

Policy Evaluation using a Hierarchical Bayesian Structural Time Series Model

Xuewen Yu

London School of Economics and Political Science



EPITOME

Evaluating Policy Implementation TO Predict MEntal health: a Bayesian hierarchical framework for quasi-experimental designs in longitudinal settings

- ▶ WP1: hierarchical statistical framework
 - ▶ Core Bayesian time series model with spatial/temporal dependency, extended to stepped wedge and matched-control designs
 - ▶ Simulation study to assess robustness and sensitivity to model and prior choices
- ▶ WP2: evaluating impact on mental health need in England
 - ▶ Apply the framework to five UK datasets to estimate effects of Universal Credit and the Hostile Environment Policy
 - ▶ Test for independent and combined effects across socioeconomic and ethnic minority groups
- ▶ WP3: developing alternative controls
 - ▶ Develop Bayesian synthetic control and negative outcome control methods for cases with no standard control group
 - ▶ Apply these to the WP2 case studies and extend simulations to guide their use
- ▶ WP4: economic evaluation
 - ▶ Extend existing forecasting models into a Bayesian framework, propagating uncertainty from WP1 to WP3
 - ▶ Use multi-state modelling and potentially Value of Information analysis to assess service demand and decision adequacy

Background: the policy

- ▶ 'Windrush scandal'[1]
 - ▶ The UK government introduced the hostile environment policy in 2012, implementing a series of measures to inspect immigration documents and limit undocumented immigrants to gain access of public services.
 - ▶ The 'Windrush' generation who arrived in the UK from Caribbean countries between 1948 and 1973: many were wrongly classified as illegal immigrants due to the destruction of landing cards.
 - ▶ The Immigration Act 2014 enforces key aspects of the hostile environment policy.
 - ▶ The Windrush scandal began to surface in 2017.



The policy

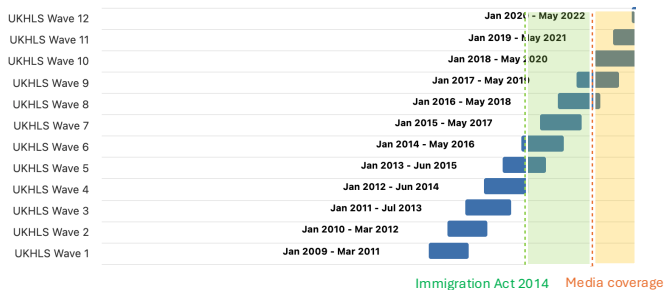


Figure 1: Timeline. Data are from the UK household longitudinal study (UKHLS) from Understanding Society.

Find the impact: frame the problem

- ▶ Goal: evaluate the effect of Immigration Act 2014 on mental health of Black Caribbeans
- ▶ Study period: from May 2009 to November 2017 (Windrush scandal media coverage).
- ▶ **Treated group:** Black Caribbean
- ▶ **Candidate control groups:** White British, White Irish, White Gypsy or Irish traveller, Other White, Mixed Asian, Other Mixed, Indian, Pakistani, Bangaldesh, Chinese, Other Asian, Black African, Other Black, Arab, Other
- ▶ Outcome: 12-item General Health Questionnaire (GHQ-12). The score varies from 0 to 36. Lower score on the 0-36 scale is associated with better mental health.
- ▶ Inclusion: We included all participants aged 16 years and older who had complete mental health data for our outcome of interest from at least one of the waves between the study period.

Overview

Find the impact of **an intervention** from an **individual-level longitudinal** dataset using a **hierarchical Bayesian Structural Time Series (HBSTS)** model:

- ▶ Review the Bayesian Structural Time Series (BSTS) model
- ▶ Define the HBSTS model
- ▶ A simulation study
- ▶ The empirical study for evaluating the impact of the hostile environment policy

Find the impact: existing approaches

Bayesian Interrupted Time Series (BITS)

- Comparing change of mental wellbeing in Black Caribbean group with change of mental wellbeing in White ethnic group
- Finding **immediate change** at the onset of the policy and **rate** of change in the following time period

Bayesian Structural Time Series (BSTS)

- Comparing mental wellbeing in Black Caribbean group with **the mental wellbeing of them had there been no intervention (the counterfactual)**
- **Covariate-treatment independence:** the covariates used to estimate the outcome when there is no intervention should not be affected by the intervention
- The counterfactual can be estimated using other ethnic groups who were not affected by the policy (**synthetic control groups**)
- Estimating effect **at every time point** and **the average effect in the whole post-intervention period**

Infer causal impacts using a BSTS model

- Idea: mimic the pre-intervention behaviour of the treat unit – using latent trend and time-varying covariates $\mathbf{x}_t = (x_{1,t}, \dots, x_{H,t})$ – and then predict the counterfactual of the treat unit for the post-intervention period

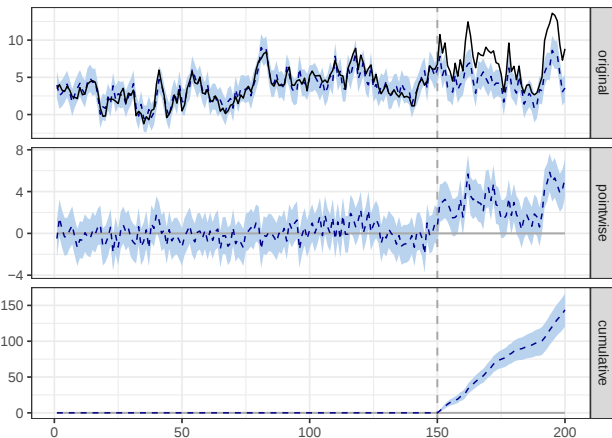


Figure 2: An example of applying the BSTS model to estimate causal impacts of an intervention happened at $t^* = 150$.

