

---

## Economía Aplicada: Problem Set N°8

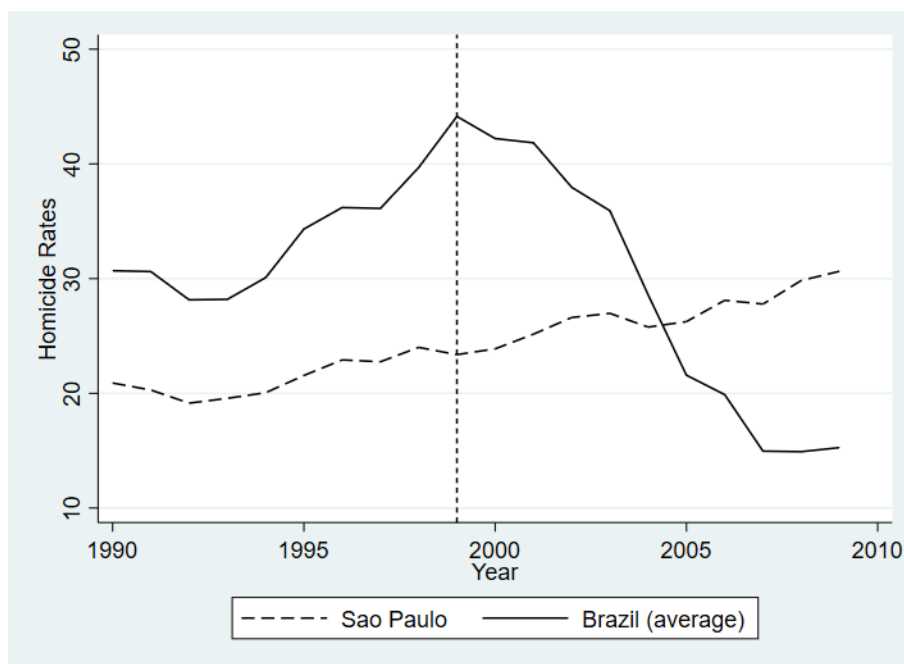
---

Milton Bronstein      Felipe García Vassallo      Santiago López      Franco Riottini

En este Problem Set vamos a replicar los primeros siete gráficos del paper “Evaluating the Effect of Homicide Prevention Strategies in São Paulo, Brazil: A Synthetic Control Approach” de Danilo Freire (2018).

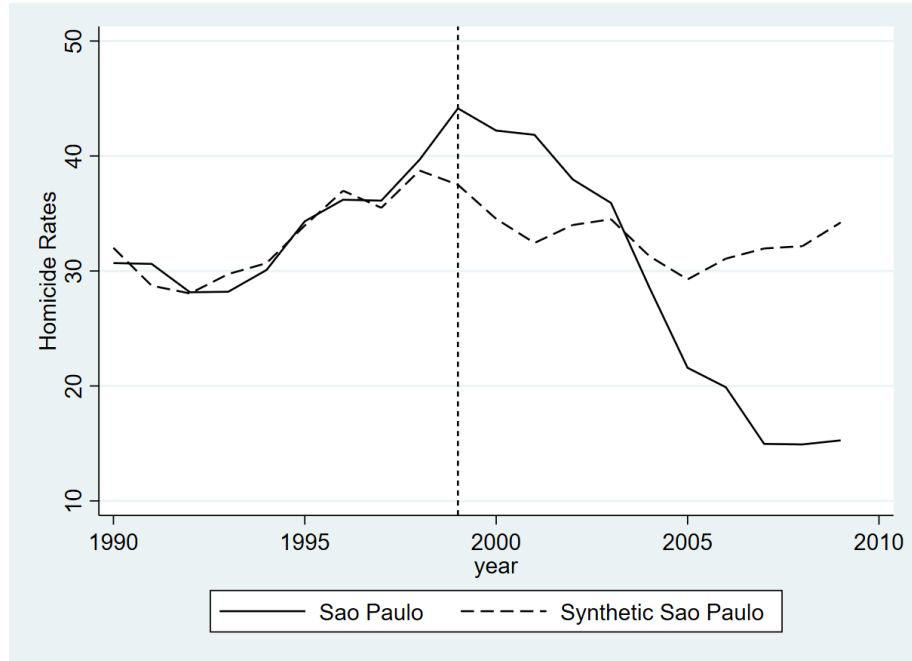
En el primer gráfico lo que observamos es que antes del año de tratamiento, en el Estado de San Pablo, la tasa de homicidios venía creciendo, y luego del cambio de política comienza a caer. En este primer gráfico se tiene como contrafactual al promedio del resto de los Estados de Brasil, en los que parece no haber cambiado la tendencia que venía teniendo la tasa de homicidios luego de 1999, el año de comienzo del tratamiento.

