Final Report

Measuring causal impact using a Bayesian structural time series model

Sai Arjun Koneru, Lucy Quirant,
Alexander Reisach, Sofia Vazquez Alferez

KEN4258: Computational Statistics

Faculty of Science and Engineering
Dpt. of Data Science and Knowledge Engineering
Maastricht University
Netherlands

Contents

1 Introduction

After launching a digital marketing campaign, companies often want to know what their return on investment was. This has sparked interest into ways of quantifying the effectiveness of marketing campaigns, which can be understood as a particular case of causal inference through structural intervention [7]. In this context, the authors of [6] proposed a fully Bayesian state space model for time series that generalizes amongst others multiple linear regression, Auto-regressive models, and Moving Average models. The model is designed for discrete interventions such as marketing campaigns, discount offers, etc. and works by predicting a synthetic counterfactual (what would have happened without an intervention) and looking at the difference between the actual outcome and this prediction. If the counterfactual prediction is close to the observed behavior, the intervention is assumed not to have had a causal impact. In this paper we investigate the causal impact approach by looking at how it behaves with synthetic data for which we know the intervention, and two different real world data sets that have been studied with different approaches by other researchers. The goal is to answer the research question *How effective is the model at predicting the impact of an intervention?* for which we have the following sub questions:

- How do the predictions of the model compare to the true impact of an intervention?
- How does the performance of the model compare to the performance of another research paper in the area? (Does the model confirm findings from other methods when applied to their data set?)

This paper is structured as follows: first, we introduce previous approaches in the field of causal inference (section 2). Next, the Bayesian structural time series model used in the chosen paper is explained in depth (section 3). Thereafter, we present the results of using the *CausalImpact* approach on simulated data under various conditions to test the model's behaviour (section 4). Then, we apply the model to real life data obtained in randomized controlled trials (RCT) (section 5.1). Finally, we compared the model performance against an alternative approach for intervention based causal inference on time series data [14] (section 5.2). The results of replicating the simulations of the original paper can be found in the Appendix. All code used for this review can be found under `https://github.com/Scriddie/causal_impact`.

2 Related Work

Causal Inference on the effect of an intervention based on observational time series data has a long and troubled history in Statistics. Generally speaking, estimating the causal impact of an intervention requires the separation of the intervention effect from all other effects that could act as confounders. Certainly, the most popular such method is the randomized control trial as advocated by R.A. Fisher himself in [8]. Modern approaches often relying on structural causal models still follow the same fundamental approach of shielding the target variable from all confounders while controlling for no more than necessary (see [10] for an overview). Still, any such method requires great scrutiny and a degree of knowledge and control over the intervention setup that can not be expected in most real-world scenarios.

For observational studies in particular, researchers have historically relied on ordinary least squares (OLS) *Difference-in-Difference* (DID) [1] estimation whilst controlling for likely and available confounder data. However, this method contains many obvious flaws, such as an often-violated parallel trends assumption for treatment and control group, as well as the disregard to serial auto correlation of ordinary OLS which assumes data to be i.i.d. [4]. *Synthetic Control Methods* attempt to leverage the available data on treated and non-treated units to find a quasi-control group with which to compare the change in the treated units, thereby combining the idea of matching from comparative case studies with DID [2]. One such method by [14] is briefly described in section 5.2.

The paper [6] examined in this review goes one step further and makes use of a *Bayesian Structural Time Series* (BSTS) model, which allows for local trends, auto correlation, seasonality, and flexible coefficients on contemporaneous covariate time series. By doing away with the structural limitations imposed by OLS and flexible incorporation of related time series, the model attempts to construct a

close fit to an absent control group. In cases where the provided related time series indeed make for a randomized control group, or only differ from one in ways that the (BSTS) can adjust for, this can allow for causal inference on observational data.

3 Bayesian Structural Time Series for Causal Impact

Structural time series models make use of different state components such as level, trend, seasonality and regression factors. Bayesian posterior sampling on this model gives rise to the name Bayesian Structural Time Series (BSTS)[12]. This type of modelling is flexible as it can represent all ARIMA models in a structural time series model. Some of the applications of Bayesian structural time series models are nowcasting, forecasting and finding relevant predictor variables using spike and slab prior[9].

3.1 State Components

A BSTS model can be generated from different state components. We first start explaining a simple BSTS model with one simple component and then further build on this model by adding other components. Note that time series data can consist of any combination of these components.

3.1.1 Local Level

Assume we have time series data that represent sales of a company throughout a year but the product they are selling is always in demand. It doesn't vary with seasons or doesn't increase or decrease steadily with respect to time. We can model this using a level component mathematically described as follows

$$\begin{aligned} y_t &= \mu_t + \epsilon_t, \\ \mu_{t+1} &= \mu_t + \eta_t \end{aligned} \tag{1}$$

where $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$ and $\eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$ are independent from one another. These account for the noises and variation along the level in our time series data. This model can also be viewed as a random walk plus noise model. We need to specify μ_1 to generate this data.

3.1.2 Local Linear Trend

Now assume that the sales of the company are increasing or decreasing, meaning that they adhere to a trend. We can incorporate this in our model by adding the trend component. The resulting model can be defined as

$$\begin{aligned} y_t &= \mu_t + \epsilon_t, \\ \mu_{t+1} &= \mu_t + \delta_t + \eta_t \\ \delta_{t+1} &= \delta_t + \xi_t \end{aligned} \tag{2}$$

where $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$, $\eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$ and $\xi_t \sim \mathcal{N}(0, \sigma_\xi^2)$ are independent from one another. To generate this data we need to specify δ_1 and μ_1. If we set $\sigma_\xi^2 = 0$ then it moves with a slope δ_1. If we set $\sigma_\eta^2 = 0$ then the trend (or slope) we observe is equal to δ_1 with an intercept equal to μ_1.

