A Gaussian Process Model for Causal Inference with TSCS Data

Nuannuan Xiang, Kevin M. Quinn

nnxiang@umich.edu, kmq@umich.edu

Department of Political Science, The University of Michigan

Motivation

- **DID** is widely used for causal inference with TSCS data, but its **parallel trend** assumption often fails.
- The **(Generalized) Synthetic Control (GSC)** Method relaxes the parallel assumption, assuming the confounder takes a **linear additive form** of latent time and unit specific factors.[a]

[a] Abadie, Diamond, and Hainmueller (2010); Xu(2017).

The GP Approach

Our approach relaxes the linearity assumption of GSC by using a Gaussian Process (GP) model with a flexible **Spectral Mixture Kernel** to impute counterfactuals.

Proposed Methodology

- **Model Untreated Units**

$$y_{it} = f_i(\mathbf{x}_{it}) + \varepsilon_{it}$$
$$\varepsilon_{it} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$$
$$f_i(\mathbf{x}_{it}) \sim \mathcal{GP}(m_i(\mathbf{x}_{it}),\ k(\mathbf{x}_{it}, \mathbf{x}_{it'}))$$
$$m_i(\mathbf{x}_{it}) = \mathbf{x}_{it}'\boldsymbol{\beta} + \mathbf{w}_{it}'\boldsymbol{\alpha}_i$$

For the following applications:

$$\mathbf{w}_{it}' = \mathbf{x}_{it}' = (1, t)$$

Spectral Mixture Kernel (Wilson and Adams 2013):

$$k(\mathbf{x}_{it}, \mathbf{x}_{it'}) = \sum_{q=1}^{Q} \omega_q \exp\{-2\pi^2(t-t')^2\nu_q\}\cos(2\pi(t-t')\mu_q)$$

- **Selected Priors**

$$\omega_q \sim p_\omega \mathbb{I}(\omega_q = 0) + (1-p_\omega)(0.5\mathcal{E}xp(r_{\omega,1}) + 0.5\mathcal{E}xp(r_{\omega,2}))$$

- **Estimate parameters and counterfactuals $y_{it}^{(0)}$ through MCMC**

Counterfactuals $y_{it}^{(0)}|\text{treat}=1$: the untreated outcomes of treated units

- **Estimate ATT, ATT$_i$, ATT$_t$**

$\widehat{ATT}_{it} = y_{it} - \hat{y}_{it}^{(0)}$, where unit i is treated at time t.

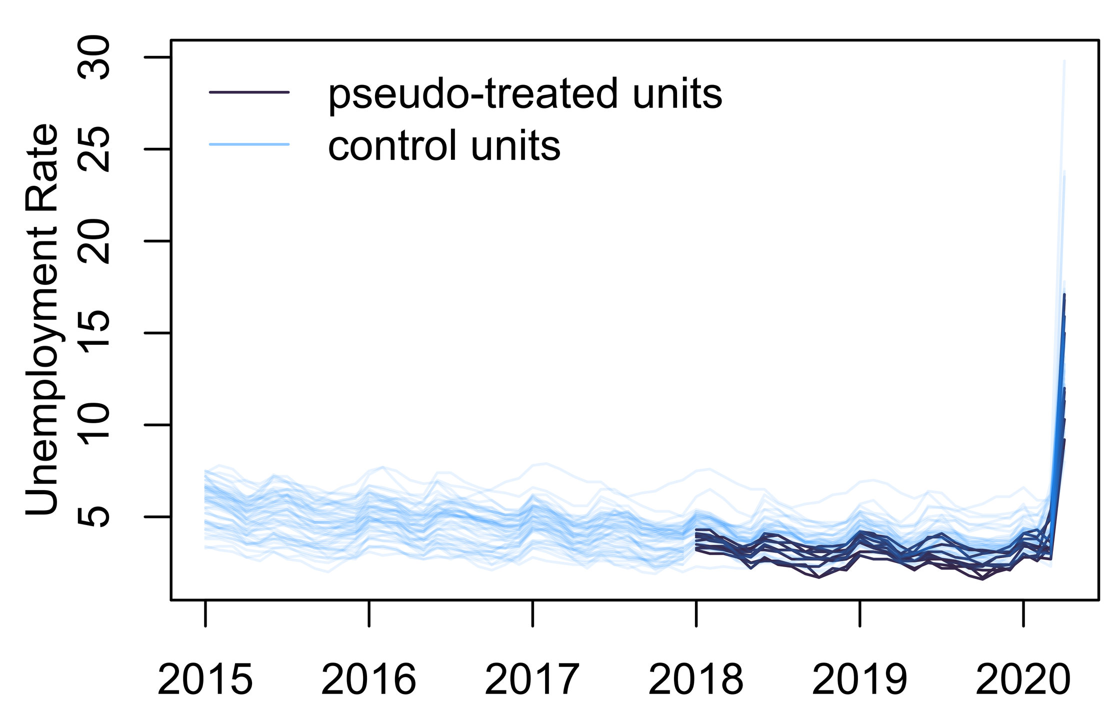
$\widehat{\text{ATT}}$, $\widehat{\text{ATT}}_i$, and $\widehat{\text{ATT}}_t$ are corresponding averages of $\widehat{\text{ATT}}_{it}$.