Find the true impact: existing approaches

Challenges of BSTS

- It's not designed for this type of data: need to combine all individual data into a single average at every timestamp — ignoring variation
- Specify the exact number of groups used to predict the counterfactual of the Black Caribbean group — hard to elicit and BSTS is **sensitive** to it

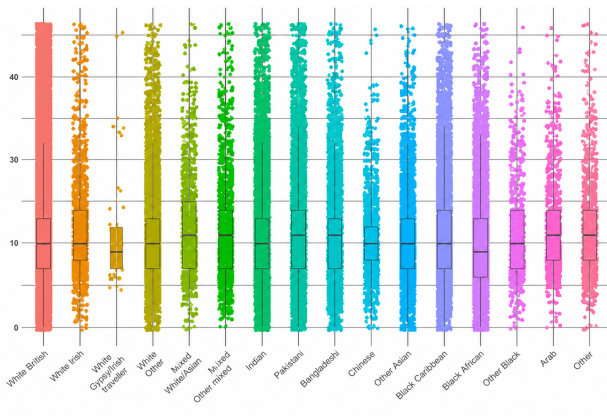


Figure 3: Distribution of GHQ-12 scores for different ethnic groups.

Our solution: A more realistic and robust model

- ▶ The extended model: we developed a hierarchical Bayesian structural time series (**HBSTS**) model.
- ▶ Novelty:

Old (BSTS)

Aggregation



New (HBSTS)

Adding a hierarchical layer to account for individual-level data

Sensitivity



Replacing the prior used for variable selection

A simulation study

1. Robustness to expected model size
2. Power of detecting a causal effect

Data Analysis: validate control groups

Some of the candidate control groups were exposed to the policy.

We examined whether they were not affected by the policy using BITS, formula see [2].

No evidence was found that these groups were affected by the intervention.

Data Analysis: The impact on mental health

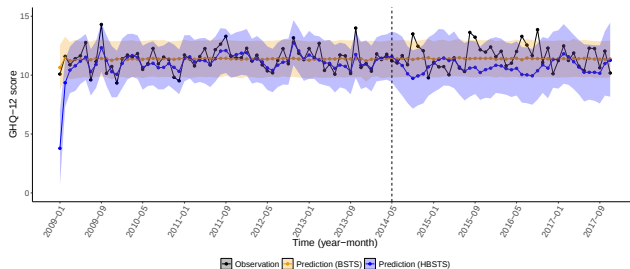


Figure 4: Posterior predictive samples from different models.

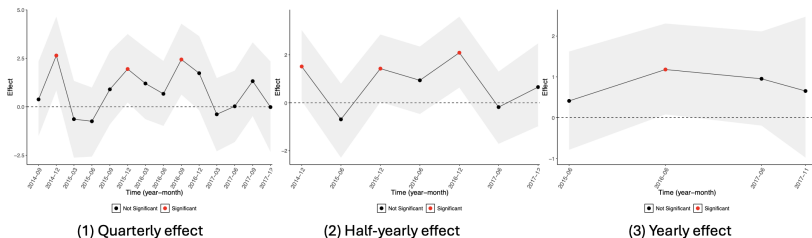


Figure 5: Pointwise effects

The key findings

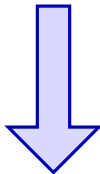
- ▶ Simulation studies:
 1. Our model showed more robustness against misspecification of the number of groups to be selected for predicting the counterfactual;
 2. Our model better detected the effect in various hypothesized scenarios.
- ▶ Real-world analysis:
 1. We found increased psychological distress among individuals from Black Caribbean background since the implementation of the Immigration Act 2014.

Limitations/Future work

1. (H)BSTS only evaluates the impact of a single intervention.
2. Spatial-temporal dependence was not considered.
3. What if control groups are not available?

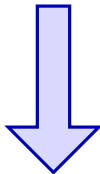
Can we use **covariates other than the control group's outcomes** to predict the counterfactual outcome for the treated group?

Covariate-treatment independence assumption



Synthetic control

Covariate-treatment independence assumption



Negative control

Thank you!

References I

- [1] Windrush scandal explained, 1999.
- [2] A. Jeffery, C. Gascoigne, J. Dykxhoorn, M. Blangiardo, S. Geneletti, G. Baio, and J. B. Kirkbride.
The effect of immigration policy reform on mental health in people from minoritised ethnic groups in england: an interrupted time series analysis of longitudinal data from the uk household longitudinal study cohort.
The Lancet Psychiatry, 11(3):183–192, 2024.