control countries

# Loop through each control country to exclude it and perform synthetic control analysis

for (i in seq\_along(control\_countries)) {

excluded\_country <- control\_countries[i]

# Prepare the data excluding one country

controls\_used <- control\_countries[control\_countries != excluded\_country]

dataprep.out <- dataprep(foo = carbontax\_updated,

predictors = c("GDP\_per\_capita", "gas\_cons\_capita", "vehicles\_capita", "urban\_pop"),

predictors.op = "mean",

time.predictors.prior = 1980:1990,

special.predictors = list(

list("CO2\_transport\_capita", 1990, "mean"),

list("CO2\_transport\_capita", 1980, "mean"),

list("CO2\_transport\_capita", 1970, "mean")

),

dependent = "CO2\_transport\_capita",

unit.variable = "Countryno",

unit.names.variable = "country",

time.variable = "year",

treatment.identifier = 10,

controls.identifier = controls\_used,

time.optimize.ssr = 1960:1990,

time.plot = 1960:2005

)

synth.out <- synth(data.prep.obj = dataprep.out, method = "All")

# Calculate synthetic control outcomes

synthetic\_outcomes <- dataprep.out$Y0plot %\*% synth.out$solution.w

# Store the results in the dataframe under the appropriate country name

results[[paste("excl", names\_to\_use[i], sep = "\_")]] <- as.vector(synthetic\_outcomes)

}

# Save or process the results dataframe as needed

save(results, file = "leave\_one\_out\_results.RData")

```

```{r}

leave\_one\_out\_path\_data <- merge(path\_data, results, by = "Year", all = TRUE) %>%

rename(Norway = X10,

Synthetic\_Norway = Synthetic)

# List of countries with weight greater than 0.001, based on your uploaded image

significant\_countries <- c("Canada", "Denmark", "France", "Greece" ,"Portugal", "Spain" ,"Switzerland", "United States", "Japan")

# Create a vector of column names to keep, based on significant countries

columns\_to\_keep <- c("Year", "Norway", "Synthetic\_Norway",

paste("excl", significant\_countries, sep = "\_"))

# Filter the leave\_one\_out\_path\_data to keep only the significant countries

significant\_leave\_one\_out <- leave\_one\_out\_path\_data[columns\_to\_keep]

# Reshape the filtered data for plotting

significant\_leave\_one\_out\_long <- melt(significant\_leave\_one\_out, id.vars = "Year",

variable.name = "Condition", value.name = "CO2\_per\_capita")

```

```{r}

# Now plot using ggplot2

leave\_one\_out\_plot <- ggplot(significant\_leave\_one\_out\_long, aes(x = Year, y = CO2\_per\_capita, color = Condition)) +

geom\_line() +

labs(title = "Leave-One-Out Analysis for Synthetic Norway with Significant Weights",

x = "Year",

y = "CO2 Emissions per Capita (metric tons)",

color = "Condition") +

theme\_minimal() +

theme(legend.position = "right") +

scale\_color\_brewer(palette = "Set1")

# Print the plot

print(leave\_one\_out\_plot)

```

```{r GDP as outcome variable}

dataprep.out <-

dataprep(foo = carbontax\_updated,

predictors = c("CO2\_transport\_capita" ,"gas\_cons\_capita" , "vehicles\_capita" ,

"urban\_pop") ,

predictors.op = "mean" ,

time.predictors.prior = 1980:1990 ,

special.predictors = list(

list("GDP\_per\_capita" , 1990 , "mean"),

list("GDP\_per\_capita" , 1980 , "mean"),

list("GDP\_per\_capita" , 1970 , "mean")

),

dependent = "GDP\_per\_capita",

unit.variable = "Countryno",

unit.names.variable = "country",

time.variable = "year",

treatment.identifier = 10, # no 10 Norway

controls.identifier = c(1:9, 12:15),

time.optimize.ssr = 1960:1990,

time.plot = 1960:2005

)

###################################################

synth.out <- synth(data.prep.obj = dataprep.out,

method = "All")

###################################################

synth.tables <- synth.tab(dataprep.res = dataprep.out,

synth.res = synth.out

)

###################################################

### Table 1: CO2 Emissions From Transport Predictor Means Before Tax Reform

###################################################

synth.tables$tab.pred[1:7, ]

###################################################

### Table 2: Country Weights in Synthetic Norway

###################################################

synth.tables$tab.w[1:14, ]

```

```{r pull path plot data}

synthetic\_control\_outcomes <- dataprep.out$Y0plot %\*% synth.out$solution.w

# Construct the Year vector from 1960 to 2005

years <- 1960:2005

# Create the comparison dataframe

gdp\_path\_data <- data.frame(

Year = years,

Actual = dataprep.out$Y1plot,

Synthetic = as.vector(synthetic\_control\_outcomes) # Convert matrix to vector

)

```

```{r pull gap plot data}

# Calculate synthetic control outcomes

synthetic\_control\_outcomes <- dataprep.out$Y0plot %\*% synth.out$solution.w

# Calculate the gap between Actual and Synthetic

gap <- dataprep.out$Y1plot - synthetic\_control\_outcomes

# Construct the Year vector from 1960 to 2005

years <- as.numeric(rownames(dataprep.out$Y0plot)) # Ensure years are correctly defined