3.1.3 Local Linear Trend with Seasonality

If the sales also vary seasonally, a reasonable assumption in practical settings, we can account for this with the seasonality component. The number of seasons depends on how we measure or generate the data. If we are talking about monthly data then the number of seasons would equal 12, whereas for

weekly data it would equal 52. Such a model can be defined as

$$\begin{aligned} y_t &= \mu_t + \gamma_t + \epsilon_t, \\ \mu_{t+1} &= \mu_t + \delta_t + \eta_t \\ \delta_{t+1} &= \delta_t + \xi_t \\ \gamma_{t+1} &= \sum_{j=1}^{s-1} \gamma_{t+1-j} + \omega_t \end{aligned} \tag{3}$$

where $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$, $\eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$, $\xi_t \sim \mathcal{N}(0, \sigma_\xi^2)$ and $\omega_t \sim \mathcal{N}(0, \sigma_\omega^2)$ are independent from one another. The expectation of the sum of seasonal effects is set to zero as we want the overall effect to even out. When generating this model we need to initialize parameters in local level trend model as well as initialize the seasonal effects under the constraint that the sum of the effects is equal to zero. Note that this is different from cyclic time series data.

3.1.4 Adding a Regression Component

The often most effective component in a BSTS model is the regression component. Suppose we have information of sales of another company and we think that our sales behave similarly but are independent, then we can use this information to perform any kind of prediction using linear regression. There are two ways in which we can think of these regression coefficients.

One way to do this is to assume that the regression coefficients are static. That is, they do not change over time. we can mathematically model this as follows

$$y_t = \mu_t + \gamma_t + \beta^T x_t \tag{4}$$

where μ_t, γ_t are defined from the equations (1),(2),(3) depending on the components we want. x_t represents our set of control time series. we assume that these are contemporaneous. Spike-and-slab prior, as explained in section 3.2.2, allows us to avoid overfitting and removes our constraint on fixing a set of covariates.

Another possible way of include regression coefficients is to allow them to vary over time[3]. These are called dynamic coefficients. we can mathematically model these as

$$\begin{aligned} x_t^T \beta &= \sum_{j=1}^{J} x_j \beta_{j,t} \\ \beta_{j,t+1} &= \beta_{j,t} + \eta_{\beta,j,t} \end{aligned} \tag{5}$$

where $\eta_{\beta,j,t} \sim \mathcal{N}(0, \sigma^2_{\beta_{j,t}})$ and $\beta_{j,t}$ represents coefficients for j's control series given covariates j = 1,.,....,J. Depending on the time series we need to choose on static or dynamic coefficients. A spike-and-slab prior allows us to infer which covariate behaves similary to our target time series.

3.1.5 Generic Representation of BSTS

Given different components we can represent the state space models in a more generic way. This allows us to represent BSTS in a concise form. It can be written as

$$\begin{aligned} y_t &= Z_t^T \alpha_t + \epsilon_t \\ \alpha_{t+1} &= T^T \alpha_t + R_t \eta_t \end{aligned} \tag{6}$$

The noise components ϵ_t and η_t are gaussian and independent of everything else. The equation used to calculate y_t is called the observation equation and the equation used to calculate α_t is called the state equation. The outcome y_t is a scalar observation, Z_t is a d-dimensional output vector, T_t is a $d \times d$ transition matrix, R_t is a $d \times q$ control matrix, ϵ_t is a scalar observation error with noise variance σ_t , and η_t is a q-dimensional system error with a $q \times q$ state-diffusion matrix Q_t , where $q \leq d$, and which represents the components of the model we want to invoke as 1's and 0's.

3.2 Bayesian Estimation

3.2.1 Prior Distribution

Whichever model components we decide to include, we will have to estimate model parameters. In order to be able to perform Bayesian inference it is therefore necessary to specify the prior distribution of said parameters, as well as the prior distribution of the initial state given the specified parameters.

The models explained in Section 3.1 mainly depend on variance parameters σ that control the diffusion (i.e. level of noise) of the individual state components. The prior distribution for such variance parameters is typically given by a Gamma distribution:

$$\frac{1}{\sigma^2} \sim G(\frac{v}{2}, \frac{s}{2}) \tag{7}$$

Where s is the prior sum of squares, and v is the weight assigned to the prior estimate. Here we assume that the incremental errors in the state process are small so small values of v and $\frac{s}{v}$ are chosen. Given the sample variance s_y^2, for seasonal and local linear trends this means choosing:

$$\frac{1}{\sigma^2} \sim G(10^{-2}, 10^{-2} s_y^2) \tag{8}$$

The model also places a spike-and-slab prior over the coefficients of the different covariates or controls that are used to predict the counterfactual. This allows the model to choose an appropriate set of most informative covariates by inducing sparsity on the coefficients.

The spike-and-slab prior is a mixture of two different distributions: the spike, which is a point-mass centered at 0 and a slab, which is a very diffuse distribution, here a Gaussian with large variance. The spike-and-slab then draws β (a coefficient) from the diffuse prior with probability p and from the pointmass with probability $1 - p$. If the data is not very informative about a non-zero value for a given coefficient this will not be able to push the coefficient away from zero and thus it will not appear in the model.