Figura 1: Tasas de homicidio por 100.000 habitantes: São Paulo y Brasil (excluyendo el estado de São Paulo).



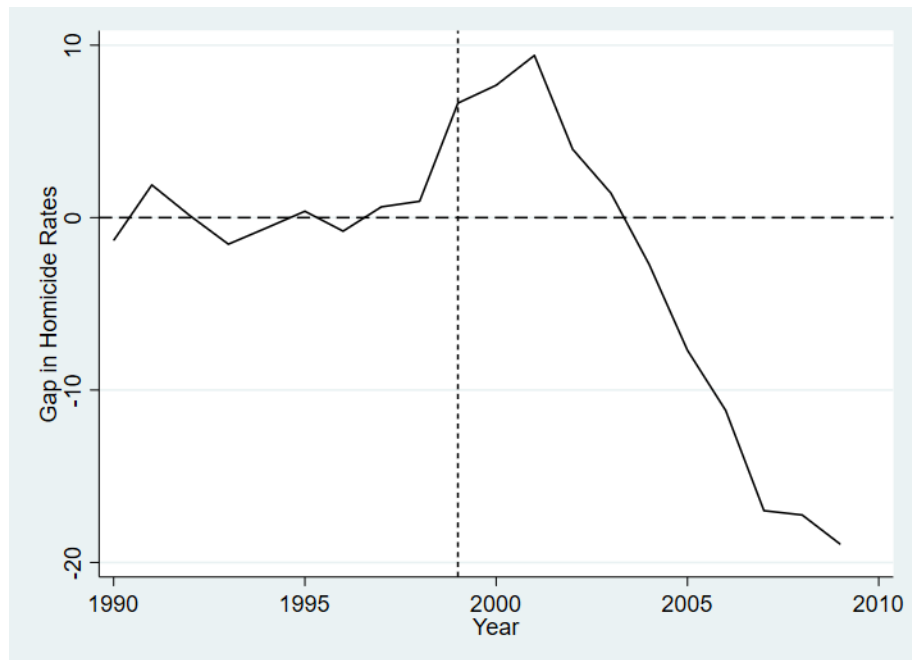
En el segundo gráfico se usa como contrafactual no al promedio simple del resto de Brasil, sino al promedio ponderado, que se convierte en un San Pablo sintético. Como se puede observar, previo a 1999 ambas tendencias tienen una figura similar, a pesar del hecho de que el sintético parece tener una caída previa a la que se observaría en San Pablo post-1999. Luego del año de aplicación de la política, se observa una divergencia en las tendencias de ambas unidades.

Figura 2: Tendencias en las tasas de homicidios: São Paulo versus São Paulo sintético.



En el siguiente gráfico, observamos la brecha en la tasa de homicidios entre San Pablo y su sintético. Podemos observar que, de forma consistente a lo expresado anteriormente, la brecha es cercana a cero hasta un año antes de la aplicación de la política analizada. Entre ese año y 2004, la brecha se mantuvo positiva, para luego tornarse negativa hasta finales del período de análisis. En el año 2009, la brecha entre San Pablo y su contrafactual sintético ascendía a casi 20 homicidios menos cada 100 mil habitantes.

