See comments at end.

Using State Space Models to Check for Time Varying Weights in Synthetic Control

Danny Klinenberg

Last Updated: 2020-03-15

Abstract

Synthetic Controls are becoming more and more used in policy analysis. One assumption of synthetic control is constant weights between the donor pool and the treated observation. This means that there are no structural breaks or shocks among any observation except the treatment of interest. I plan to use basic state space modeling as a check for changing weights. The proposed method is to compute a likelihood ratio test between a constant coefficient state space model and a random coefficient state space model using a donor pool previously chosen via synthetic control.

You might also want to simulate a zero treatment effect but changing weights

Don't single space body

1 Brief Overview

Synthetic Controls have become a staple in econometric analysis. They allow for inference on a small group of treated observations when there is treatment endogeneity. The intuitive idea behind synthetic control is to construct a counterfactual for the treated observations using a weighted average of the untreated. If the constructed, or synthetic, observation fits the treated observation "well before" the intervention, then the synthetic observation is said to have worked and qualifies as a counterfactual. "Fits the treated observation" means the synthetic control is almost perfectly matching in the pretreatment. "Well before" has been defined differently by different authors, but a general consensus is at least ten periods. A causal interpretation can then be made on the effect of the treatment in following periods by comparing the synthetic observation to the actual.

You should distinguish between nonconstant before and non-constant after policy switch

Causal inference is jeopardized when the weights used for the counterfactual are non-constant. The original method assumes that the weights are unchanging throughout the research period. Papers limit their potential controls, or donor pool, to observations that do not have major shocks or disturbances. However, this requires the researcher to have an in depth knowledge of possible societal, economic, and political changes of each observation as well as the insight to know their effects. In addtion, gradual changes in the relationship between observations may be overlooked or assumed away. Synthetic control deals mainly with aggregate level data. This includes country level GDP (e.g., (Campos, Coricelli, and Moretti 2019a)) or state level issues (Abadie, Diamond, and Hainmueller 2010). Aggregate level data paired with pre-treatment intervals spanning at least ten periods add doubt to the constant weights assumption.

You don't need to attack the plausibility of constant weights.

I propose using state space models as a check for constant weights after running a synthetic control analysis.

2 ADH Synthetic Control

not just for macro

In many macroeconomic events, a true counterfactual does not exist. In addition, there are relatively few treated observation. This means a researcher cannot use traditional tools such as difference in differences, regression discontinuity, or simple randomization. Synthetic Control was proposed as an alternative in which observations unaffected by a policy are pooled together to create a counterfactual for the treated observations. This was first employed in (Abadie and Gardeazabal 2003) and formalized in a followup paper (Abadie, Diamond, and Hainmueller 2010). The basic setup of this method (now referred to as *ADH synthetic control*) is as follows:

Suppose there is an outcome of interest, Y_{it}. Suppose there are N+1 observations observed over T years with one observation being exposed to treatment. Without loss of generality, assume observation $i = 1$ is treated starting at time T_0. Furthermore, define the treatment D such that $D_{1t} = 1 * I(t \geq T_0)$ with treatment effect $\alpha_{1,t}$. Using the potential outcomes framework, assume there are two states of the world: untreated $Y_{it}(0)$ and treated $Y_{it}(1) = \alpha_{it} D_{it} + Y_{it}(0)$. The objective is to measure:

$$\alpha_{1t} = Y_{1t}(1) - Y_{1t}(0)$$

Furthermore, assume that $Y_{it}(0)$ comes from the following data generating process:

$$Y_{it}(0) = \delta_t + \theta_t Z_i + \lambda_t \mu_i + \epsilon_{it}$$

where δ_t is an unknown common time factor constant across units, θ_t is a (1 x r) of unknown parameters, Z_t is a (r x 1) vector of observed covariates that are not affected by treatment, λ_t is a (1 x F) vector of unobserved common factors, μ_i is a (F x 1) vector of unobserved factor loadings, and ϵ_{it} is the error term.