3.2.2 MCMC

Brodersen et al. sample from the posterior distribution using Markov Chain Monte Carlo (MCMC) since no closed-form calculation is possible for the chosen model [6]. A Markov chain is defined as a chain of states where each state is only dependent on its predecessor. Given the transition probability from state i to state j called $q_{ij} = P(X_{n+1} = j | X_n = i)$, we can define a transition matrix Q for all states. A distribution is said to be stationary once the probability of a state X_j is equal to the sum of all products of transition probability q_{ij} and the probability of a given other state X_i for all j. This entails that the Markov chain will maintain its stationary distribution. The goal of MCMC is to build a Markov chain such that it leads to the desired stationary distribution. To this end, an auxiliary Markov chain is created where an algorithm assigns higher probability to those transitions that agree with the stationary distribution. Monte Carlo refers to the algorithm sampling repeatedly from this Markov chain to obtain the posterior probability distribution [13]. In our case the Markov chain has a stationary distribution of $p(\theta, \alpha | \mathbf{y}_{1:n})$ with model parameters θ and state vector α which are simulated based on observed data $\mathbf{y}_{1:n}$. A Gibbs sampler alternates between data-augmentation where it simulates from $p(\alpha | \mathbf{y}_{1:n}, \theta)$, and parameter-simulation where it simulates from $p(\theta | \mathbf{y}_{1:n}, \alpha)$. Due to equation conjugacy, the posterior distribution remains a Gamma distribution (section 3.2.1) [6].

3.3 Causal Impact

3.3.1 Assumptions

CausalImpact's ability to correctly evaluate causal impact relies on several assumptions. First and foremost, the provided related time series which in combination form the foundation for the synthetic control

estimation need to have all the characteristics of a randomized control group, save for those the BSTS model can adjust for. The BSTS itself requires continuity of the model structure for the estimations to be accurate. Moreover, the BSTS model must be able to capture or approximate the underlying data generating process and the Priors chosen must be able to converge to their true value with the amount of data available.

3.3.2 Posterior Inference

With MCMC (section 3.2.2) we calculated the posterior over model parameters and states $p(\theta, \boldsymbol{\alpha} \,|\mathbf{y}_{1:n})$; which is relevant to calculate the posterior incremental effect, also called the posterior predictive density, a distribution over all counterfactual datapoints.

$$p(\mathbf{y}_{n+1:m}|\mathbf{y}_{1:n}, \mathbf{x}_{1:m}) \tag{9}$$

Counterfactuals represent how the treated market/ group would have looked like, had the intervention not taken place. These counterfactuals depend on the activity of the treated group before the intervention took place, on the activity of similar control groups (before and during intervention), and on the priors. The posterior is not dependent on other parameter estimates or regression covariates.

The impact of the intervention can be calculated pointwise for each draw τ, by subtracting y_t (at a particular timepoint $t = n+1, ..., m$) from the corresponding $\widetilde{y}_t$ sampled from the posterior predictive distribution $\phi_t^{(\tau)} = y_t - \widetilde{y}_t$. To observe the cumulative effect of an intervention, two metrics are possible.

Firstly, the cumulative sum (equation 10a) of causal increments is used to quantify a flow quantity (such as additional number of quantity of interest observed in a certain timespan). It is useful when we assume that an impact might first have a strong effect, and then slowly reduce, the cumulative impact will approach the true value as the forecasting period is increased.

$$(a) \sum_{t'=n+1}^{t} \phi_{t'}^{(\tau)} \qquad\qquad (b) \quad \frac{1}{t-n} \sum_{t'=n+1}^{t} \phi_{t'}^{(\tau)} \tag{10}$$

Secondly, the posterior running average (equation 10b) effect is applicable for a flow or a stock quantity (total number of quantity of interest at a certain point in time). However, unlike the cumulative sum, it will eventually approach zero as the forecasting period is extended. In general, as the forecasting period increases, results (for both metrics) will become more uncertain.

4 Simulated Data

In this section we present experiments with simulated data to illustrate how the model works. We generate data with a linear trend according to Equation 2, as well as data with seasonality according to Equation 3. Then we test the model on this data with and without covariates, and give a brief interpretation of the results.

4.1 Generated Data

4.1.1 Linear trend

We generated the data with linear trend according to Equation 2 for 1000 time steps with the following parameters: $\mu_0 = 67, \delta_0 = 5, \epsilon_t \sim \mathcal{N}(0, 5), \eta_t \sim \mathcal{N}(0, 500), \xi_t \sim \mathcal{N}(0.08, 1)$. [1] For our intervention we add an extra 2000 to every value of μ after the first 800. We map each time step to a day, which gives a range from 2019-01-01 to 2021-03-11 without intervention and 2021-03-12 to 2021-09-27 after the intervention. Next we fit a BSTS model on this data and compute the counterfactual using the time series without intervention as a covariate. The results can be seen in 1.

[1] Notice that the 0.08 in $\xi_t \sim \mathcal{N}(0.08, 1)$ controls how strong the linear trend is. If we decrease this parameter from 0.08 to 0.01 we get more erratic behaviour.