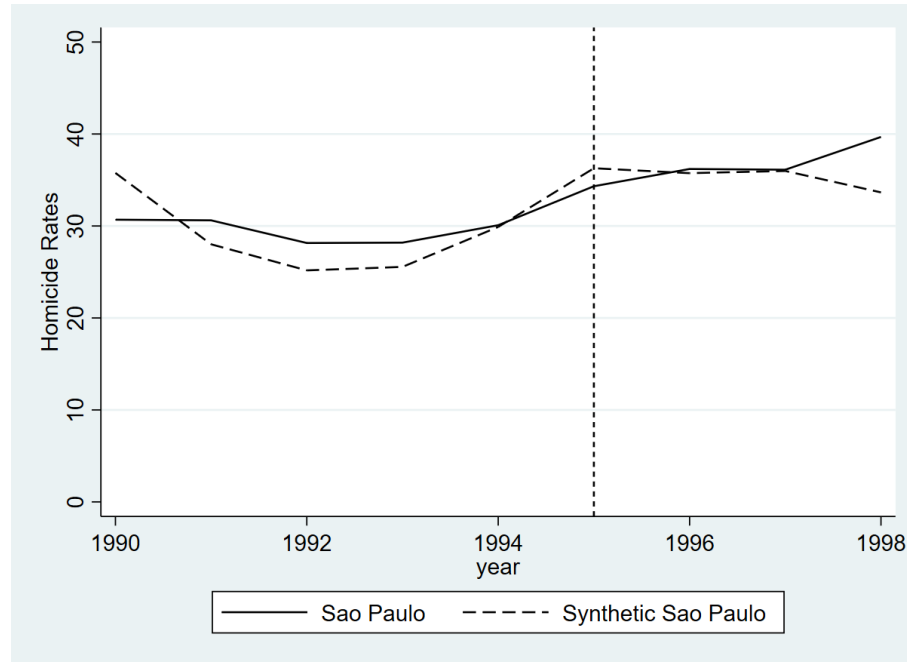
Figura 3: Brecha de tasas de homicidio entre São Paulo y São Paulo sintético.



## Robustness Checks

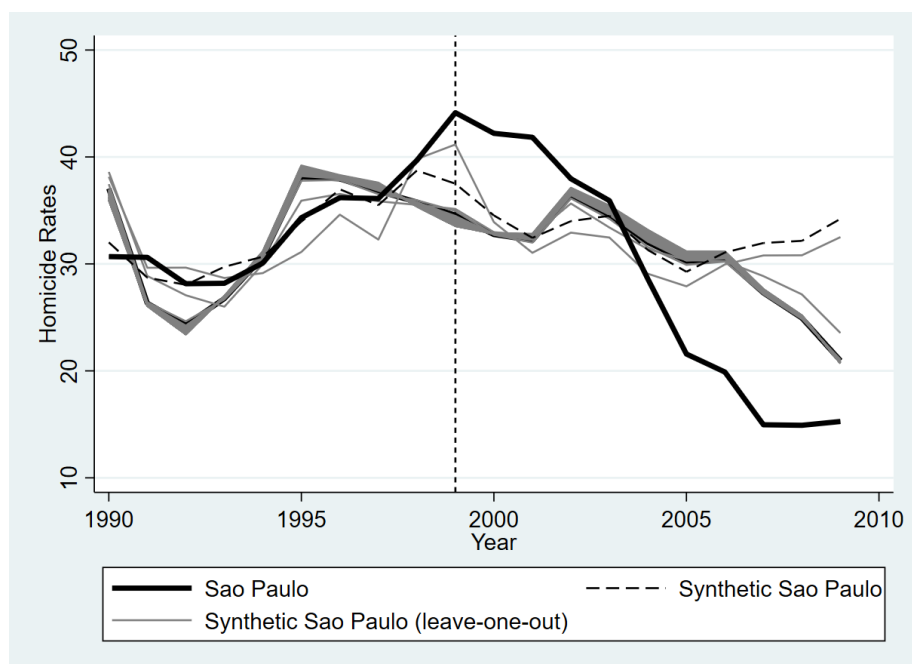
En el gráfico 4 se lleva a cabo un primer ejercicio de robustez que consiste en una prueba placebo. En vez de tratar a San Pablo en 1999 como efectivamente ocurrió, se desplaza el año de tratamiento a 1994 con el objetivo de comprobar si se observan efectos de tratamientos falsos en los años previos al mismo. Como se puede observar, si bien hay una brecha entre San Pablo y el control sintético, la misma no parece ser significativa, lo que da mayor confianza en los resultados principales del ejercicio.

