Bayesian Synthetic Control Methods

Sungjin Kim, Clarence Lee, Sachin Gupta

Presentation by Sungjin Kim

Sungjin Kim is a Doctoral Candidate (e-mail: sk2756@cornell.edu), Clarence Lee is Assistant Professor (e-mail: clarence.lee@cornell.edu), and Sachin Gupta is the Henrietta Johnson Louis Professor of Management (e-mail: Sachin.gupta@cornell.edu), at the Johnson Graduate School of Management, S.C. Johnson College of Business, Cornell University.

We seek to make a methodological contribution in the domain of econometric methods for panel data to measure treatment effects. In particular, we propose Bayesian Synthetic Control Methods (BSCMs) which offer multiple advantages over extant Synthetic Control Methods (SCMs). Since the introduction of SCMs by Abadie and Gardeazabal in their article in the *American Economic Review* in 2003, these methods have been widely employed in a number of fields such as marketing, economics, political science, statistics, public health, medicine, and so forth. The original SCMs have also been improved upon, such as in a recent *NBER* working paper by Doudchenko and Imbens (2016).

In the current study, we identify several limitations of extant SCMs and propose solutions based on a Bayesian framework. Specifically, the proposed BSCMs do not impose any restrictive constraints on the parameter space a priori. Moreover, they allow statistical inference in a straightforward manner, and a natural mechanism to deal with the "large p, small n" and sparsity problems through MCMC procedures. BSCMs show how the effect of the treatment and its uncertainty evolve over time, as opposed to a single aggregate estimate of the treatment effect such as the average treatment effect.

We demonstrate strong evidence of the superiority of our proposed methods over extant methods in an extensive simulation study. We find that the proposed BSCMs almost always dominate other models in terms of predictive accuracy in a variety of data situations. Even when the data generating process (DGP) is consistent with all the standard SCM constraints, the proposed models perform better when "large p, small n" or sparsity concerns exist. When the DGP violates the SCM constraints, the standard SCM is not identified whereas the proposed models show excellent predictive accuracy.

We show how the proposed methods can be applied to real world data by estimating the impact of a tax on soda in the state of Washington in 2010. The proposed models show good in-sample fit in the pre-treatment periods and better predictive accuracy compared to extant models. We find that the tax in Washington led to a 5.7% increase in retail price and 5.5~5.8% decrease in sales of soda. We also find that retailers in Washington over-shifted the tax to consumers, with a pass-through rate of about 121%.