Published in final edited form as:

Epidemiology. 2018 July ; 29(4): 503–505. doi:10.1097/EDE.0000000000000837.

A new tool for case studies in Epidemiology - the Synthetic Control Method

David H. Rehkopf and Sanjay Basu

Primary Care and Population Health, Stanford University School of Medicine, Stanford, CA (DHR); Primary Care and Outcomes Research, Stanford University School of Medicine, Stanford, CA (SB)

With the publication of “Repeal of Comprehensive Background Check Policies and Firearm Homicide and Suicide,”¹ the synthetic control method joins a group of useful statistical methods for the analysis of observational data imported to the field of epidemiology from economics and political science. These approaches include fixed effects (2), difference in differences (2), regression discontinuity (3) and instrumental variables (4). This expansion of methodological tools has been critical for the development for epidemiology as the discipline has expanded to consider the analysis of exposures that are not easily subjected to true random assignment, and for which unmeasured confounders plausibly exist. Furthermore, for social epidemiology, these methods help to facilitate a consequential approach (5,6), as the subfield moves to evaluating the impacts of social and economic policies on health. The importance of considering the synthetic control method for epidemiologic questions is that when randomization is difficult or impossible, and there are no available instrumental variables, the tools available for analysis are currently limited. This is particularly true when only one or a small number of units are treated.

The synthetic control method is most closely related to a difference in differences approach. The difference in differences comparison is based on defining a treated unit (e.g., the population of a state that is exposed to a policy) and a control unit (e.g., a single state or multiple states), and comparing pre/post change between these units. The novel contribution of this method is that, rather than aggregating all comparison groups into one control unit, potential control groups (e.g., the ‘donor pool’ of unaffected states) are considered separately, and a weighted combination of states that are the best pre-treatment match with the treatment unit are used as the controls. Inference is then based on a post-treatment difference between the treatment and the weighted combination of control states. Thus, while difference in differences requires a parallel trend assumption for the pre/post comparison, the synthetic control method relies on there being a close pre-treatment match with the donor pool. The “synthetic” term refers to the fact that the control group is a weighted average of the available control states, such that typically some states (because they diverge too much from the treatment state in their pre-treatment characteristics) are not

Corresponding author: David H. Rehkopf, drehkopf@stanford.edu.

Invited Commentary on: Repeal of Comprehensive Background Check Policies and Firearm Homicide and Suicide” by Rose M.C. Kagawa, Alvaro Castillo-Carniglia, Jon S. Vernick, Daniel Webster, Cassandra Crifasi, Kara E. Rudolph, Magdalena Cerdá, Aaron Shev, Garen J. Wintemute

used at all, and other states contribute different amounts to the synthetic control population. For example, as shown in table 2 of Kagawa et al., when analyzing firearm homicide in Indiana, the synthetic control is a combination of Illinois (weighted as 56% of the synthetic control group) and Iowa (the remaining proportion).

Epidemiologists, perhaps more than most disciplines given the field's development of case-control methods, are acutely aware of the importance of appropriately selecting controls, and this is a critical decision to be made with the synthetic control method. The user selects the pool of eligible control regions. Kagawa et al. appropriately select only eligible states, and because the treatment is repeal of a policy, they are limited to only the eight other states that currently have the policy. This approach is most consistent with the selection of the control population in a case-control study with respect to the outcome of interest, selecting randomly from among those states at risk for the outcome (7). Recent synthetic control method applications used as few as 16 (8) but as many as 130 (9) regions in the donor pool. In addition, based on the matching shown in table 2 of Kagawa et al., half of the eight states are not very good matches, with virtually all of the control states across outcomes made up of Illinois, Iowa, Michigan, and North Carolina. The primary advantage in the synthetic control is taking a data-driven approach to getting a close pre-treatment match between treatment and control regions, but this advantage is more challenging to leverage when the sample of donor pool states is small. In scenarios where only a small number of regions match, robustness checks with the synthetic control method have been performed by excluding the best matches and showing that results were consistent with the second best selection of control regions (10).

The importance of the donor pool is magnified with the synthetic control method because inference about whether the results are due to random error also depends on the choice of the donor pool. This is perhaps the most novel and challenging aspect of the synthetic control method; there is not a validated, parametric approach to establishing confidence intervals on the matched post-treatment difference between the treatment and synthetic control as there is with difference in differences. Kagawa *et al.* take the generally recommended approach of a placebo test where each of the control states is used to establish the range of the post-treatment level of the outcome that could be expected in the absence of treatment, since the control states did not receive the intervention. While there is a rough analogy here to simulated confidence intervals from approaches such as the bootstrap, the additional challenge for synthetic control method is that the standard assumptions for estimating an effect parameter are also required assumptions for valid inference on whether the effect parameter estimated is within the range of natural variation observed in the placebo controls. That is, the following assumptions must be met for valid inference about the *precision* of the estimates: (i) stable unit treatment value assumption, (ii) assumption of random assignment to the treatment, and (iii) assumption that the potential outcomes for regions are fixed but *a priori* unknown (11). Imagine, for example, that New York and Massachusetts are regions that differ substantially from Indiana and Tennessee (which is consistent with the fact that they don't serve as best matched control regions). These regions are then used to establish the natural variation that would be expected for Indiana and Tennessee if no intervention had occurred. The conclusion that there is no impact of the repeal of comprehensive background checks is dependent on the assumption that Indiana

and Tennessee are randomly selected for treatment from the sample of control states. This is a particular challenge when study conclusions are that there was no impact of a policy change. We are thus concerned about the potential power for Kagawa et al. to detect an effect, particularly on a rare outcome, using the synthetic control method approach. This limitation may be able to be addressed in the future by modifications to the method that produce confidence intervals (11–13). A further recommended approach is to assess placebo treatments based on time (8).