Unfortunately, $Y_{1t}(1)$ and $Y_{1t}(0)$ are never observable at the same time. Therefore, a counterfactual must be created. This is done through a convex set of the control observations over the pre-treatment period. The control observations are referred to as the *donor pool*. Formally, the donor pool is defined as $Y_{jt}(0) \forall j \neq 1$. If a study involved the policy implications on a country's GDP, the control observations would be other countries' GDP.

why this restriction

Suppose there exists a weighting matrix $W = (w_2, ..., w_{n+t})'$ such that $w_j \in [0, 1]$ and $\sum_{j=2}^{n+1} w_j = 1$. Choosing an optimal weighting matrix, w^* would create an unbiased estimator. Namely:

ADH Assumption: The estimator of $Y_{i1}(0)$ is unbiased if :

$$\sum_{j=2}^{n+1} w_j^* Z_j = Z_1$$

$$\sum_{j=2}^{n+1} w_j^* \mu_j = \mu_1$$

Using w^* would yield:

$$Y_{1,1}(0) = \sum_{i=2}^{n+1} w_i^* Y_{i,1}(0)$$

$$Y_{1,2}(0) = \sum_{i=2}^{n+1} w_i^* Y_{i,2}(0)$$

$$\vdots$$

$$Y_{1,T_0}(0) = \sum_{i=2}^{n+1} w_i^* Y_{i,T_0}(0)$$

The equality above implies that the weighted average of the other observations creates a perfect counterfactual for the treated unit. It is impossible to test if the μ condition is satisfied because μ is never observed.

However, (Abadie, Diamond, and Hainmueller 2010) argues this holds under fairly regular conditions. Assuming the relationship holds past T_0, a causal effect can be calculated.

On an intuitive level, the notion of constant weights means the relationship between the treated unit and controls is constant throughout the analysis period. If a synthetic control analysis is being employed to study country GDP (as is quite common), each country cannot have a significant change throughout the analysis. This means no major governmental shifts, policy reforms, or economic booms/busts. Given that synthetic control recommends a long pre-treatment window of ten periods and post of five, countries cannot undergo major changes in 15 periods. Since most country level data is reported at the yearly level, this translates to a 15 year constant relationship between the treatment and control.

At the writing of this proposal, the researcher is unaware of any other reports focusing specifically on the assumption of constant weights. Understanding if the constant weights assumption holds is essential to applying synthetic controls. As of now, the method is done via expert knowledge. Researchers filter out observations in which major events took place to disrupt the relationship. If a researcher is performing a synthetic control using country GDP as an outcome, then he/she would omit countries that had coups, trade wars or other economic shocks during the period of interest. In addition, comparing pre-treatment covariates to synthetic covariates are used. If there is a "close" fit, then the assumption is said to hold. (Abadie, Diamond, and Hainmueller 2010) does this via visual inspection of the outcome variable and inspection of the covariates. The proposed research will add to the synthetic control literature by adding a formalized check for constant weights.

3 Checking for Time Varying Weights Using State Space Modeling

The problem described above is a latent variable estimation problem. State space modeling is a time series concept that allows for modeling latent variables. This means modeling unobserved components like time trends, seasonality, and time varying coefficients. Thinking back to synthetic control, this would be the elements of μ_i. A state space model is composed of an observation equation and transition equation. A general form of these equations are:

$$\begin{aligned} y_t &= Z_t'\alpha_t + \epsilon_t \quad \text{observation equation} \\ \alpha_{t+1} &= T_t\alpha_t + \eta_t \quad \text{transition equation} \end{aligned}$$

y_t is the observed data and α_t is a combination of observed data (e.g. control variables) and unobserved components (e.g. trend and cycle). The transition equation can be matricized to incorporate multiple transition equations. This will be used below. State space models are structural models. The assumptions necessary for state space models are:

1) the observation equation and transition equation are linear. not for all state-space, but is true to apply Kalman Filter

2) $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ and $\eta_t \sim N(0, \sigma_\eta^2)$. The errors are also assumed to be ~~mutually and serially uncorrelated.~~ not absolutely
This is because they are meant to be random disturbances within the model. If α_{t+1} is matricized, then $\eta_t \sim N(\mathbf{0}, \mathbf{\Sigma})$ with $\mathbf{0}$ a vector of zeroes and Σ a diagonal matrix.

3) The errors must be normal.

4) the transition equations can be of lag order 1. Any additional lag orders can be rewritten as order 1 using the state space framework.

A benefit of state space models is the ability to explicitly model latent variables such as trend and seasonality. This comes from the model's parametric assumptions and utilizing the Kalman filter and smoother.