Figura 4: Implementación de la política de placebo en 1994: São Paulo versus São Paulo sintético.



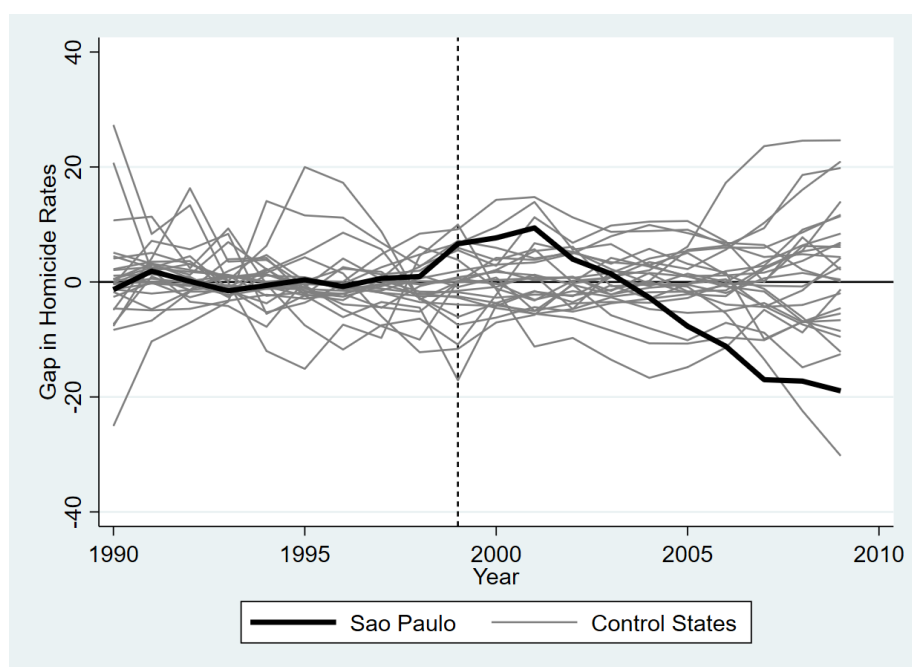
En el gráfico 5 se realiza una prueba de robustez eliminando de a uno los estados que componen el control sintético. De esta forma minimizamos el riesgo de que sea un solo estado, de los pertenecientes al control, el que esté impulsando nuestros resultados. De ser esto así, probablemente el contrafactual no sea razonable. Como se puede observar en el gráfico, ningún resultado está sesgando las estimaciones debido a que las tendencias del tratamiento y los controles se mantienen estables.

Figura 5: Distribución leave-one-out del control sintético para São Paulo



En el siguiente gráfico, lo que observamos es la diferencia en las tasas de homicidio entre las unidades tratadas y sus controles sintéticos. Sólo uno de los controles sintéticos tiene al final del período una brecha menor que San Pablo. El objetivo de este test, según Freire, es verificar si hay alguna otra tendencia inobservada a nivel nacional o regional que pueda estar impulsando el resultado original. El movimiento aleatorio de las líneas de control sintético nos indica que la dinámica de los resultados de San Pablo no puede ser explicada como parte de una tendencia más general.