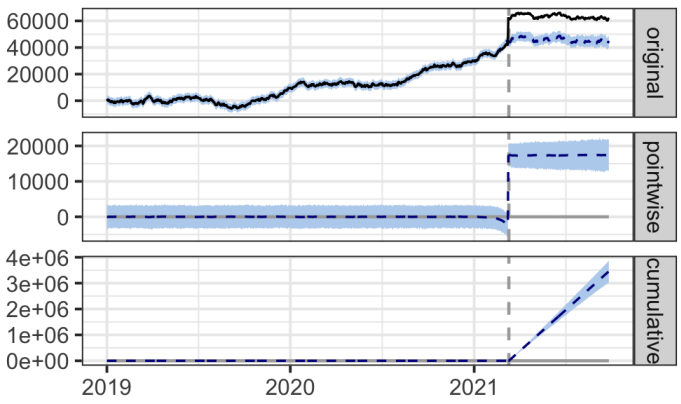

Figure 1: Visualization of the model predictions for the counterfactual with 95% credible intervals, the pointwise difference, and the cumulative difference.

In the uppermost plot, the dashed line denotes the fitted synthetic counterfactual, the solid line denotes the actual data for the treatment group. The other plots show the difference between them.

During the post-intervention period, the response variable's average value was around 63.29K. In contrast, without intervention, our model expects an average response of 45.92K. The 95% interval of this counterfactual prediction is [43.93K, 48.13K]. Subtracting the counterfactual prediction from the observed response gives the estimate of the causal effect the intervention on the response variable, which was 17.37K with a 95% interval of [15.17K, 19.37K]. Given that this was simulated data, we know that the effect of the intervention was 20K which falls outside of the 95% credible interval by around 700 units. Thus, although not impossible, the model considers the correct value of the intervention to be unlikely, showing that there is still room for misjudgement.

In cumulative terms, if we sum up the individual data points during the post-intervention period, the overall value of the response variable was 12.60M whilst that of the counterfactual prediction was 9.14M, with [8.74M, 9.58M] as the 95% credible interval. Thus, the response variable showed an increase of +38% with a 95% interval of [+33%, +42%]. The model considers this a statistically significant effect with a Bayesian one-sided tail-area probability p = 0.001, which is correct since there was indeed an intervention.

4.1.2 Seasonality

In a similar manner, we generate data for seasonality (7 seasons) according to Equation 3 for 100 time steps with the following parameters: $\mu_0 = 67, \gamma_0 = 5, \delta_0 = (-10, 15, -5, 50, -20, -30, 0), \epsilon \sim \mathcal{N}(0, 1), \eta_t \sim \mathcal{N}(0, 1), \xi_t \sim \mathcal{N}(0.08, 1), \omega_t \sim \mathcal{N}(0, 1)$. Adding an intervention at time step 80 of an extra 200 to the observations and mapping the time to dates yields the following results:

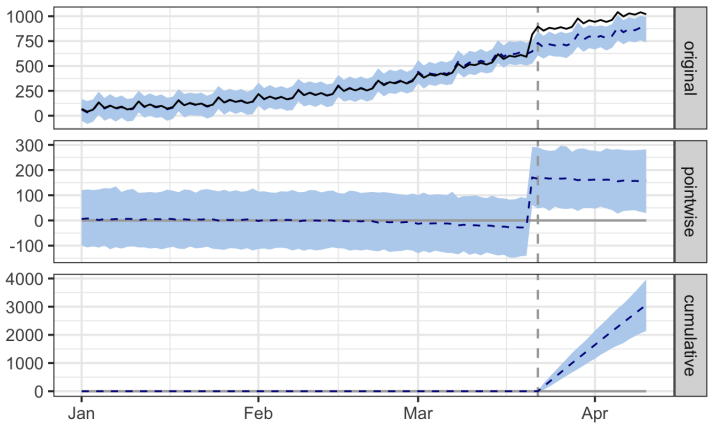

Figure 2: Visualization of the model predictions for the counterfactual with 95% credible intervals, the pointwise difference, and the cumulative difference for data with seasonality.
In the uppermost plot, the dashed line denotes the fitted synthetic counterfactual, the solid line denotes the actual data for the treatment group. The other plots show the difference between them.

During the post-intervention period we get an average value of 951.97 for the observed data, and 790.53 for the counterfactual with a 95% interval of [743.16, 839.45]. The causal effect of the intervention according to the model is 161.44 with a 95% interval of [112.52, 208.81]. So the true impact of the intervention (200) is contained within the interval. Thus the model seems to approximate the impact of the intervention closely.

It can be seen that BSTS are indeed a very flexible model, and that the inclusion of covariates indeed helps the prediction, as can be seen if we do the analysis taking a model that does not include a regression. In that case, we very quickly loose confidence in the prediction. For our linear trend, if we don't use regressions we get the results (predicting 200 time steps) in Figure 3.

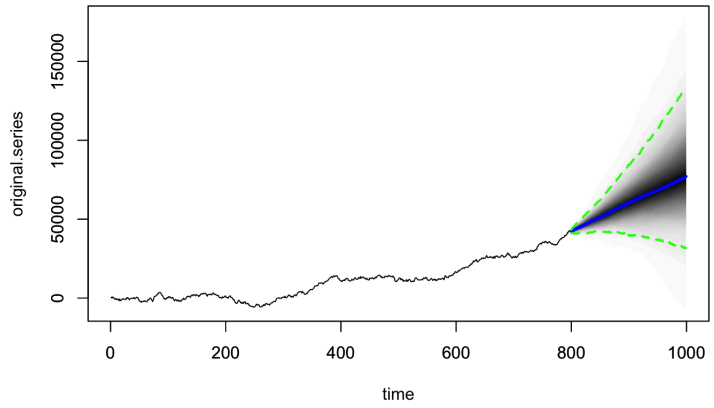

Figure 3: Visualization of the forecast from the model with no regression. 95% confidence intervals in green.