(Scott and Varian, n.d.) first proposed using state space models in combination with Spike and Slab variable selection for high dimensional time series analysis. This was then extended by (Brodersen et al. 2015) as an alternative for synthetic control. This paper opts to follow the state space model used by these papers (excluding the spike and slab setup):

$$y_{1t} = \xi_t + \sum_{j=2}^{n+1} w_j y_{j,t} + \epsilon_t \tag{1}$$

$$\xi_{t+1} = \xi_t + v_t + \eta_{1t} \tag{2}$$

$$v_{t+1} = v_t + \eta_{2t} \tag{3}$$

This is a basic structural model with a time varying intercept. The observable data are the y values. ξ_t is some trend term assumed to be linear with stochastic level and slope. ξ, v, η, and ϵ are not observed. If the researcher believes that the weights are nonconstant, the model can be rewritten as:

$$y_{1t} = \xi_t + \sum_{j=2}^{n+1} w_{j,t} y_{j,t} + \epsilon_t \tag{4}$$

$$\xi_{t+1} = \xi_t + v_t + \eta_{1t} \tag{5}$$

$$v_{t+1} = v_t + \eta_{2t} \tag{6}$$

$$w_{j,t+1} = \omega_{j,t} + \eta_{w_j,t} \quad \forall j \tag{7}$$

In this example, the weights are being estimated as a random walk. Notice if $\sigma^2_{\eta_{w_j}} = 0$ for all j, then the model simplifies back to equations (1)-(3). Estimating whether the weights are shifting simply becomes a likelihood ratio test where the null is $\sigma^2_{\eta_{w_j}} = 0$ for all j. The purpose of the test is to determine if the weights are changing at all during the pre-treatment estimation period. Therefore, it is sufficient to test if the standard deviation is nonzero. A downside to modeling the evolution of the weights in this fashion is that there is no way to know how the weights are evolving. A benefit is there are less parameters to estimate.

huh?

w ~~wtr~~ smoothed estimates

Equations (4)-(7) can be rewritten in matrix notation using the transition and observation equations described at the beginning of the section:

$$\alpha_t = \begin{bmatrix} \xi_t \\ v_t \\ w_{2,t} \\ \cdot \\ \cdot \\ \cdot \\ w_{n+1,t} \end{bmatrix}_{(n+2)\ x\ 1}$$

$$T = \begin{bmatrix} 1 & 1 & & \\ & 1 & 0 & \\ & & \ddots & \\ & 0 & & 1 \end{bmatrix}_{(n+2)\ x\ (n+2)}$$

$$Z' = \begin{bmatrix} 1 & 0 & y_{1,t} & \dots & y_{n,t} \end{bmatrix}_{1\ x\ (n+2)}$$

$$\eta_t = \begin{bmatrix} \eta_{1,t} & & & & \\ & \eta_{2,t} & & 0 & \\ & & \eta_{w_2,t} & & \\ & & & \ddots & \\ & 0 & & & \eta_{w_{n+1},t} \end{bmatrix}_{(n+2)\ x\ (n+2)}$$

$$y_t = y_t$$

$$\epsilon_t = \epsilon_t$$

The model can then be estimated with a Kalman filter and a likelihood ratio test can be performed[1]. The hyper parameters of the model are: σ_ϵ^2, $\sigma_{\eta_1}^2$, $\sigma_{\eta_2}^2$, and $\sigma_{w_j}^2$ for all j. The likelihood ratio test follows a χ^2 distribution with the number of restrictions as degrees of freedom. In this application, the number of restrictions will be the number of nonzero weights from the synthetic control.

4 Estimation

This paper will focus on a single treated time series with preselected control time series. In other words, I will begin this analysis with control time series already chosen by a synthetic control process. In addition, the paper is restricted to the outcomes of the observed units, without considering underlying covariates. A growing body of literature has supported synthetic control analysis without covariates. (Athey and Imbens 2017) and (Doudchenko and Imbens 2016) argue the outcomes tend to be far more important than covaritates in terms of predictive power. They further argue that minimizing the difference between treated outcomes and control outcomes prior to treatment tend to be sufficient to construct a synthetic control. (Kaul et al., n.d.) also shows that covariates become redundant when all lagged outcomes are included in ADH approach. (Botosaru and Ferman 2019) show that the counterfactual estimated by using only pre-treatment outcomes is very close to the original ADH that directly attempted to match on these covariates. In addition, (Brodersen et al. 2015) does not use covariates in their structural approach.

The purpose of matching on observed covariates and unobserved factors is to ensure an unbiased estimate of the outcome variable. Testing the relationship of the outcome variable is a check on the methodology's effectiveness. In other words, the purpose of this test is to ensure the constant weights assumption holds for the outcome variable. Therefore, this test will focus exclusively on the outcome variable..

The estimation process will be as follows:

1) select a past synthetic control analysis.

2) limit the control time series to those with non-zero weighting.

3) Perform a likelihood ratio test comparing the state space model with time-varying coefficients to the state space model without time-varying coefficients. Test the hypothesis:

$$\begin{aligned} H_0 : \sigma_{\eta_{w_j}}^2 &= 0 \qquad \forall j \\ H_1 : \sigma_{\eta_{w_j}}^2 &\neq 0 \quad \text{for some j} \end{aligned}$$

This will follow a $\chi_{(j)}^2$ distribution.

