Intro
oo

SCMs
oooooo

SCMs with SA
oooo

CM and SCMs
oooooo

Future Directions
ooo

Synthetic Control Methods for Causal Inference in Observational Social and Behavioral Sciences

John M. Felt, Ph.D.

Assistant Research Professor
Center for Healthy Aging
The Pennsylvania State University

Talk presented for
Modern Modeling Methods Conference
University of Connecticut
June 28th, 2023

Intro
00

SCMs
000000

SCMs with SA
0000

CM and SCMs
000000

Future Directions
000

Acknowledgements

- ▶ Collaborators
  - ▶ Chad Shenk
  - ▶ Kenneth Shores
  - ▶ Zack Fisher
  - ▶ Nilam Ram
  - ▶ Ulziimaa Chimed-Ochir
  - ▶ Anneke Olson

Intro
○○

SCMs
○○○○○○

SCMs with SA
○○○○

CM and SCMs
○○○○○○

Future Directions
○○○

Outline

Causal Inference in Observational Research

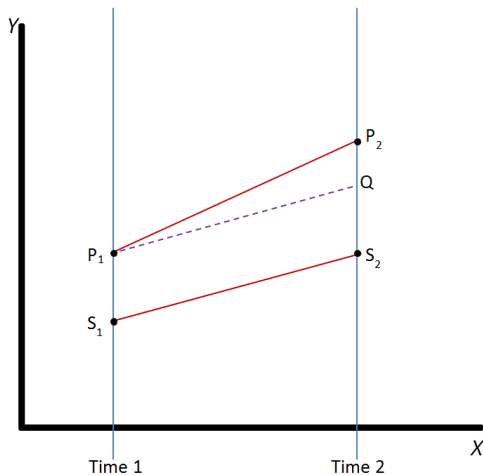Synthetic Control Methods (SCMs)

SCMs with Staggered Adoption

Child Maltreatment and SCMs

Future Directions

Causality

Intro
○●

SCMs
○○○○○○

SCMs with SA
○○○○

CM and SCMs
○○○○○○

Future Directions
○○○

Difference-in-Differences

Intro
○○
SCMs
●○○○○○
SCMs with SA
○○○○
CM and SCMs
○○○○○○
Future Directions
○○○

The Synthetic Control Method[1]

Example Data Used for Synthetic Control Method

—— Control Unit 1 —— Control Unit 2 —— Treated Unit

[1] Abadie et al. (2010)