4.1.3 Inclusion probabilities of covariates

Spike-and-Slab prior allows us to look at which covariates had more influence in the prediction. For this experiment we created three additional dummy covariates and plotted their inclusion probabilities in the end. We are using the same target series and covariate data that was generated in section 4.1. The three dummy covariates are generated as follows

- True predictor(cov) as generated in section 4.1
- First dummy covariate(dummy1) has the value of its index. It is simply a y=x line.
- Second dummy covariate(dummy2) is a sine curve on the index.
- Third dummy covariate(dummy3) is gaussian noise with mean 0 and variance 1

Figure 4 shows the inclusion probabilities of the following covariates.

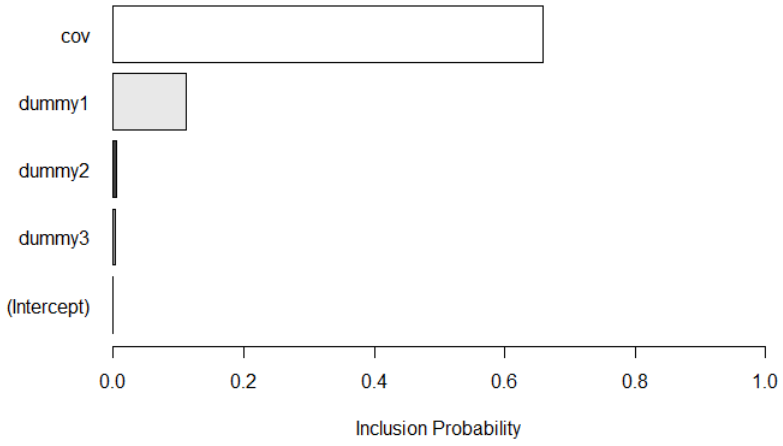

Figure 4: Inclusion probabilities of covariates

It can be seen that the actual covariate has the highest probability. dummy1 also is used in prediction because the target series has a trend component. It also moves similarly with our dummy1 as it is also a time series with smooth trend and zero noise. The other simulated covariates as expected contribute minimally to the prediction of our target series.

5 Real World Data

5.1 Randomized Controlled Trial

Randomized controlled trials (RCT) are considered to be the golden standard for causal inference. Through the construction of a randomly assigned control group and intervention group (in a simplistic setup), it is assumed that all other confounding factors will be evenly distributed amongst the two groups, thereby negating their effect [5]. We test the *CausalImpact* package with RCT data, to examine how the model's causal impact estimate compares to standard methodology.

The dataset chosen measures the adherence to antiretroviral (ART) therapy in Tanzania for pregnant HIV-positive women. The intervention consisted of paper-based appointment tracking and community outreach systems to decrease the rate of missed appointments. This dataset was chosen as it has many

timepoints (data collected frequently over 2 years) and in addition, has numerous datapoints before the intervention started (a pre-intervention period) which many RCT datasets lack [11].

Ross-Degnan and his colleagues applied interrupted time series segmented regression models to compare differences in monthly rates between the intervention and control group (fig 5). In particular, they looked at differences 6 months after the start of the intervention (chosen as strong differences were seen). Their net estimated differences in outcomes was a decrease of 13.7% (95% CI (-15.4, -12.1)) [11] (table 1).

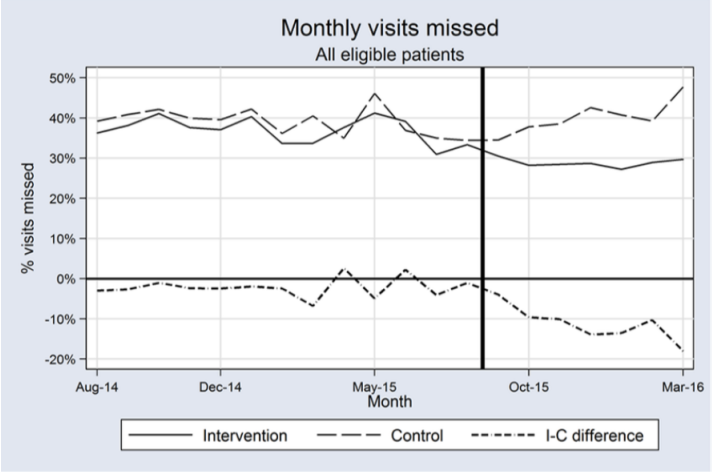

Figure 5: Original analysis of the rate of missed visits between intervention and control group [11]

	Absolute effect	CI.lower	CI.upper	p.value
ITS	-13.7%	-15.4%	-12.1%	0.001
BSTS	-8.5%	-12%	-4.6%	0.001

Table 1: Impact estimation using interrupted time series model (ITS) and *CausalImpact* (BSTS)

The *CausalImpact* model was applied to the same dataset. For this purpose, data is aggregated per month and the percentage of missed meetings is mapped from 0 to 1, 0 representing that no meeting was missed and 1 representing that 100% of meetings were missed. The counterfactuals are composed of joint estimation based on the pre-intervention period of the treated group and the control group as predictive covariate series (fig 6). In the figure, first the original data is represented with the intervention as a solid black line steadily decreasing, and the counterfactuals in a dotted blue line remaining approximately stable with some fluctuation. The second plot, represents the pointwise comparison of the intervention to the counterfactuals, where a marked dip is seen post-intervention. Thirdly, we see the cumulative effect of the intervention. As time progresses, the sum of the pointwise comparisons, shows a strong difference in the intervention to counterfactuals.

Comparing the *CausalImpact* figure to figure 5, both resemble one another. Similarly to the original RCT analysis, the rate of missed appointments significantly decreased in the intervention group. The intervals observed are larger but this is due to the fact that *CausalImpact* uses credible intervals instead of confidence intervals.