4) Discuss thoroughly.

Initial analysis would begin with a simulation study. I'd first determine if the level of fluctuation necessary for the state space model to detect differences. From there, the proposed paper will revisit (Billmeier and Nannicini 2013). This paper suggested a positive relationship between economic liberalisation and welfare. The study covered 30 individual cases from 1963 to 2000. The study performed a synthetic control approach on trade liberalization. Trade liberalization was defined as when a country became economically open. A country was considered economically closed if:

i) average tariffs exceeded 40%

[1]This section is a bit thin still because I am having trouble understanding how to set up the Kalman filter with unknown starting values. Will be more flushed out when I get through chapter 5 of Koopman and Durbin.

ii) nontariff barriers cover more than 40% of imports

iii) socialistic economic system

iv) black market premium on the exchange rate exceeded 20%

v) many of the exports are controlled by a state monopoly

This is an appealing dataset to use because there are 30 individual case studies to use and the data is publically available.

5 Conclusion

This proposal lays out a check for the synthetic control constant weights assumption. The check will be performed by running a basic state space model to test if the weights are random coefficients. Initially, the test will be performed on simulated data. This will be done to understand the effectiveness of the test. Then the test will be applied to (Billmeier and Nannicini 2013), a popular synthetic control paper.

Work Cited and 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. https://doi.org/10.1198/jasa.2009.ap08746.

Abadie, Alberto, and Javier Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review* 93 (1): 113–32. https://doi.org/10.1257/000282803321455188.

Athey, Susan, and Guido W. Imbens. 2017. "The State of Applied Econometrics: Causality and Policy Evaluation." *Journal of Economic Perspectives* 31 (2): 3–32. https://doi.org/10.1257/jep.31.2.3.

Billmeier, Andreas, and Tommaso Nannicini. 2013. "Assessing Economic Liberalization Episodes: A Synthetic Control Approach." *Review of Economics and Statistics* 95 (3): 983–1001. https://doi.org/10.1162/REST_a_00324.

Botosaru, Irene, and Bruno Ferman. 2019. "On the Role of Covariates in the Synthetic Control Method." *The Econometrics Journal*, January, utz001. https://doi.org/10.1093/ectj/utz001.

Brodersen, Kay H., Fabian Gallusser, Jim Koehler, Nicolas Remy, and Steven L. Scott. 2015. "Inferring Causal Impact Using Bayesian Structural Time-Series Models." *The Annals of Applied Statistics* 9 (1): 247–74. https://doi.org/10.1214/14-AOAS788.

Campos, Nauro F., Fabrizio Coricelli, and Luigi Moretti. 2019a. "Institutional Integration and Economic Growth in Europe." *Journal of Monetary Economics* 103 (May): 88–104. https://doi.org/10.1016/j.jmoneco.2018.08.001.

———. 2019b. "Institutional Integration and Economic Growth in Europe." *Journal of Monetary Economics* 103 (May): 88–104. https://doi.org/10.1016/j.jmoneco.2018.08.001.

Doudchenko, Nikolay, and Guido Imbens. 2016. "Balancing, Regression, Difference-in-Differences and Synthetic Control Methods: A Synthesis." w22791. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w22791.

Durbin, J., and S. J. Koopman. 2012. *Time Series Analysis by State Space Methods*. 2nd ed. Oxford Statistical Science Series 38. Oxford: Oxford University Press.

Kaul, Ashok, Stefan Kloßner, Gregor Pfeifer, and Manuel Schieler. n.d. "Synthetic Control Methods: Never Use All Pre-Intervention Outcomes Together with Covariates," 24.

Kinn, Daniel. 2018. "Synthetic Control Methods and Big Data." *arXiv:1803.00096 [Econ]*, February. http://arxiv.org/abs/1803.00096.

Scott, Steven L, and Hal Varian. n.d. "Predicting the Present with Bayesian Structural Time Series," 21.

Scott, Steven, and Hal Varian. 2013. "Bayesian Variable Selection for Nowcasting Economic Time Series." w19567. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w19567.

Comments to Danny

1. This is a really exciting project. You should definitely pursue it. I would be pleased to see more as it develops.
2. One big point that it missing: What happens during the policy period? What weights get applied? I can think of a variety of approaches, but the answer really isn't obvious.
3. The issue shouldn't be a test for constant weights. You want something that gives evidence that you get a better estimate of the policy effect by allowing for nonconstant weights.
4. You should set up a Monte Carlo platform to study what happens under various circumstances.