1 Introduction

Synthetic control is a method for estimating causal effects in comparative case studies when there exists only one treatment unit, such as a single state, country, school, etc. It was first developed by Alberto Abadie and Javier Gardeazabal in their 2003 AER (Abadie and Gardeazabal, 2003). Robust methods for inference were added in the 2010 JASA (Abadie, Diamond and Hainmueller, 2010). While the estimation of the causal effect is a canned procedure in STATA, the inference program is not canned and requires a little more effort.[1] The purpose of this document is to show the reader how to estimate causal effects using synth, including perform inference described in Abadie, Diamond and Hainmueller (2010). I use as an illustrating example a Texas natural experiment in which the state significantly expanded its prison capacity by the building of new prisons and expansion of prison capacity on old prisons in 1993 (Cornwell and Cunningham, 2016).

2 Texas experiment

Texas expanded its operational capacity in prisons starting in 1993. For three years – 1993, 1994, 1995 – operational capacity expanded 35% per year causing its overall operational capacity to approximately double in three years. This expansion is described in more detail in Cornwell and Cunningham (2016) and Perkinson (2010). We use the 1993-1995 expansion as a natural experiment to examine the effect of prison expansion on incarceration rates.

3 Files

There are two files you will need to download: `texas.do` and `texas.dta`.[2] Briefly, `texas.do` is a STATA do-file which runs the synthetic control program, including inference procedure, and `texas.dta` is the STATA data file saved in STATA 14 format. Before you can run this do-file, you will need to install the `synth` program.[3]

[1] Note, `synth` is also available as an `R` package. See Abadie, Diamond and Hainmueller (2011).
[2] I have also included a STATA 13 version entitled `texas13.dta`
[3] Type `ssc install synth` from the STATA prompt.

4 Variables

There are 16 variables in `texas.dta`. They are listed below.

Table 1 Variable names

variable name	type	format
statefip	float	%9.0g
year	float	%8.0g
bmprison	double	%10.0g
wmprison	double	%10.0g
bmpop	long	%12.0g
wmpop	long	%12.0g
alcohol	float	%9.0g
income	long	%12.0g
ur	float	%9.0g
poverty	float	%9.0g
black	float	%9.0g
perc1519	float	%9.0g
aidscapita	float	%9.0g
state	str20	%20s
bmprate	float	%9.0g
wmprate	float	%9.0g

5 Program directory

The do-file makes extensive use of sub-directories. Create a main directory with two sub-directories: /Do and /Data. Place `texas.do` in the /Do subdirectory and `texas.dta` in the /Data subdirectory. Create a new subdirectory in the /Data subdirectory entitled /Data/synth. All synthetic control estimates will be saved to /Data/synth. Create a subdirectory off your main directory entitled /Inference. All inference output as a text file and an excel spreadsheet will be saved to /Inference.[4] Create a figures directory off the main subdirectory entitled /Figures. All figures will be saved to this subdirectory. To summarize your directory structure should be as follows:

[4] see lines 116-127

- Do
- Data
  - synth
- Inference
- Figures

6 Output

You can now implement the synthetic control estimator, including inference, by running the `texas.do` file from your /Do subdirectory. The do-file will create several files that are then used to create figures. Synthetic control weights and balance in covariates can be recovered by running only lines 10-23 (which also creates the main synthetic control figure). The causal effect graph is created by running lines 28-38. Inference is conducted in lines 41-139. Figures of the causal effect including placebo lines is created by running lines 142-255.

Please contact with me errors or questions at the following address. And if you find ways to improve on this do-file, please share it with me as well.

Scott Cunningham
Associate Professor
Department of Economics
Baylor University
Waco, TX 76798-8003
http://www.scunning.com
scunning@gmail.com

References

Abadie, Alberto, Alexis Diamond and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105(490):493–505.

Abadie, Alberto, Alexis Diamond and Jens Hainmueller. 2011. "Comparative Politics and the Synthetic Control Method." Unpublished Manuscript.

Abadie, Alberto and Javier Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review* 93(1):113–132.

Cornwell, Christopher and Scott Cunningham. 2016. "Mass Incarcration's Effect on Risky Sex." Unpublished Manuscript.

Perkinson, Robert. 2010. *Texas Tough: The Rise of America's Prison Empire*. First ed. Picador.