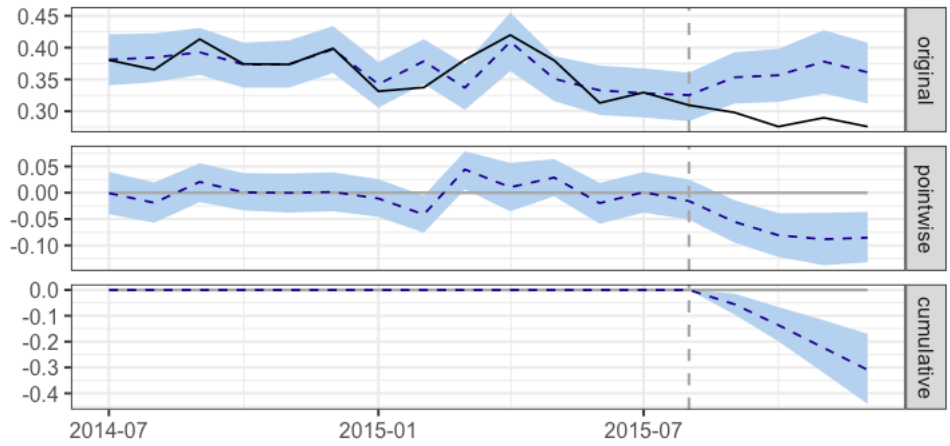

Figure 6: *CausalImpact* analysis of missed visits including the control group as covariate series
The x-axis shows missed meetings as fraction of all meetings, the y-axis shows time passed. In the uppermost plot, the dashed line denotes the fitted synthetic counterfactual, the solid line denotes the actual data for the treatment group. The other plots show the difference between them.

As *CausalImpact* is looking at the full data set rather than an estimate at a specific time point, the effect was seen as less pronounced at -8.5% than the original measurements. In figure 6, it seems that the effects of the intervention start building up slowly before 2016-01 and fade away somewhat after. Another reason for *CausalImpact* estimating smaller overall effect is the fact that there is a slight decrease in the treated group already before the intervention takes place. It might thus be the case that, similar to the case with synthetic data in 4.1.1, the model's great flexibility reduces its accuracy to some extent. Despite the difference in magnitude, *CausalImpact* is able to confirm the findings of the RCT successfully and finds a significant negative impact of the intervention on missed antiretroviral treatment appointments.

5.2 Comparison to Interactive Fixed Effects on Observational Data

In this section we will compare some of the results obtained in [14] using an *Interactive Fixed Effects* (IFE) model with the results given by *CausalImpact* when run on the same data set. Specifically, we will estimate the causal impact on turnout of a policy change from mandatory voter registration to no voter registration requirement in the states of Maine, Minnesota and Wisconsin in the year 1976. For this purpose we will use time series data on US federal election voter turn outs on the state level from 1920 to 1992. During this time period, no other states introduce the same measures, and the voter turnout in the remaining US states will be used as covariates. This analysis differs slightly from the original one available in [14] since we do not look at the impact of all such laws introduced, but only at this specific subset. We use this setup, because *CausalImpact* relies on contemporaneous covariates. This assumption would be violated if we were to analyze effects that set on at different times as done in the original study.

5.2.1 Interactive Fixed Effect Model Estimation

As a first step, we customize the original analysis to be able to deal with the altered data set and run it with otherwise unchanged parameters. The model therefore becomes:

$$turnout_{it} = X_{it}\hat{\beta} + \hat{\lambda}_i' f_t$$

where $turnout_{i_t}$ is the voter turnout in state i at time t, X_{it} describes characteristics of state i at time t (in our case whether or not a no-registration policy has been enabled), and f_t captures the values of unobservable covariates for the treatment group. $\hat{\beta}$ and $\hat{\lambda}$ are estimated on the control group only, λ_i is estimated on the treatment group for all pre-treatment periods. The number of unobservable covariates

is determined in a cross-validation procedure. For more details see the online appendix of [14]. The results can be seen in fig 7. In line with the original findings, and as hypothesized by the author, who expects this subset to exhibit the strongest causal impact, we see a clear change of trends at beginning of the intervention. The average treatment effect on the treated (ATT) is positive significantly different from zero accordingly.