Application Examples

- Two TSCS data sets of US states;
- We give selected states in selected time periods a **pseudo treatment status**;
- The counterfacturals, $y_{it}^{(0)}|\text{treat}=1$, are known;
- The true ATT, ATT$_i$ and ATT$_t$ are all 0.

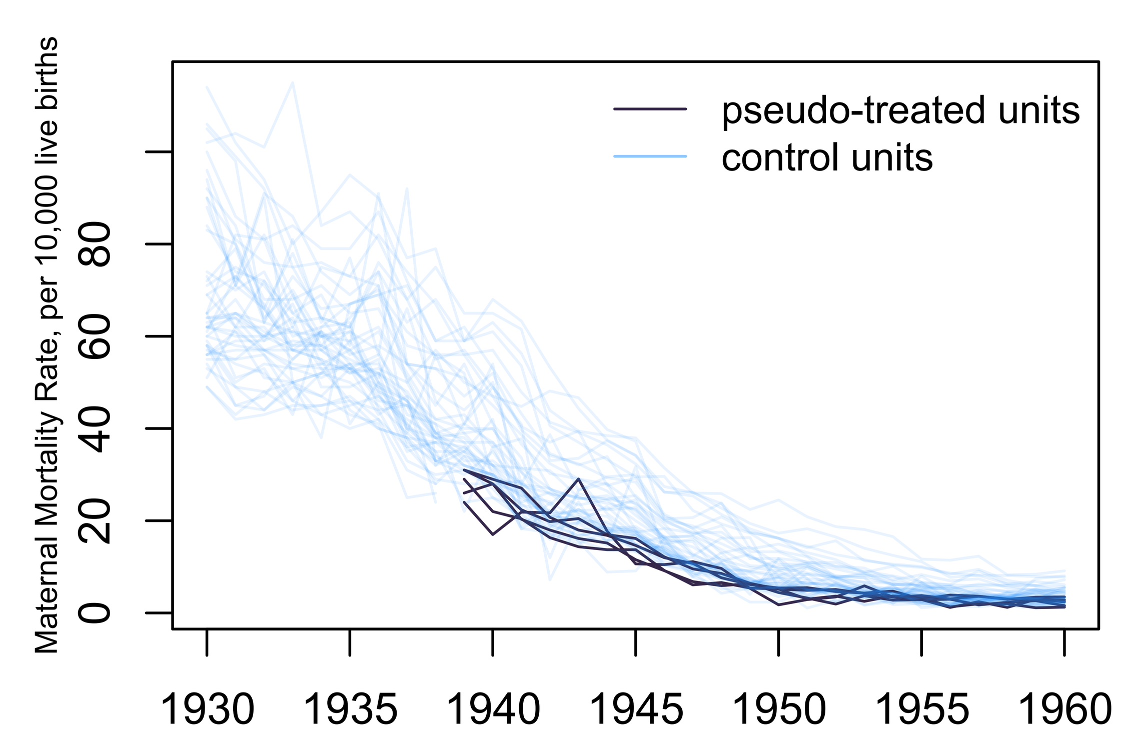
Case 1: Monthly Unemployment Data

- From Jan. 2015 to Apr. 2020;
- States with 2nd, 4th, 6th, ..., 20th lowest mean unemployment rates in 2017 are "treated" since Jan. 2018.