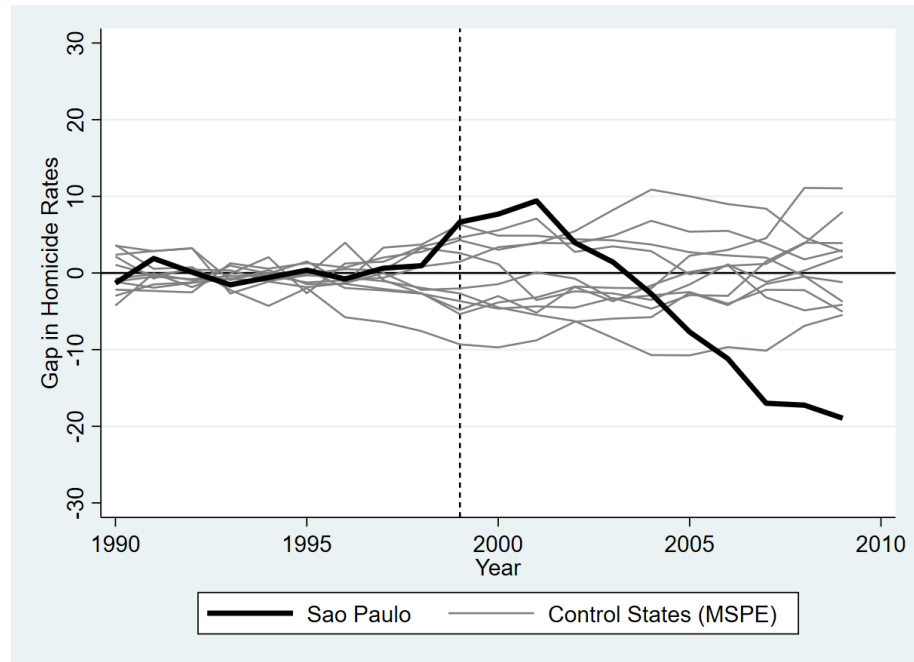
Figura 6: Permutation test: Brechas en la tasa de homicidios en São Paulo y veintiséis estados de control.



Por último, en el gráfico 7 vemos el mismo test que se presentó anteriormente, pero ahora para un grupo seleccionado de los controles sintéticos. La selección no es de manera aleatoria, sino que se seleccionan aquellos

sintéticos que tienen un error cuadrático medio predecido no superior al doble del de San Pablo. La intención es utilizar sólo los sintéticos con una buena coincidencia. Como se puede ver, en este caso la diferencia con los sintéticos es mayor y, como señala Freire, una mejor evidencia a favor de que existió una brecha considerable en las tasas de homicidios al final de período.

Figura 7: Permutation test: Brechas en la tasa de homicidios en São Paulo y estados de control seleccionados





```

/*****

```

# Semana 9: Control sintético

Universidad de San Andrés  
Economía Aplicada

2022

```

>
*****/
*      Bronstein      García Vassallo      López      Riottini
/*****/
Este archivo sigue la siguiente estructura:

```

0) Set up environment

1) Replicación de gráficos

```

*****/
* 0) Set up environment
*=====

```

```

global main "C:\Users\felip\Documents\UdeSA\Maestría\Aplicada\Problem-Sets\PS 8"
global output "$main/output"
global input "$main/input"

```

```

cd "$main"

```

\* 1) Empiezo a replicar

```

* Uso el csv en forma de panel
import delimited "$input/df.csv", encoding(UTF-8) clear

```

```

* Defino el panel
tsset code year

```

```

* Instalo el paquete para controles sintéticos
*ssc install synth

```

```

line homiciderates year

```

```

collapse (mean) homiciderates if code!=35, by(year)

```

```

gen code=1

```

```

save "$input/brasil.dta", replace

```

```

import delimited "$input/df.csv", encoding(UTF-8) clear

```

```

append using "$input/brasil.dta"

```

```

twoway (line homiciderates year if code==1, lcolor(grey) lpattern(dash)) (line homicid
> erates year if code==35, lcolor(black)), ytitle("Homicide Rates") xtitle("Year") xli
> ne(1999, lpattern(shortdash) lcolor(grey)) legend(label(1 "Sao Paulo") label(2 "Braz
> il (average)"))
graph export "$output/1.png", replace

```

\* Gráfico 2

```

drop if code==1

```

```

synth homiciderates yearsschoolingimp statgdpcapita homiciderates proportionextremepo
> verty giniimp populationprojectionln statgdpgrowthpercent, trunit(35) trperiod(1999
> ) nested fig keep(loo-resout28, replace)