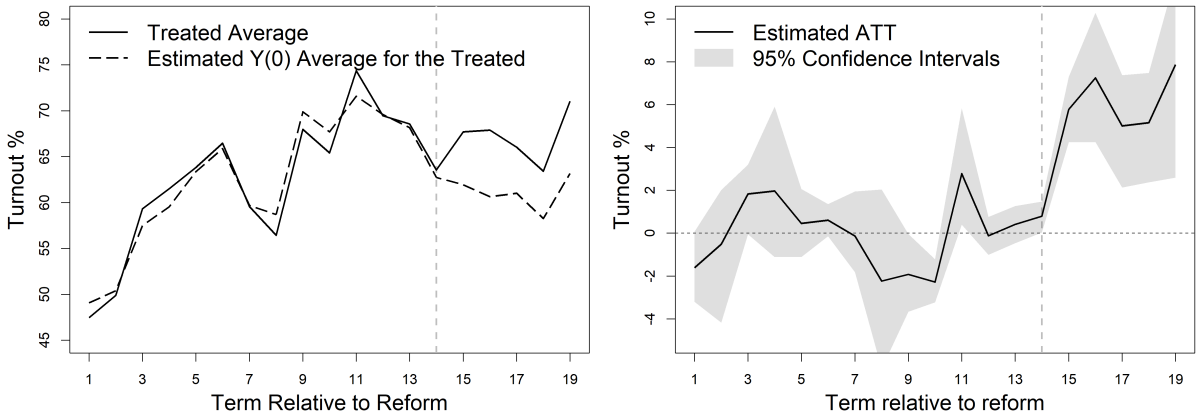

Figure 7: Synthetic controls using an interactive fixed effect model

5.2.2 CausalImpact Estimation

In a next step, we feed the same data set to CausalImpact. In the upper row of figure 8 we can see that the remaining US states make for highly predictive covariates. The second and third row in figure 8 show that the causal impact as estimated by *CausalImpact* is not quite as big as the one in figure 7. The intervals also appear wider, but this comes as no surprise given that they indicate credible intervals rather than confidence intervals.

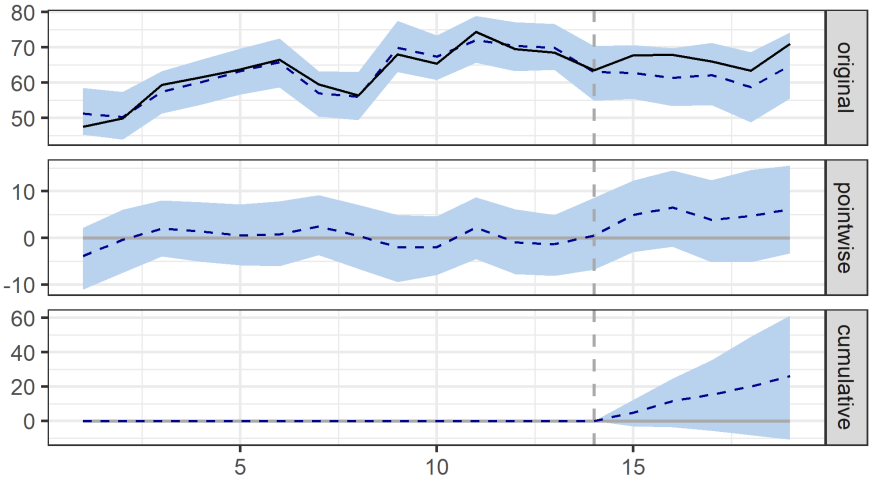

Figure 8: Synthetic controls using a bayesian state-space model
The x-axis shows average voter turnout, the y-axis shows time in election cycles. In the uppermost plot, the dashed line denotes the fitted synthetic counterfactual, the solid line denotes the actual data for the treatment group. The other plots show the difference between them.

5.2.3 Comparison

For the IFE, we seem to find a clear causal impact of the policy intervention. According to the *CausalImpact* analysis, the impact is similar in magnitude, but markedly less significant due to the fact that we are using credible intervals.

	Absolute Effect	S.E.	CI.lower	CI.upper	p.value
IFE	6.22	1.37	3.29	8.50	0.00
BSTS	5.39	3.60	-1.90	12	0.08

Table 2: Impact estimation using IFE vs. BSTS (CausalImpact)

First and foremost, it is worth noting, that both methods find a causal impact of similar magnitude. The difference in significance is mostly due to the fact that the models use a different notion of significance and is therefore hard to compare directly. A very interesting aspect of the comparison are potential covariates that may confound a causal interpretation of the impact. While *CausalImpact* allows for time-varying local trends in the latent states of the state-space model, it assumes that the model structure otherwise remains intact throughout the period. The accuracy of the model structure in turn implicitly relies on the notion that the covariates provided capture any general trends that also apply to the treatment group. In case of a lack of suitable covariates, the model can only use local trends to make up for otherwise explainable variation. In our case, the covariates are unlikely to be of the same quality as in a randomized control trial, since policy change and policy effect are very likely to be simultaneous.

Furthermore, the results obtained by [14] indicate that there may have been some structural changes during the pre-treatment period which the combination of local linear trend and static regression coefficients can likely not capture. As in the original study, our interactive fixed effects model provides best results when equipped with $r = 2$ unobserved factors, which appear to capture the removal of Jim Crow laws in southern states and the onset of a long-term negative trend in voter turnout around the year 1965. The bayesian state-space model, for all its flexibility, shows no hint of capturing these effects in an interpretable manner. Conversely, the interactive fixed effects model does capture them, but all while carrying the implicit assumption that these are factors to be controlled for when estimating the causal impact which, depending on the actual causal chain, may itself introduce spurious correlation.

Ultimately, we are left with the choice between *CausalImpact* which may mistake structural changes for local trends and auto-correlation, and *Interactive Fixed Effects*, which is better suited to capturing structural changes but may use them in all the wrong ways. While both shortcomings show the caveats of estimating causal impact without a causal diagram or a randomized control trial, they do end up with remarkable similar results and may just be the best chance we have at causal inference in real world scenarios when applied to an appropriate setting.

6 Conclusion

We successfully applied the model proposed by [6] to synthetic data and new data sources. In terms of synthetic data we found the flexibility of the model to be somewhat limiting to how accurately it is capable of estimating the counterfactual, as was expected. Finding the optimal tradeoff of model flexibility and prediction accuracy therefore remains a challenge. Nonetheless, the model came close to estimating the true values of the intervention and was capable of handling different structures in the data.