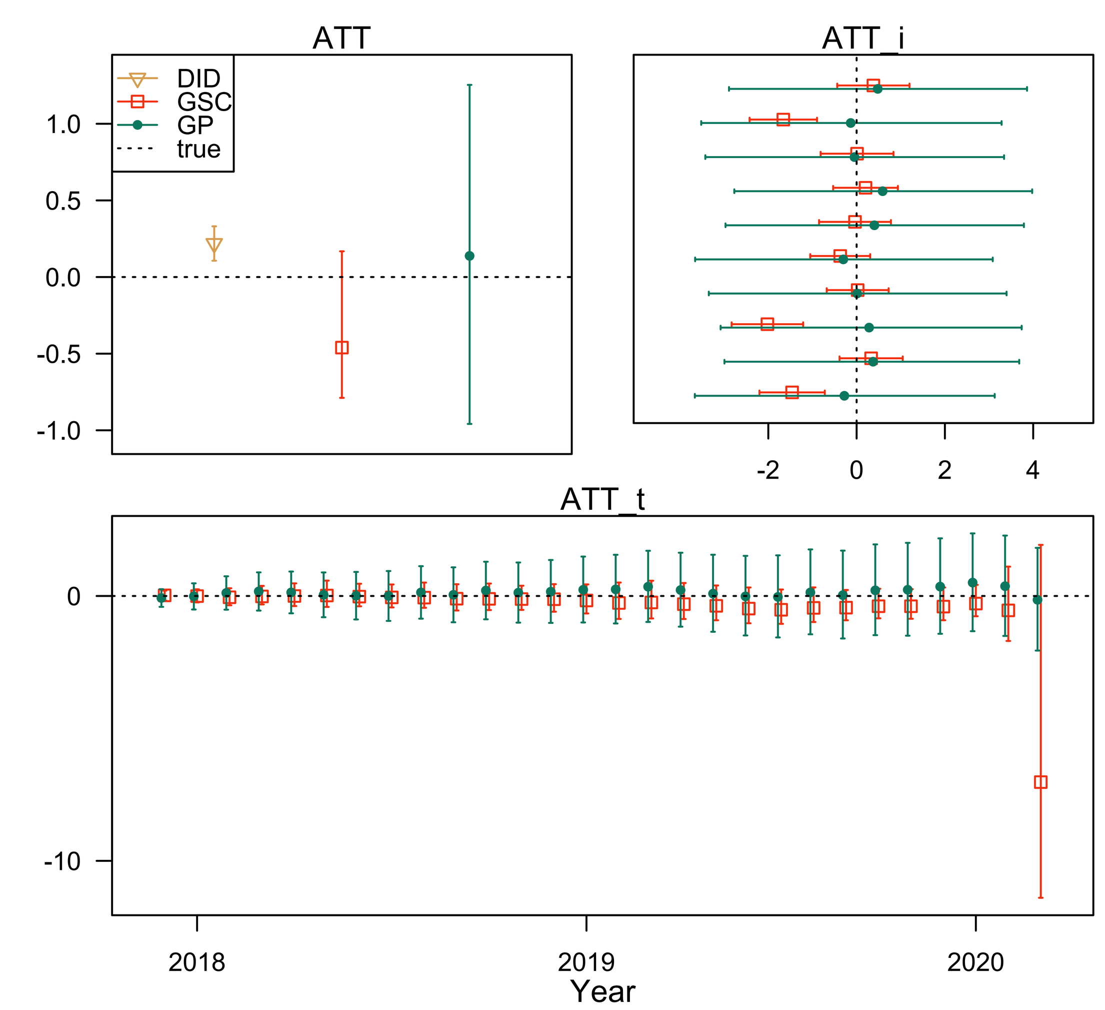


Case 2: Yearly Maternal Mortality

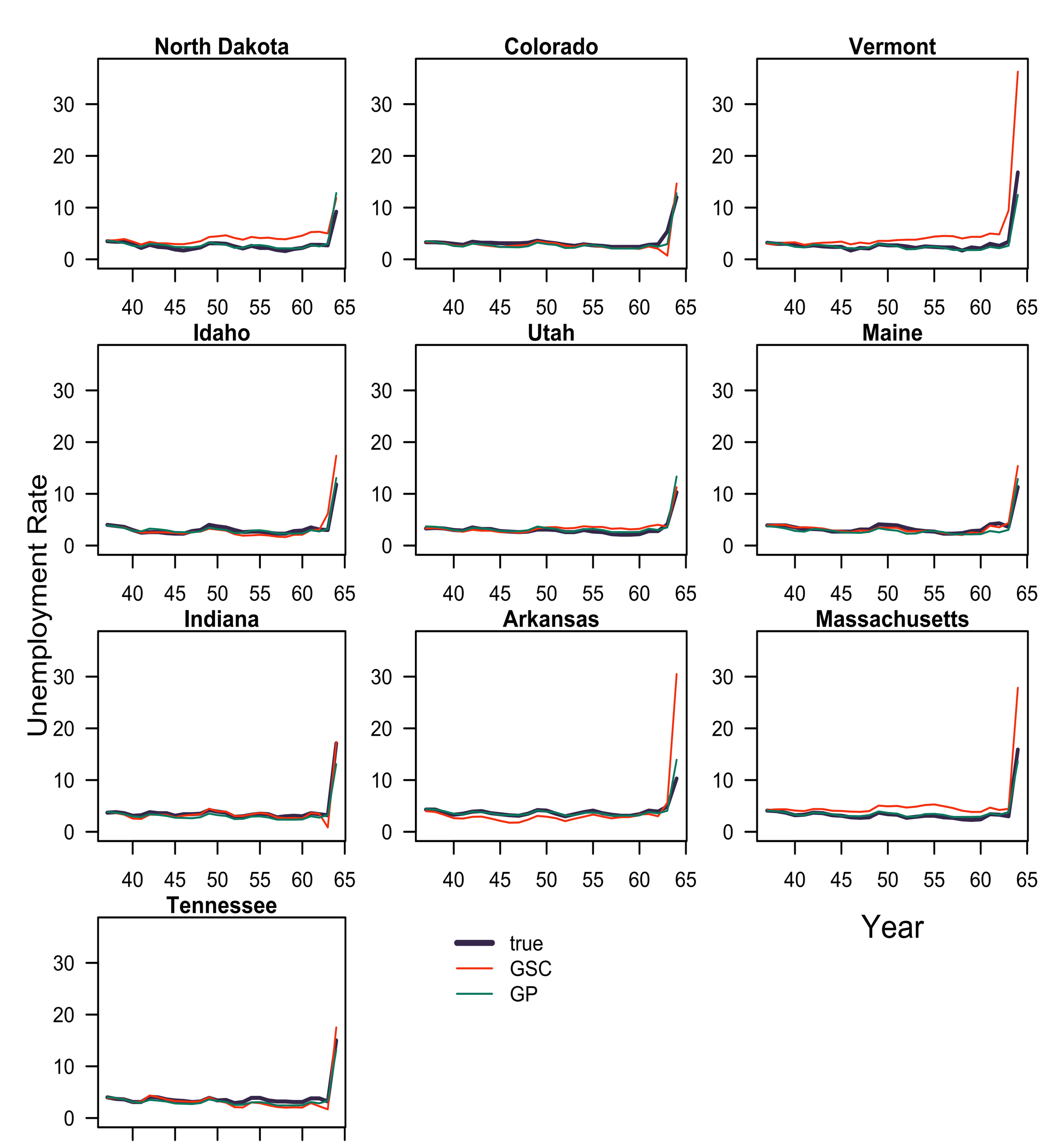
- From 1930 to 1960;
- The 2nd, 4th, ..., 10th states with lowest mortality rates in 1938 are "treated" since 1939.



