Bayesian Shrinkage Among Time Varying Coefficients in Counterfactual Analysis

Motivation

- Synthetic Control methods have become a major empirical tool (ADH approach)
  - Most development been on the frequentist side
  - Abadie and Gardeazabal (2003), Abadie, Diamond, and Hainmueller (2010) initially
  - Developments include: Athey et al. (2018), Xu (2017), Powell (2018), Kaul et al. (2018) to name a very very limited few
- Other approaches to the synthetic control framework
  - Brodersen et al. (2015)
  - Uses Bayesian variable selection and state space models (Kalman Filter)
  - Based off of Scott and Varian (2013)
  - This paper builds off of this literature

Basic Setup

- There is a unit $Y_{0,t}$ (aka country GDP)
- An endogenous intervention, d, occurs at $t = T_0$ (aka policy change) to $Y_{0,t}$ and stays forever
  - $Y_{0,t}(1)$: unit 0 when a treatment has occurred
  - $Y_{0,t}(0)$: unit 0 when a treatment has not occurred
  - Observe $Y_{0,t} = dY_{0,t}(1) + (1 - d)Y_{0,t}(0)$
- Want to estimate $Y_{0,t}(0)$ when $d = 1$ using untreated units (aka other countries)
  - $Y_{0,t} = f(Y_{1,t}, ..., Y_{J,t})$

Benefits of a Time Series Bayesian Approach

- Allows modeling time series component explicitly
- Allows modeling nonconstant relationships between control and treatment
  - Time varying coefficients
- Allows stronger structure to use "expert knowledge"

Brodersen et al. (2015) Basic Model

$$
\begin{aligned}
y_t &= \mu_t + X_t\beta + \epsilon_t & \epsilon_t &\sim \mathcal{N}(0, \sigma^2) & (1)\\
\mu_{t+1} &= \delta_t + \mu_t + \eta_{1,t} & \eta_{1,t} &\sim \mathcal{N}(0, \sigma^2_{\eta_1}) & (2)\\
\delta_{t+1} &= \delta_t + \eta_{2,t} & \eta_{2,t} &\sim \mathcal{N}(0, \sigma_{\eta_2}) & (3)
\end{aligned}
$$

- Flexible model allows for changes in the linear trend term
- Allows for β to be time varying
- Bayesian priors (spike and slab) used on β to avoid overfitting
- Problem: Coefficients are either time varying or constant
  - Bayesian priors only apply to β, but can decompose β into time varying and time invariant parts

My Initial Simulation

$$
\begin{aligned}
y_{0,t}(0) &= \sum_{j=1}^{J} w_{j,t} \left(c_j t + z_j + \gamma_j sin\left(\frac{\pi t}{\rho_j}\right)\right) + \epsilon'_{1.t} \\
w_{1,t} &= w_{1,t-1} + \frac{.6}{T} \\
w_{2,t} &= w_{2,t-1} - \frac{.6}{T} \\
w_{j,t} &= w_{j,t-1} & j \notin \{1,2\}
\end{aligned}
$$

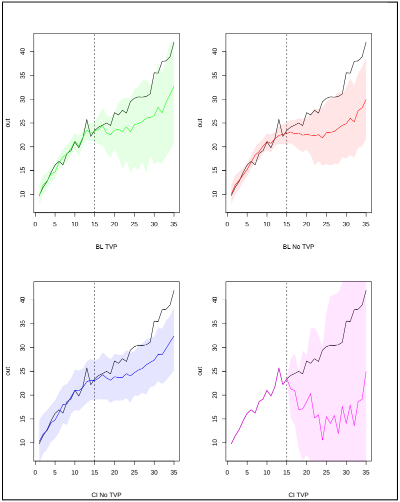
with initial conditions:

$$
\begin{aligned}
w_{1,0} &= .2 \\
w_{2,0} &= .8 \\
w_{j,0} &= 0 \quad j \notin \{1,2\}
\end{aligned}
$$

Parameters - Linear Weights

1) $c_{1,t} = .75$, $c_{2,t} = .25$, and $c_{j,t} \sim U[0,1]$ for all $j \notin \{1,2\}$
2) $z_1 = 25$, $z_2 = 5$ and z_j is sampled from $\{1,2,3,4,...,50\}$
3) $\epsilon'_{j,t} \sim N(0,1)$
4) T $= 35$, $T_0 = 15$
5) J $= 15$
6) $w_{1,t} = .2 + .6\frac{t}{T}$, $w_{2,t} = 1 - w_{1,t}$, and $w_{j,t} = 0$ for all else
7) $\gamma_j = 0 \; \forall j$

