Appendix 8 (as supplied by the authors): Synthetic control methods

The synthetic control method estimates the effect of an event or treatment in a country by comparing the outcome of interest with its counterfactual level in the absence of the treatment.[1,2] The method differs from traditional approaches to comparative causal inference by making comparisons with an estimated counterfactual in the treated country rather than comparing it with a different un-treated country or unit.

The counterfactual 'synthetic control' is estimated using the values of predictors of the outcome – in our case caloric sweeteners – in un-treated comparison countries. The countries that form the synthetic control are selected from a pool of countries based on their similarity to the treated country before the treatment with respect to relevant predictors of the outcome, including treatment covariates and past realizations of the outcome variable. The matching algorithm selects comparison units by assigning a weight to each potential comparison country, and selecting the combination of units and weights that minimises the distance between the value of outcome predictors in the combined synthetic control unit and their observed values in the treated country before the intervention. To estimate the value of the outcome in the synthetic control these predictors are also assigned a second set of weights according to their contribution to predictive power on the outcome. Thus the algorithm that choses the comparison countries and their weights W* that minimises:

Equation 1. $\sum_{m=1}^{k} v_m(X_{1m} - X_{0m}W)^2$

Where m is the country (1 through k), X_i is the value of outcome predictor X in country m with treatment status i (1 = treated, 0 = untreated), W is a vector of weights assigned to the predictors in country m, and v_m is a vector of weights assigned predictors based on their predictive power. The estimated effect size is given by comparison of post-treatment outcomes between the treated unit, Y_1, and the synthetic control of the outcome, Y_0w^*. This can be estimated in a given year or averaged

across multiple years by calculating the Root Mean Squared Prediction Error (RMSPE, see footnote 1[1]).

The synthetic control method has multiple advantages for evaluating the effects of FTAs compared with traditional cross-country analyses and natural experiment designs. First, it offers several strengths of qualitative case studies that pursue a detailed understanding of the policy and context in question whilst also offering precise statistical inference, comparable numerical results, and selection of comparison units in a way that as closely as possible reproduces the counterfactual of the country of interest.[2,3] This is particularly important when evaluating FTAs. Their complex and varied nature – evidenced by the subtle tariff changes within NAFTA – calls for a granular approach, whilst potential unobserved confounding and pre-treatment selection bias call for more rigorous econometric analyses of their effects.[4,5]

Second, the synthetic control method for estimating the counterfactual outcome of the treated country is transparent to readers but precludes undesirable 'cherry picking' by the researcher due to the use of a matching algorithm.[6] Third, the algorithm-based matching method is advantageous compared with model-dependent extrapolations that regression based inferences are often based on, including fixed-effects 'difference-in-differences' (DiD) strategies in which the choice of specification for the DiD estimation can have a major impact on the point estimates and estimated statistical significance of estimated policy effects.[7,8]

Like all methods, these advantages also come with limitations. One is that standard techniques are not suitable for assessing the significance of the results. Following suggestions by Abadie and colleagues we therefore use placebo tests to compute p-values.[2] This approach assigns a dummy treatment to all units that are un-treated in the sample. P-values are then calculated as the proportion of effect sizes that are as large as the one obtained for the treated unit. This inferential exercise reduces to

[1] The RMSPE measures the lack of fit between the path of the outcome in a country and its synthetic counterpart across a time period T. For example, the RMSPE for the pre-intervention time-period with number T_0 years, comparison countries J yielding synthetic control estimates of Y, $\sum_{j=2}^{J+1} w_j^* Y_{it}$, is given by:

$$RMSPE = \left(\frac{1}{T_0}\sum_{t=1}^{T_0}\left(Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{it}\right)^2\right)^{1/2}$$

classical randomization inference with no rules-of-thumb for evaluating whether a given p-value yields significance or not. Instead, the p-value has an interpretation as the probability of obtaining an estimate at least as large as the one obtained for the unit representing the case of interest when the intervention is reassigned at random within the data set.[9]

References

1. Abadie A, Gardeazabal J. The economic costs of conflict: A case study of the Basque country. *Am Econ Rev*. 2003;93(1):113-132. doi:10.1257/000282803321455188.
2. Abadie A, Diamond A, Hainmueller J. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *J Am Stat Assoc*. 2010;105(490):493-505. doi:10.1198/jasa.2009.ap08746.http.
3. Abadie A, Diamond A, Hainmueller J. Comparative Politics and the Synthetic Control Method. *Am J Pol Sci*. 2015;59(2):495-510. doi:10.1111/ajps.12116.
4. Baier SL, Bergstrand JH. Do free trade agreements actually increase members' international trade? *J Int Econ*. 2007;71(1):72-95. doi:10.1016/j.jinteco.2006.02.005.
5. Baier SL, Bergstrand JH. Estimating the effects of free trade agreements on international trade flows using matching econometrics. *J Int Econ*. 2009;77(1):63-76. doi:10.1016/j.jinteco.2008.09.006.
6. King G, Keohane RO, Verba S. *Designing Social Inquiry: Scientific Inference in Qualitative Research*.; 2001. doi:10.2307/2076556.
7. King G, Zeng L. The dangers of extreme counterfactuals. *Polit Anal*. 2006;14(2):131-159. doi:10.1093/pan/mpj004.
8. Ryan AM, Burgess JF, Dimick JB. Why we should not be indifferent to specification choices for difference-in-differences. *Health Serv Res*. 2015;50(4):1211-1235. doi:10.1111/1475-6773.12270.
9. Ho DE, Imai K. Randomization Inference With Natural Experiments. *J Am Stat Assoc*. 2006;101(475):888-900. doi:10.1198/016214505000001258.