Intro
○○

SCMs
○●○○○○○

SCMs with SA
○○○○

CM and SCMs
○○○○○○

Future Directions
○○○

The Synthetic Control Method

Intro
○○

SCMs
○○●○○○

SCMs with SA
○○○○

CM and SCMs
○○○○○○

Future Directions
○○○
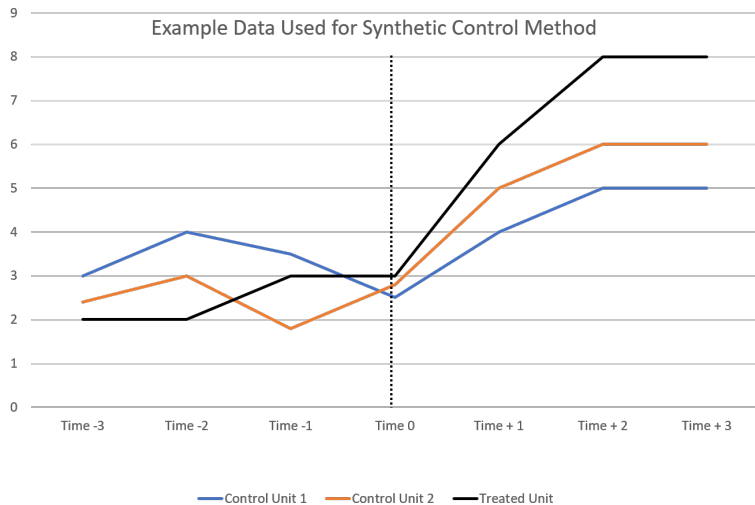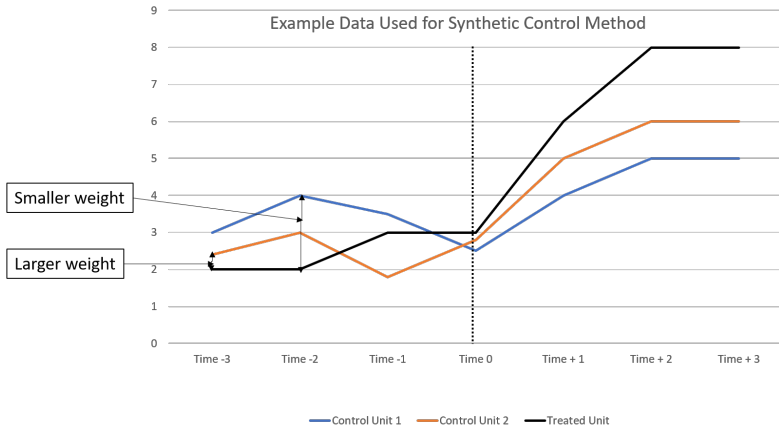
$$\min_{\mathbf{W}} \sum_{i=1}^{n} \frac{(Y_{it} - \sum_{j=1}^{p} W_j X_{jt})^2}{V_{it}} \quad \text{subject to} \quad \sum_{j=1}^{p} W_j = 1, \quad W_j \geq 0$$

n is the number of time periods

p is the number of control units

Y_{it} is the outcome variable of interest for the treated unit i at time t

X_{jt} is a vector of pre-treatment characteristics for control unit j at time t

V_{it} is a weight or variance term that represents the measurement error or uncertainty in the outcome variable Y_{it}

$\mathbf{W} = (W_1, W_2, \ldots, W_p)$ is a vector of weights assigned to the control units.

The Synthetic Control Method

Intro
○○

SCMs
○○○○○●○

SCMs with SA
○○○○

CM and SCMs
○○○○○○

Future Directions
○○○
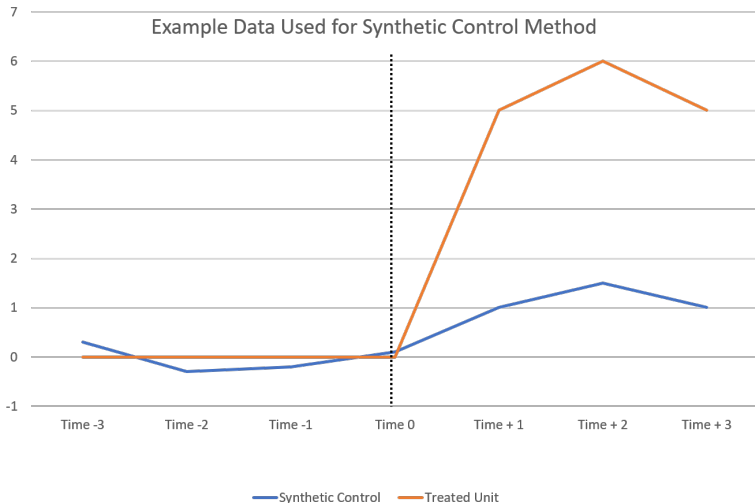
$$SATT = \frac{\sum_{t=t_1+1}^{T} w_{j,t}(Y_{j,t} - \bar{Y}t) - \sum_{t=t_1+1}^{T} w_{i,t}(Y_{i,t} - \bar{Y}t)}{\hat{\sigma}\epsilon},$$

$Y_{i,t}$ is the observed outcome for the treated unit i at time t $Y_{j,t}$ is the observed outcome for the control unit i at time t $w_{j,t}$ is the weight assigned to control unit j at time t

$\bar{Y}_t$ is the average outcome for all units at time t in the pre-treatment period

$\hat{\sigma}_\epsilon$ is the estimated residual variance, which is typically calculated based on the pre-treatment period.