In terms of application to real world data, the inherent flexibility of the BSTS based model proves capable of estimating the causal impact measured in a randomized controlled trial. Furthermore, it obtains very similar results to an interactive fixed effects model when applied to the same data, all without the explicit assumption of unobservable fixed effects. This indicates that, given the right predictive covariates, the expressivity of the approach is sufficient for modeling complex relationships. All these results have been obtained still without making use of dynamic coefficients and using very specific sets

of predictive time series as well as off-the-shelf priors. A more detailed study of model behaviour given different initialisations and varying degrees of predictiveness in the covariate time series given to the model would help further illuminate the potential of the approach. In the cases examined in this report, *CausalImpact* has shown itself a reliable tool for causal impact estimation in real world scenarios, provided a set of suitable predictive time series.

References

[1] Alberto Abadie. "Semiparametric difference-in-differences estimators". In: *The Review of Economic Studies* 72.1 (2005), pp. 1–19.

[2] Alberto Abadie, Alexis Diamond, and Jens Hainmueller. "Comparative politics and the synthetic control method". In: *American Journal of Political Science* 59.2 (2015), pp. 495–510.

[3] Sudipto Banerjee, Robert J Kauffman, and Bin Wang. "Modeling Internet firm survival using Bayesian dynamic models with time-varying coefficients". In: *Electronic Commerce Research and Applications* 6.3 (2007), pp. 332–342.

[4] Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan. "How much should we trust differences-in-differences estimates?" In: *The Quarterly journal of economics* 119.1 (2004), pp. 249–275.

[5] Laura E Bothwell et al. "Assessing the gold standard—lessons from the history of RCTs". In: *N Engl J Med* 374.22 (2016), pp. 2175–2181.

[6] Kay H Brodersen et al. "Inferring causal impact using Bayesian structural time-series models". In: *The Annals of Applied Statistics* 9.1 (2015), pp. 247–274.

[7] Frederick Eberhardt and Richard Scheines. "Interventions and causal inference". In: *Philosophy of Science* 74.5 (2007), pp. 981–995.

[8] Ronald Aylmer Fisher et al. "Statistical methods for research workers." In: *Statistical methods for research workers.* 5th Ed (1934).

[9] Hemant Ishwaran, J Sunil Rao, et al. "Spike and slab variable selection: frequentist and Bayesian strategies". In: *The Annals of Statistics* 33.2 (2005), pp. 730–773.

[10] Judea Pearl et al. "Causal inference in statistics: An overview". In: *Statistics surveys* 3 (2009), pp. 96–146.

[11] Dennis Ross-Degnan et al. *A group randomized trial using an appointment system to improve adherence to ART at reproductive and child health clinics implementing Option B+ in Tanzania*. Sept. 2017. DOI: 10.1371/journal.pone.0184591. URL: https://plos.figshare.com/articles/A_group_randomized_trial_using_an_appointment_system_to_improve_adherence_to_ART_at_reproductive_and_child_health_clinics_implementing_Option_B_in_Tanzania/5451829/1.

[12] Steven L Scott and Hal R Varian. "Predicting the present with bayesian structural time series". In: *Available at SSRN 2304426* (2013).

[13] Christof Seiler. *KEN4258: Computational Statistics - Topic 4: Markov Chain Monte Carlo*. Feb. 2020.

[14] Yiqing Xu. "Generalized synthetic control method: Causal inference with interactive fixed effects models". In: *Political Analysis* 25.1 (2017), pp. 57–76.

7 Appendix

7.1 Reproduction of the Paper

To begin the project, we recreated the analysis done by the authors of [6]. Here we provide the results of such recreation.

We simulate a time series Y that is related to two sinusoidal covariates C_1, C_2, a the local level value and two regression coefficients β_1, β_2, by the following equation:

$$Y_t = 20 + linearValues + \mathcal{N}(0, 0.1) + (C_1)(\beta_1) + (C_2)(\beta_2) \tag{11}$$

where the elements evolve as follows:

$$linearValues_t = \mathcal{N}(linearValues_{t-1}, 0.1) \tag{12}$$

$$\beta_{1t} = \mathcal{N}(\beta_{1,t-1}, 0.1) \tag{13}$$

$$\beta_{2t} = \mathcal{N}(\beta_{2,t-1}, 0.1) \tag{14}$$

$$C_{1t} = \sin \frac{\pi(t)}{45} \tag{15}$$

$$C_{2t} = \sin \frac{\pi(t)}{180} \tag{16}$$

The generated series Y can be seen in Figure 9.

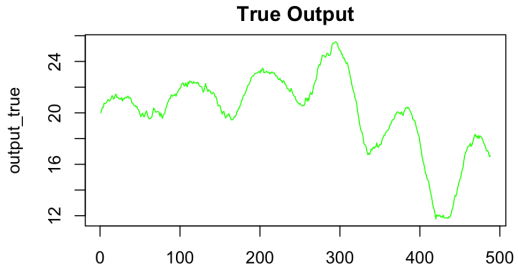

Figure 9: Simulated time series

An intervention was added at 389 days by multiplying it with e.

Applying the model yields the following results:

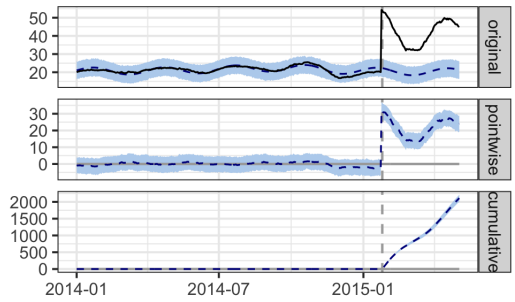

Figure 10: Causal impact for data generated above

Notice that using the covariates gives a very narrow 95% interval.