# Create the gap\_data data frame

gdp\_gap\_data <- data.frame(

Year = years,

Gap = as.vector(gap) # Convert matrix to vector if not already

)

```

```{r}

gdp\_path\_plot <- ggplot(gdp\_path\_data, aes(x = Year)) +

geom\_line(aes(y = X10, color = "Norway", linetype = "Norway"), size = 0.5) + # Solid line for Norway (Actual data)

geom\_line(aes(y = Synthetic, color = "Synthetic Norway", linetype = "Synthetic Norway"), size = 0.5) + # Dashed line for Synthetic control

geom\_vline(xintercept = 1991, linetype = "dotted", color = "black", size = 0.5) + # Dotted line at the year 1990

geom\_rect(aes(xmin = 1987, xmax = 1992, ymin = -Inf, ymax = Inf), fill = "grey", alpha = 0.01) + # Shaded area for the recession

labs(title = NULL,

x = NULL,

y = "GDP",

color = "Legend",

linetype = "Legend") + # Label for linetype legend

theme\_classic() +

theme(legend.position = c(0.15, 0.13), # Position the legend at the bottom left

legend.title = element\_blank(), # Remove the legend title for simplicity

legend.text = element\_text(size = 8), # Set text size for legend

legend.background = element\_rect(fill = "white", colour = "black"),

legend.key = element\_blank()) + # No boxes around the legend symbols

annotate("text", x = 1986, y = max(gdp\_path\_data$Synthetic, na.rm = TRUE) \* 0.27, label = "Carbon tax", hjust = 0, size = 3, color = "black") +

scale\_y\_continuous(limits = c(0, 50000), breaks = seq(0, 50000, by = 10000)) +

scale\_x\_continuous(limits = c(1960, 2005), breaks = seq(1960, 2005, by = 5)) +

scale\_color\_manual(values = c("Norway" = "black", "Synthetic Norway" = "black")) + # Set line colors

scale\_linetype\_manual(values = c("Norway" = "solid", "Synthetic Norway" = "dashed")) # Set line types

gdp\_path\_plot

```

```{r}

gdp\_gap\_data <- gdp\_gap\_data %>%

rename(gdp\_gap = Gap)

# Combine the data frames into one for plotting

combined\_data <- merge(gap\_data, gdp\_gap\_data, by = "Year")

# Rename columns as necessary

names(combined\_data)[names(combined\_data) == "gdp\_gap"] <- "GDP\_Deviation"

names(combined\_data)[names(combined\_data) == "Gap"] <- "CO2\_Deviation"

```

```{r gdp and co2 deviation from synthetic}

library(ggplot2)

# Assuming 'combined\_data' is already your merged and cleaned data frame

# Calculate the range for each variable to set the secondary axis appropriately

gdp\_range <- range(combined\_data$GDP\_Deviation, na.rm = TRUE)

co2\_range <- range(combined\_data$CO2\_Deviation, na.rm = TRUE)

# The ratio will help to align the secondary axis

ratio <- gdp\_range[2] / co2\_range[2]

# Create the combined plot

combined\_plot <- ggplot(combined\_data, aes(x = Year)) +

geom\_line(aes(y = GDP\_Deviation, color = "GDP Deviation")) +

geom\_line(aes(y = CO2\_Deviation \* ratio, color = "CO2 Deviation")) +

geom\_hline(yintercept = 0, linetype = "dashed", color = "black", size = 0.5) + # Corrected to horizontal line at y = 0

scale\_y\_continuous(

name = "GDP Deviation",

limits = range(c(gdp\_range, co2\_range \* ratio)), # Set limits to accommodate both variables

sec.axis = sec\_axis(~./ratio, name = "CO2 Deviation", breaks = seq(floor(co2\_range[1]), ceiling(co2\_range[2]), length.out = 5))

) +

labs(title = "Comparison of GDP and CO2 Emissions Deviations from Synthetic Control",

x = "Year",

color = "Variable") +

theme\_minimal() +

scale\_color\_manual(values = c("GDP Deviation" = "grey", "CO2 Deviation" = "black")) +

theme(

legend.title = element\_blank(),

axis.text.x = element\_text(angle = 45, hjust = 1), # To prevent overlapping x-axis labels

legend.position = "bottom"

)

# Print the plot

print(combined\_plot)

```