Intro
oo

SCMs
ooooo●

SCMs with SA
oooo

CM and SCMs
oooooo

Future Directions
ooo

Data Requirements for SCM

ID	Time	Ever Treat	When Treat
1	1	1	0
1	2	1	0
1	3	1	1
1	4	1	1
2	1	0	0
2	2	0	0
2	3	0	0
2	4	0	0
3	1	1	0
3	2	1	1
3	3	1	1
3	4	1	1

Synthetic Controls with Multiple Treated Units[2]

- ▶ Also referred to as the Partially Pooled Synthetic Control Method

- ▶ Takes into account the imbalance for each treated united separately and the imbalance for the average of the treated units

- ▶ This partially pooled SCM generates weights to balance both of these sources of error

- ▶ Includes an intercept shift to account for average pre-exposure differences

---

[2]Ben-Michel, Feller, & Rothstein (Working Paper)

Intro
00

SCMs
000000

SCMs with SA
0●00

CM and SCMs
000000

Future Directions
000

Partially Pooled SCMs

Construct a synthetic control group for each unit using
pre-treatment data up to the time of adoption. Let $\mathbf{W}_{it}$ be the
weights assigned to each control unit for the synthetic control
group of treated unit i at time t.

$$\mathbf{X}it = \sum j = 1^N \mathbf{W}ijt\mathbf{X}jt$$

$\mathbf{X}_{it}$ is a vector of pre-treatment characteristics for treated unit i
at time t

$\mathbf{X}_{jt}$ is a vector of pre-treatment characteristics for control unit j
at time t

N is the number of control units

Intro
00

SCMs
000000

SCMs with SA
0000

CM and SCMs
000000

Future Directions
000

Partially Pooled SCMs

Estimate the treatment effect for each unit at the time of
adoption using the difference between the observed outcome
and the predicted outcome from the synthetic control group:

$$\text{TE}i = Yi, t_i - \mathbf{X}i, t_i{}^\top \mathbf{W}it_i$$

Y_{i,t_i} is the observed outcome for treated unit i at the time of
adoption t_i

X_{i,t_i} is the observed outcome for the control unit i at the time
of adoption t_i

$\mathbf{W}_{it_i}$ is the vector of weights assigned to each control unit for
the synthetic control group of treated unit i at time t_i

Intro
○○

SCMs
○○○○○○

SCMs with SA
○○○●

CM and SCMs
○○○○○○

Future Directions
○○○

Partially Pooled SCMs

Combine the estimates of treatment effects across units using a weighted average:

$$\text{ATT} = \frac{\sum_{i=1}^{M} \text{TE}i \cdot w_i}{\sum i = 1^M w_i}$$

TE_i is the estimated treatment effect for treated unit i
w_i is the weight assigned to treated unit i
M is the number of treated units.

Child Maltreatment and Externalizing Behavior Problems

- ▶ Child maltreatment is a worldwide health concern consisting of abuse (physical, emotional, and sexual) and neglect.

- ▶ More likely to develop externalizing behavior problems such as aggression[3]

- ▶ Control groups have typically been created through demographic matching with children who have never been maltreated

---

[3]Cicchetti & Valentino (2006)

Child Maltreatment and LONGSCAN

▶ Multisite multiyear study of child maltreatment (N = 1354)

▶ Substantiated maltreatment and externalizing behavior problems measured every 2 years from age 4 to 16