```

\* Gráfico 3

```

matrix gaps=e(Y_treated) -e(Y_synthetic)
matrix Y_treated=e(Y_treated)
matrix Y_synthetic=e(Y_synthetic)
keep year
svmat gaps
svmat Y_treated
svmat Y_synthetic

twoway (line gaps1 year, lcolor(black)), xline(1999, lpattern(shortdash) lcolor(grey))
> yline(0, lpattern(dash) lcolor(black)) ytitle("Gap in Homicide Rates") xtitle("Year"
> )
graph export "$output/3.png", replace

* Gráfico 4

import delimited "$input/df.csv", encoding(UTF-8) clear

tsset code year

synth homiciderates yearsschoolingimp stategdpcapita homiciderates proportionextremepo
> verty giniimp populationprojectionln stategdpgrowthpercent, trunit(35) trperiod(1995
> ) resultsperiod(1990(1)1998) nested fig

* Gráfico 5

egen id=group(code)

save "$input/df.dta", replace

use "$input/df.dta", clear

tsset id year

cd "$input/loo"

tempname resmat
local i 20
qui synth homiciderates yearsschoolingimp stategdpcapita homiciderates proport
> ionextremepoverty giniimp populationprojectionln stategdpgrowthpercent, trunit(`i')
> trperiod(1999) keep(loo-resout`i', replace)

        forvalues j=1/27 {
            if `j'==20 {
                continue
            }
            use "$input/df.dta", clear
            tsset id year
            drop if id==`j'
            qui synth homiciderates yearsschoolingimp stategdpcapita homiciderates proport
> ionextremepoverty giniimp populationprojectionln stategdpgrowthpercent, trunit(20) t
> rperiod(1999) keep(loo-resout`j', replace)
        }

forvalues i = 1/28 {
    use "$input/loo/loo-resout`i'.dta", clear
    ren _Y_synthetic _Y_synthetic`i'
    ren _Y_treated _Y_treated`i'
    gen _Y_gap`i'=_Y_treated`i'-_Y_synthetic`i'
    save "$input/loo/loo-resout`i'.dta", replace
}

use "$input/loo/loo-resout1.dta", clear
forvalues i = 2/28 {
    merge 1:1 _Co_Number _time using "$input/loo/loo-resout`i'.dta", nogen
}

```

```

twoway (line _Y_synthetic_1 _time, lcolor(gray)) (line _Y_synthetic_2 _time, lcolor(gr
> ay)) (line _Y_synthetic_3 _time, lcolor(gray)) (line _Y_synthetic_4 _time, lcolor(gr
> ay)) (line _Y_synthetic_5 _time, lcolor(gray)) (line _Y_synthetic_6 _time, lcolor(gr
> ay)) (line _Y_synthetic_7 _time, lcolor(gray)) (line _Y_synthetic_8 _time, lcolor(gr
> ay)) (line _Y_synthetic_9 _time, lcolor(gray)) (line _Y_synthetic_10 _time, lcolor(g
> ray)) (line _Y_synthetic_11 _time, lcolor(gray)) (line _Y_synthetic_12 _time, lcolor
> (gray)) (line _Y_synthetic_13 _time, lcolor(gray)) (line _Y_synthetic_14 _time, lcol
> or(gray)) (line _Y_synthetic_15 _time, lcolor(gray)) (line _Y_synthetic_16 _time, lc
> olor(gray)) (line _Y_synthetic_17 _time, lcolor(gray)) (line _Y_synthetic_18 _time,
> lcolor(gray)) (line _Y_synthetic_19 _time, lcolor(gray)) (line _Y_synthetic_21 _time
> , lcolor(gray)) (line _Y_synthetic_22 _time, lcolor(gray)) (line _Y_synthetic_23 _ti
> me, lcolor(gray)) (line _Y_synthetic_24 _time, lcolor(gray)) (line _Y_synthetic_25 _ti
> me, lcolor(gray)) (line _Y_synthetic_26 _time, lcolor(gray)) (line _Y_synthetic_27
> _time, lcolor(gray)) (line _Y_treated_20 _time, lcolor(black) lwidth(thick)) (line
> _Y_synthetic_28 _time, lcolor(black) lpattern(dash)), xline(1999, lpattern(shortdash
> ) lcolor(grey)) legend(order(27 "Sao Paulo" 28 "Synthetic Sao Paulo" 3 "Synthetic Sa
> o Paulo (leave-one-out)")) xtitle("Year") ytitle("Homicide Rates")