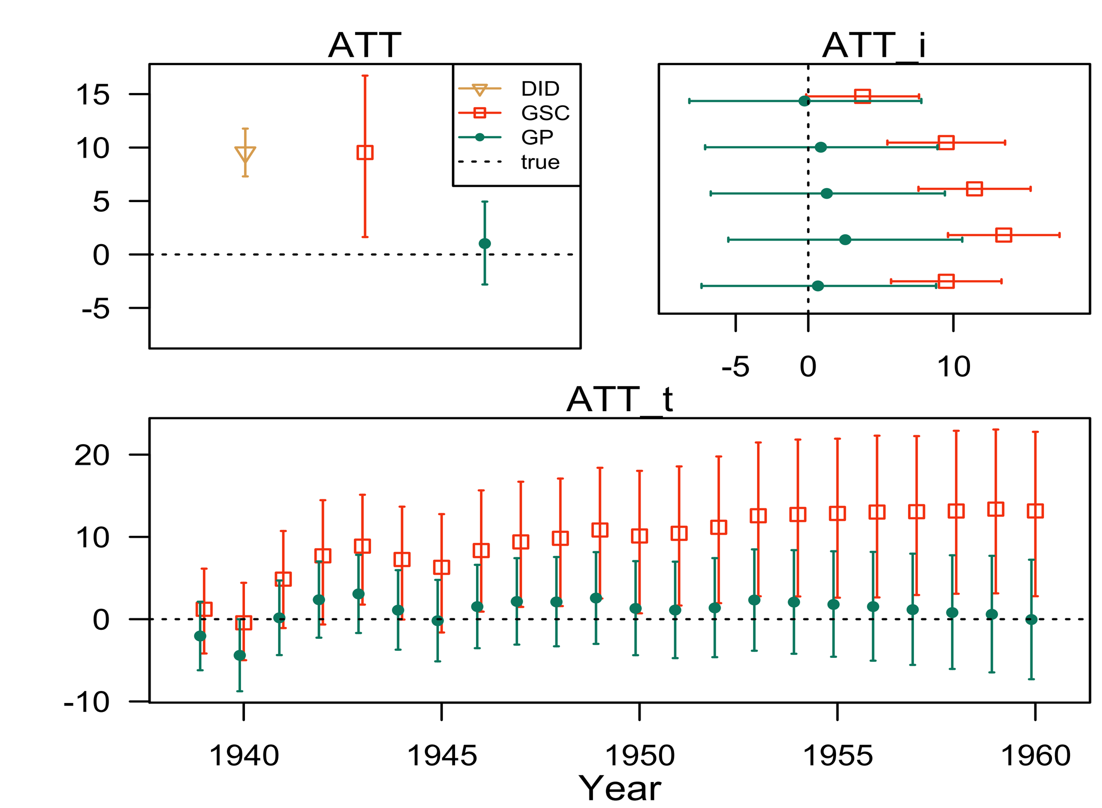
Results of Case 1: Unemployment



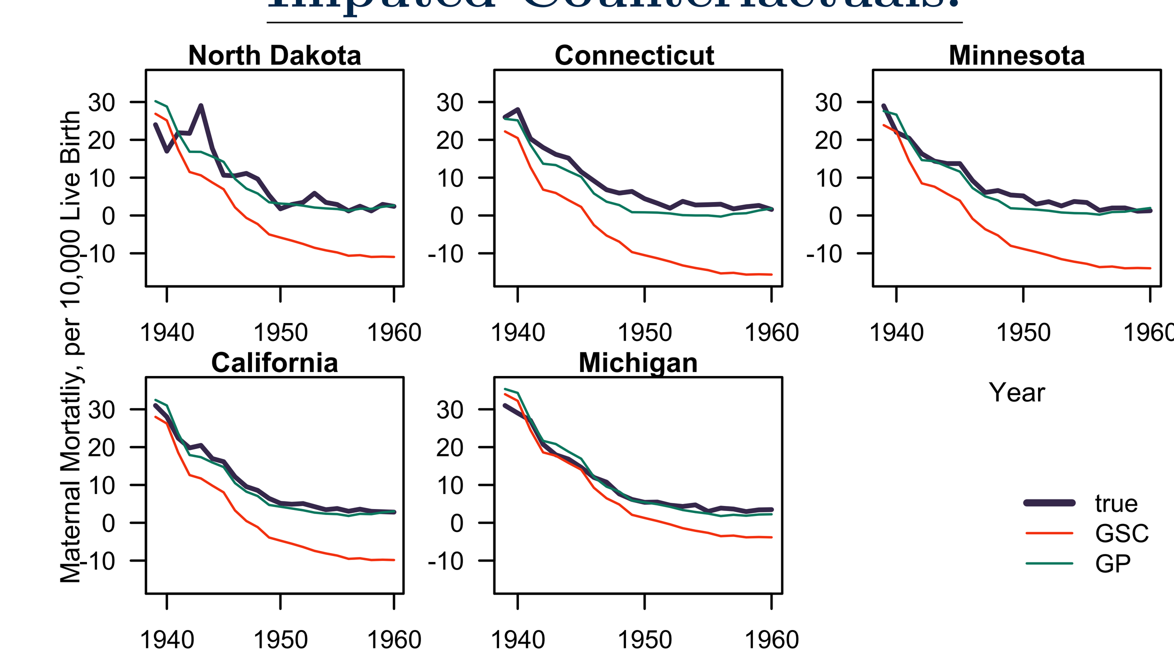
Imputed counterfactuals:



Results of Case 2: Mortality



Imputed Counterfactuals:



500 Monte Carlo Results

Based on the original data, we simulate 500 datasets:

Case	Model	Bias	SD	MSE	95% Coverage [a]
Case 1	GP	-0.086	0.134	0.027	0.994
	GSC	0.098	0.306	0.103	0.988
	DID	0.246	0.121	0.076	0.118
Case 2	GP	-0.142	0.872	0.779	1
	GSC	4.938	5.192	51.284	0.764
	DID	7.651	2.420	64.380	0.008

[a] A valid 95% Bayesian credible interval will not necessarily cover the truth in 95% of repeated samples.

Takeaways

- Like the two examples, many social science applications feature complicated dynamics that are not easily captured by standard modeling approaches.
- The Spectral Mixture Kernel employed by our GP approach naturally captures a wide range of dynamics and allows our method to succeed where other approaches fail.