In addition to selecting the potential control units, another set of critical decisions involves what variables should be used to match units, and an evaluation of how close the synthetic match is. Since pre-treatment levels of the outcome are almost always the strongest predictor of post-treatment levels of the outcome, critical decisions revolve around how to include these in the model. Kagawa et al. fit several alternatives with respect to pre-treatment outcome levels, deciding on using a single year measure of the outcome directly prior to treatment based on matching fit, an important consideration to avoid false positive findings (14). However, the difficulty of finding a good match across all covariates is in part driven by the fact that there are very few matching units available, so in several cases the best matched synthetic controls are ones that differ substantially in their pre-treatment levels of the outcome. For example, the rate of firearm suicide pre-treatment for the synthetic control group for Tennessee is 54% lower than the pre-treatment level in actual Tennessee (shown in eTable 2). This creates a situation where firearm suicide is likely to appear higher in Tennessee after the treatment. The difficulty in pre-treatment matching is made more difficult by the instability in the outcome and the sharply declining trend in the firearm homicide rate just prior to treatment (Kagawa Figure 1A). Prior applications of the synthetic control method have shown less volatility in the outcome and a generally higher pre-treatment match (8–10), for example a 1% difference in the primary models and a range of 0%–3% differences in alternative specifications (10).

Abadie et al., who developed the method, present the synthetic control method as a bridge between qualitative and quantitative research in that it provides more rigor for the comparison groups chosen for comparative case studies (8). This assessment emphasizes that when there are only one or two treated units and a small number of control units, the power for inference will always be limited. Without robust confidence intervals and a clear understanding of statistical power, the estimates from the method should thus be considered similar to estimates from a case study when considering the strength of evidence from this approach. We do not use the term “case study” pejoratively here, but rather to frame how we consider the contribution that estimates from synthetic control method contribute to causal questions. The synthetic control method provides a useful tool for epidemiology when there are only a few treated units. It is a tool that, as demonstrated by Kagawa et al., increases the transparency of how closely a control unit matches pre-treatment, a worthwhile advance over analyses using single unit or aggregated unit control groups.

Acknowledgments

Funding: NIH AG047280 (DHR); NIH DP2MD010478 (SB)

Biographies

David H. Rehkopf is an Assistant Professor of Medicine at the Stanford University School of Medicine. His research interests are in understanding the role of social and economic policies in chronic disease risk and in the biological pathways through which those processes occur.

Sanjay Basu is an Assistant Professor of Medicine at the Stanford University School of Medicine. His research interests are in improving strategies for the primary and secondary prevention of cardiovascular disease.

References

1. Kagawa RMC, Castillo-Carniglia A, Vernick JS, Webster D, Crifasi C, Rudolph KE, Cerdá M, Shev A, Wintemute GJ. Repeal of Comprehensive Background Check Policies and Firearm Homicide and Suicide. *Epidemiology*. 2018; 29:xxx.
2. Strumpf, EC., Harper, S., Kaufman, JS. *Methods in Social Epidemiology*. Jossey-Bass; San Francisco CA: 2017. Fixed effects and difference in differences.
3. Bor J, Moscoe E, Mutevedzi P, Newell M-L, Bärnighausen T. Regression discontinuity designs in epidemiology: causal inference without randomized trials. *Epidemiology (Cambridge, Mass)*. 2014; 25(5):729.
4. Greenland S. An introduction to instrumental variables for epidemiologists. *International journal of epidemiology*. 2000; 29(4):722–9. [PubMed: 10922351]
5. Galea S. An argument for a consequentialist epidemiology. *American journal of epidemiology*. 2013; kwt172.
6. Nandi A, Harper S. How consequential is social epidemiology? A review of recent evidence. *Current epidemiology reports*. 2014; 2(1):61–70.
7. Robins JM, Gail MH, Lubin JH. More on "Biased selection of controls for case-control analyses of cohort studies". *Biometrics*. 1986:293–9. [PubMed: 3741971]
8. Abadie A, Diamond A, Hainmueller J. Comparative politics and the synthetic control method. *American Journal of Political Science*. 2015; 59(2):495–510.
9. Gobillon L, Magnac T. Regional policy evaluation: Interactive fixed effects and synthetic controls. *Rev Econ Stat*. 2016; 98(3):535–51.
10. Pinotti P. The economic costs of organised crime: Evidence from Southern Italy. *The Economic Journal*. 2015; 125(586)
11. Firpo S, Possebom V. Synthetic control method: Inference, sensitivity analysis and confidence sets. 2017
12. Xu Y. Generalized synthetic control method: Causal inference with interactive fixed effects models. *Political Analysis*. 2017; 25(1):57–76.
13. Athey S, Bayati M, Doudchenko N, Imbens G, Khosravi K. Matrix completion methods for causal panel data models. *arXiv preprint arXiv:171010251*. 2017
14. Ferman B, Pinto C, Possebom V. Cherry picking with synthetic controls. 2017