```

\* Gráfico 6

```

use "$input/df.dta", clear
tsset id year

cd "$input/pt"

tempname resmat
local i 20
qui synth homiciderates yearsschoolingimp stategdpcapita homiciderates proport
> ionextremepoverty giniimp populationprojectionln stategdpgrowthpercent, trunit(`i')
> trperiod(1999) keep(resout`i', replace)
matrix `resmat' = nullmat(`resmat') \ e(RMSPE)
local names ``names' ``i'""
mat colnames `resmat' = "RMSPE"
mat rownames `resmat' = `names'
matlist `resmat' , row("Treated Unit")

drop if id==20

forvalues i = 1/27 {
    if `i'==20 {
        continue
    }
    qui synth homiciderates yearsschoolingimp stategdpcapita homiciderates proport
> ionextremepoverty giniimp populationprojectionln stategdpgrowthpercent, trunit(`i')
> trperiod(1999) keep(resout`i', replace)
matrix `resmat' = nullmat(`resmat') \ e(RMSPE)
local names ``names' ``i'""
}

mat colnames `resmat' = "RMSPE"
mat rownames `resmat' = `names'
matlist `resmat' , row("Treated Unit")

forvalues i = 1/27 {
    use "$input/pt/resout`i'.dta", clear
    ren _Y_synthetic _Y_synthetic_`i'
    ren _Y_treated _Y_treated_`i'
    gen _Y_gap_`i' = _Y_treated_`i' - _Y_synthetic_`i'
    save "$input/pt/resout`i'.dta", replace
}

use "$input/pt/resout1.dta", clear
forvalues i = 2/27 {
    merge 1:1 _Co_Number _time using "$input/pt/resout`i'.dta", nogen
}

save "$input/pt.dta", replace

use "$input/df.dta", clear

tsset id year

