Evaluating mask policy of the Korean government using Bayesian Structural Time Series Model

SEOKMIN KIM*, MINSOO KIM, KYOUNGJAE LEE

Department of Statistics, INHA University

Introduction

- From the beginning of this year, as Covid-19 became a pandemic disease, public's interest in mask supply increased. In addition, anxiety of the mask supply increased because supply became unstable with a sharp increase in demand.
- In response, the government enforced mask-related policies, such as limiting the number of purchases per time at pharmacy(Feb. 27) and five-day rotation system based on birth year(Mar. 6).
- In this study, using Bayesian Structural Time Series, we confirm how the mask policy had affected the public's anxiety over the mask supply and how the policy impact had evolved over time.

Methodology

<Bayesian Structural Time Series model>

$$y_t = Z_t^T \alpha_t + \beta^T x_t + \epsilon_t \qquad \epsilon_t \sim N(0, H_t) \qquad (1)$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \qquad \eta_t \sim N(0, Q_t) \qquad (2)$$

- Equation (1) is called observation equation with observed data y_t and latent variables called state α_t. And equation (2) is called transition equation because it defines how the latent state evolves over time. One of the structural time series model, basic structural model can be represented as below [1].

$$y_t = \underbrace{\mu_t}_{trend} + \underbrace{\tau_t}_{seasonality} + \beta^T x_t + \epsilon_t \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2)$$

$$\mu_t = \mu_{t-1} + \delta_{t-1} + \upsilon_t \qquad \upsilon_t \sim N(0, \sigma_\upsilon^2)$$

$$\delta_t = \delta_{t-1} + \nu_t \qquad \nu_t \sim N(0, \sigma_\nu^2)$$

$$\tau_t = -\sum_{s=1}^{S-1} \tau_{t-s} + \omega_t \qquad \omega_t \sim N(0, \sigma_\omega^2)$$

<Prior of the parameters>

$$\gamma \sim \prod_{k=1}^{K} \pi_k^{\gamma_k} (1-\pi_k)^{1-\gamma_k} \qquad \beta_\gamma | \sigma_\epsilon^2, \gamma \sim N\left(b_\gamma, \sigma_\epsilon^2 \left(\Omega_\gamma^{-1}\right)^{-1}\right) \qquad 1/\sigma_\epsilon^2 \sim Ga\left(\frac{v_\epsilon}{2}, \frac{s_\epsilon}{2}\right)$$

$$1/\sigma_\upsilon^2 \sim Ga(10^{-2}, 10^{-6} s_y^2) \qquad 1/\sigma_\nu^2 \sim Ga(10^{-2}, 10^{-6} s_y^2) \qquad 1/\sigma_\omega^2 \sim Ga(10^{-2}, 10^{-6} s_y^2)$$

where $s_y^2 = \sum_{t=1}^{n} \frac{(y_t - \bar{y})^2}{n-1}$ $\quad s_\epsilon = (1-R^2) s_y^2 v_\epsilon$ $\quad \Omega_\gamma^{-1} = \kappa(\omega X_\gamma^T X_\gamma + (1-\omega) diag(X_\gamma^T X_\gamma))/n$

<Parameter learning>

- The θ denotes the set of model parameters σ_υ^2, σ_ν^2 and σ_ω^2.

1. Simulate the latent state α from $p(\alpha | \theta, \beta, \sigma_\epsilon^2, y)$ using Durbin and Koopman method.
2. Simulate $\theta \sim p(\theta | \alpha, \beta, \sigma_\epsilon^2, y)$. <Inverse gamma distribution>
3. Simulate β and σ_ϵ^2 from Markov chain with stationary distribution $p(\beta, \sigma_\epsilon^2 | \alpha, \theta, y)$ [1].

Step 3 is specifically shown in **<Gibbs sampling process of β and σ_ϵ^2>**.

<Gibbs sampling process of β and σ_ϵ^2>

$y^* = (y_1^*, \cdots, y_n^*)$ $\qquad y_t^* = y_t - Z_t^T \alpha_t$ from observation equation (1).