Visual Comparison - Linear Weights



Simulation Results - Linear Weights

Table 1: Mean Results of 100 Simulations

Model	pre.treat.mse	pre.treat.coverage	post.treat.mse	post.treat.coverage	CI.Spread	Bias
Causal Impact No TVP	0.828	1.000	17.427	0.838	10.295	8.574
Causal Impact TVP	0.000	1.000	30.994	0.954	13.396	8.950
Bayesian Lasso No TVP	0.549	0.948	24.676	0.589	9.361	2.043
Bayesian Lasso TVP	0.267	0.991	17.468	0.772	10.176	3.171

Simulation Results - Linear Weights

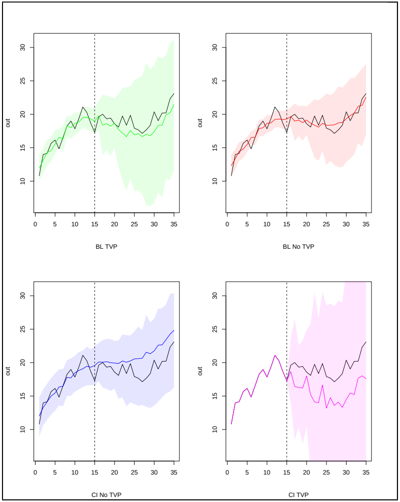
Table 2: Median Results of 100 Simulations

Model	pre.treat.mse	pre.treat.coverage	post.treat.mse	post.treat.coverage	CI.Spread	Bias
Causal Impact No TVP	0.779	1.000	16.137	0.853	10.258	8.704
Causal Impact TVP	0.000	1.000	19.429	1.000	13.145	9.026
Bayesian Lasso No TVP	0.512	0.941	23.192	0.647	9.013	1.856
Bayesian Lasso TVP	0.216	1.000	13.901	0.853	9.995	2.847

Parameters - Constant Weights

1) $c_{1,t} = .75$, $c_{2,t} = .25$, and $c_{j,t} \sim U[0,1]$ for all $j \notin \{1,2\}$
2) $z_1 = 25$, $z_2 = 5$ and z_j is sampled from $\{1,2,3,4,...,50\}$
3) $\epsilon'_{j,t} \sim N(0,1)$
4) T $= 35$, $T_0 = 15$
5) J $= 15$
6) $w_{1,t} = w_{1,t-1}$, $w_{2,t} = 1 - w_{1,t}$, and $w_{j,t} = 0$ for all else
7) $\gamma_j = 0 \ \forall j$

Visual Comparison - Constant Weights



Simulation Results - Constant Weights

Table 3: Mean Results of 100 Simulations

Model	pre.treat.mse	pre.treat.coverage	post.treat.mse	post.treat.coverage	CI.Spread	Bias
Causal Impact No TVP	0.702	0.999	6.814	0.898	6.452	2.638
Causal Impact TVP	0.000	1.000	16.372	0.933	9.787	1.131
Bayesian Lasso No TVP	0.619	0.838	2.714	0.926	6.322	1.002
Bayesian Lasso TVP	0.317	0.978	5.133	0.919	8.468	1.108

Simulation Results - Constant Weights

Table 4: Median Results of 100 Simulations

Model	pre.treat.mse	pre.treat.coverage	post.treat.mse	post.treat.coverage	CI.Spread	Bias
Causal Impact No TVP	0.725	1.000	5.585	1.000	6.486	2.706
Causal Impact TVP	0.000	1.000	7.912	1.000	9.286	0.852
Bayesian Lasso No TVP	0.651	0.824	1.993	0.941	6.323	1.041
Bayesian Lasso TVP	0.258	1.000	3.185	1.000	8.201	1.352

Work Cited

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, Mohsen Bayati, Nikolay Doudchenko, Guido Imbens, and Khashayar Khosravi. 2018. "Matrix Completion Methods for Causal Panel Data Models." *arXiv:1710.10251 [Econ, Math, Stat]*, September. http://arxiv.org/abs/1710.10251.

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.