```



```
synth homiciderates yearsschoolingimp stategdpcapita homiciderates proportionextremepo
> verty giniimp populationprojectionln stategdpgrowthpercent, trunit(20) trperiod(1999
> ) nested
```

```
matrix gaps=e(Y_treated) -e(Y_synthetic)
matrix Y_treated=e(Y_treated)
matrix Y_synthetic=e(Y_synthetic)
keep year
svmat gaps
svmat Y_treated
svmat Y_synthetic
gen _Co_Number=_n
gen _time=year
save "$input/pt/resout28", replace

use "$input/pt/resout1.dta", clear
forvalues i = 2/28 {
merge 1:1 _Co_Number _time using "$input/pt/resout`i'.dta", nogen
}
```

```
twoway (line _Y_gap_1 _time, lcolor(gray)) (line _Y_gap_2 _time, lcolor(gray)) (line _
> Y_gap_3 _time, lcolor(gray)) (line _Y_gap_4 _time, lcolor(gray)) (line _Y_gap_5 _tim
> e, lcolor(gray)) (line _Y_gap_6 _time, lcolor(gray)) (line _Y_gap_7 _time, lcolor(gr
> ay)) (line _Y_gap_8 _time, lcolor(gray)) (line _Y_gap_9 _time, lcolor(gray)) (line _
> Y_gap_10 _time, lcolor(gray)) (line _Y_gap_11 _time, lcolor(gray)) (line _Y_gap_12 _
> time, lcolor(gray)) (line _Y_gap_13 _time, lcolor(gray)) (line _Y_gap_14 _time, lcol
> or(gray)) (line _Y_gap_15 _time, lcolor(gray)) (line _Y_gap_16 _time, lcolor(gray))
> (line _Y_gap_17 _time, lcolor(gray)) (line _Y_gap_18 _time, lcolor(gray)) (line _Y_g
> ap_19 _time, lcolor(gray)) (line _Y_gap_21 _time, lcolor(gray)) (line _Y_gap_22 _tim
> e, lcolor(gray)) (line _Y_gap_23 _time, lcolor(gray)) (line _Y_gap_24 _time, lcolor(
> gray)) (line _Y_gap_25 _time, lcolor(gray)) (line _Y_gap_26 _time, lcolor(gray)) (li
> ne _Y_gap_27 _time, lcolor(gray)) (line gaps1 _time, lcolor(black) lwidth(thick)), x
> line(1999, lpattern(shortdash) lcolor(grey)) legend(order(27 "Sao Paulo" 2 "Control
> States")) xtitle("Year") ytitle("Gap in Homicide Rates") yline(0, lcolor(black))
```

```
graph export "$output/6.png", replace
```

```
* Gráfico 7
```

```
use "$input/df.dta", clear
tsset id year
```

```
keep if code==13 | code==15 |code==17|code==21|code==23|code==24|code==25|code==31|cod
> e==41|code==42|code==43|code==53|code==35
```

```
cd "$input/pt2"
```

```
egen id2=group(code)
```

```
tsset id2 year
```

```
tempname resmat
local i 9
qui synth homiciderates yearsschoolingimp stategdpcapita homiciderates proport
> ionextremepoverty giniimp populationprojectionln stategdpgrowthpercent, trunit(`i')
> trperiod(1999) keep(resout`i', replace)
matrix `resmat' = nullmat(`resmat') \ e(RMSPE)
local names ``names' ``i'""""
mat colnames `resmat' = "RMSPE"
mat rownames `resmat' = `names'
matlist `resmat' , row("Treated Unit")
```

```
drop if id2==9
```

```
forvalues i = 1/13 {
if `i'==9 {
continue
}
qui synth homiciderates yearsschoolingimp stategdpcapita homiciderates proport
> ionextremepoverty giniimp populationprojectionln stategdpgrowthpercent, trunit(`i')
> trperiod(1999) keep(resout`i', replace)
```

```

matrix `resmat' = nullmat(`resmat') \ e(RMSPE)
local names ``names' ``i'""'
}

mat colnames `resmat' = "RMSPE"
mat rownames `resmat' = `names'
matlist `resmat' , row("Treated Unit")

forvalues i = 1/13 {
use "$input/pt2/resout`i'.dta", clear
ren _Y_synthetic _Y_synthetic`i'
ren _Y_treated _Y_treated`i'
gen _Y_gap`i'=_Y_treated`i'-_Y_synthetic`i'
save "$input/pt2/resout`i'.dta", replace
}

use "$input/pt2/resout1.dta", clear
forvalues i = 2/13 {
merge 1:1 _Co_Number _time using "$input/pt2/resout`i'.dta", nogen
}

save "$input/pt2.dta", replace

use "$input/df.dta", clear

tsset id year

synth homiciderates yearsschoolingimp statgdpcapita homiciderates proportionextremepo
> verty giniimp populationprojectionln statgdpgrowthpercent, trunit(20) trperiod(1999
> ) nested

matrix gaps=e(Y_treated) -e(Y_synthetic)
matrix Y_treated=e(Y_treated)
matrix Y_synthetic=e(Y_synthetic)
keep year
svmat gaps
svmat Y_treated
svmat Y_synthetic
gen _Co_Number=_n
gen _time=year
save "$input/pt2/resout14", replace

use "$input/pt2/resout1.dta", clear
forvalues i = 2/14 {
merge 1:1 _Co_Number _time using "$input/pt2/resout`i'.dta", nogen
}

twoway (line _Y_gap_1 _time, lcolor(gray)) (line _Y_gap_2 _time, lcolor(gray)) (line _
> _Y_gap_3 _time, lcolor(gray)) (line _Y_gap_4 _time, lcolor(gray)) (line _Y_gap_5 _tim
> e, lcolor(gray)) (line _Y_gap_6 _time, lcolor(gray)) (line _Y_gap_7 _time, lcolor(gr
> ay)) (line _Y_gap_8 _time, lcolor(gray)) (line _Y_gap_10 _time, lcolor(gray)) (line
> _Y_gap_11 _time, lcolor(gray)) (line _Y_gap_12 _time, lcolor(gray)) (line gaps1 _tim
> e, lcolor(black) lwidth(thick)), xline(1999, lpattern(shortdash) lcolor(grey)) legen
> d(order(12 "Sao Paulo" 2 "Control States (MSPE)") xtitle("Year") ytitle("Gap in Hom
> icide Rates") yline(0, lcolor(black))

```