▶ Repeated assessments make SCMs an ideal method

Intro
○○

SCMs
○○○○○○

SCMs with SA
○○○○

CM and SCMs
○○●○○○

Future Directions
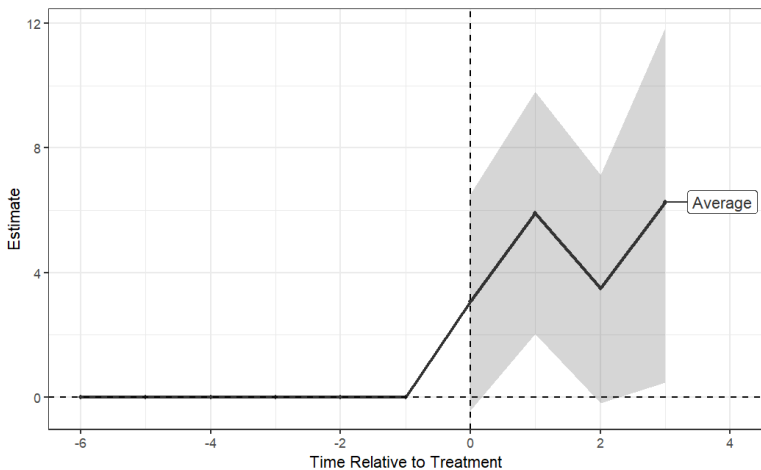○○○

Data Requirements for SCM

- ▶ For SCMs, including partially pooled for staggered adoption, requires repeated measures prior to maltreatment

- ▶ Data restricted to those with at least 3 pre-maltreatment times (N = 63)

- ▶ N = 571 children in control group

Results

TimeSince	ATT	SE	CIL	CIU
0	3.06	1.74	-0.35	6.35
2	**5.91**	**1.87**	**2.65**	**9.80**
4	3.49	1.80	-0.03	7.03
6	**6.27**	**2.93**	**0.36**	**12.00**

Average ATT (of post intervention times) = 4.652 (SE = 1.70)

Results

Intro
oo

SCMs
oooooo

SCMs with SA
oooo

CM and SCMs
oooooo●

Future Directions
ooo

Weights

	V17	V18	V19	V20	V21	V22	V23
EA40015	-3.909506e-08	0.0012600979	0.0074696094	1.790062e-02	3.799893e-05	0.0004148545	0.0001107647
EA40056	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40091	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40129	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40157	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40185	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	7.232834e-04	0.0000000000	0.0000000000
EA40190	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40247	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40251	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40304	-9.826916e-08	0.0028835375	0.0013200143	6.203384e-04	3.466743e-04	0.0006975510	0.0010110625
EA40312	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40353	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40399	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40401	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40414	1.790066e-01	0.0000000000	0.0000000000	3.091038e-04	0.000000e+00	0.0000000000	0.0000000000
EA40436	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40450	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40469	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40475	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000
EA40497	0.000000e+00	0.0000000000	0.0000000000	0.000000e+00	0.000000e+00	0.0000000000	0.0000000000

Intro
oo

SCMs
oooooo

SCMs with SA
oooo

CM and SCMs
oooooo

Future Directions
●oo

Future Considerations

► Methodological

▶ Number of times pre-
and post- exposure

▶ Strength of average
effect

▶ Heterogeneity in effect
across persons

▶ Number of control units

► Child Maltreatment

▶ More Assessments
Earlier in Development

▶ Handling
"Contamination"

Intro
oo

SCMs
oooooo

SCMs with SA
oooo

CM and SCMs
oooooo

Future Directions
o●o

References

▶ Abadie et al. (2010): Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of American Statistical Association*

▶ Ben-Michael, Feller, Rothstein (Working Paper). Synthetic Controls with Staggered Adoption.

▶ Ben-Michael, Feller, Rothstein (2021). The Augmented Synthetic Control Method. *Journal of the American Statistical Association*

▶ Kim, Lee, & Gupta (2020). Bayesian Synthetic Control Methods. *Journal of Marketing Research*

▶ Shadish, Cook, & Campbell (2002). Experimental and Quasi Experimental Designs for Generalized Causal Inference