1. $\gamma | y^* \sim C(y^*) \frac{|\Omega_\gamma^{-1}|}{|V_\gamma^{-1}|^{\frac{1}{2}}} \frac{p(\gamma)}{SS_\gamma\left(\frac{N}{2}-1\right)}$
2. $1/\sigma_\epsilon^2 | \gamma, y^* \sim Ga\left(\frac{N}{2}, \frac{SS_r}{2}\right)$
3. $\beta_\gamma | \sigma_\epsilon^2, \gamma, y^* \sim N\left(\tilde{\beta}_r, \sigma_\epsilon^2 \left(V_\gamma^{-1}\right)^{-1}\right)$, β_γ denotes the subset of elements of β where $\beta_k \neq 0$.

where $V_\gamma^{-1} = (X^T X)_\gamma + \Omega_\gamma^{-1}$ $\qquad N = v_\epsilon + n$

$\widetilde{\beta_\gamma} = \left(V_\gamma^{-1}\right)^{-1} (X_\gamma^T y^* + \Omega_\gamma^{-1} b_\gamma)$ $\qquad SS_\gamma = s_\epsilon + y^{*T} y^* + b_\gamma^T \Omega_\gamma^{-1} b_\gamma - \widetilde{\beta_\gamma}^T V_\gamma^{-1} \widetilde{\beta_\gamma}$

<Forecasting>

- Let $y = (y_1, \cdots, y_n)$, $\tilde{y} = (y_{n+1}, \cdots, y_m)$ and $\phi = (\theta, \beta, \sigma_\epsilon^2, \alpha)$.
- Posterior predictive distribution of $\tilde{y}$ is $p(\tilde{y}|y) = \int p(\tilde{y}|\phi) p(\phi|y) d\phi$.
- $\phi^{(1)}, \phi^{(2)}, \cdots$ are a set of random draws from **<Parameter learning>**, then one samples from $p(\tilde{y}|y)$ by sampling from $p(\tilde{y}^{(g)} | \phi^{(g)})$, which is done by simply iterating equations (1) and (2) forward from $\alpha_n^{(g)}$, with parameters $\theta^{(g)}, \beta^{(g)}$ and $\sigma_\epsilon^{2(g)}$ [1].
- With $\tilde{y}^{(g)}$s, we can get posterior mean and posterior predictive interval.

<Estimating Causal Impact>

- First, using the response variables before intervention, we fit the Bayesian Structural Time Series model to obtain posterior samples for model parameters. And using these samples with **<Forecasting>** method, obtain samples from the posterior predictive distribution for the response variables up to the time point 'm' after the intervention.
- And then, with the difference between the samples and real values of the response variables, up to the point 'm' after the intervention, we can get the point estimates and credible intervals of the causal impact, cumulative causal impact and average causal impact for each time point [2].

$\varphi_t^{(i)} = y_t - \tilde{y}_t^{(i)}$, $\qquad \tilde{y}_t^{(i)}$ is i-th sample at time t from posterior predictive distribution.

Cumulative causal impact at time t $= \frac{1}{B} \sum_{i=1}^{B} \sum_{t'=n+1}^{t} \varphi_{t'}^{(i)}$ $\qquad$ B is the number of posterior samples

Average causal impact at time t $= \frac{1}{B} \sum_{i=1}^{B} \frac{1}{t-n} \sum_{t'=n+1}^{t} \varphi_{t'}^{(i)}$ $\qquad$ B is the number of posterior samples

<Control time series>

- When using Bayesian Structural Time Series to estimate the intervention impact, the time series of covariates x_t should not be affected by intervention. In addition, the relationship between covariates and response variable before intervention is equal to the relationship between covariates and the counterfactual response variable after intervention. A covariate that satisfies these conditions is called 'control' [2].

Real Data Analysis

- Data analysis proceeded as follows. First, we assumed mask search volume as a surrogate of anxiety about mask supply that is not observed, and we predicted the counterfactual mask search volume after government's mask policy. Then comparing it with the real value, the causal impact according to the time point was obtained. **Figure 1** is a DAG (Directed Acyclic Graph) indicating the relationship between variables.

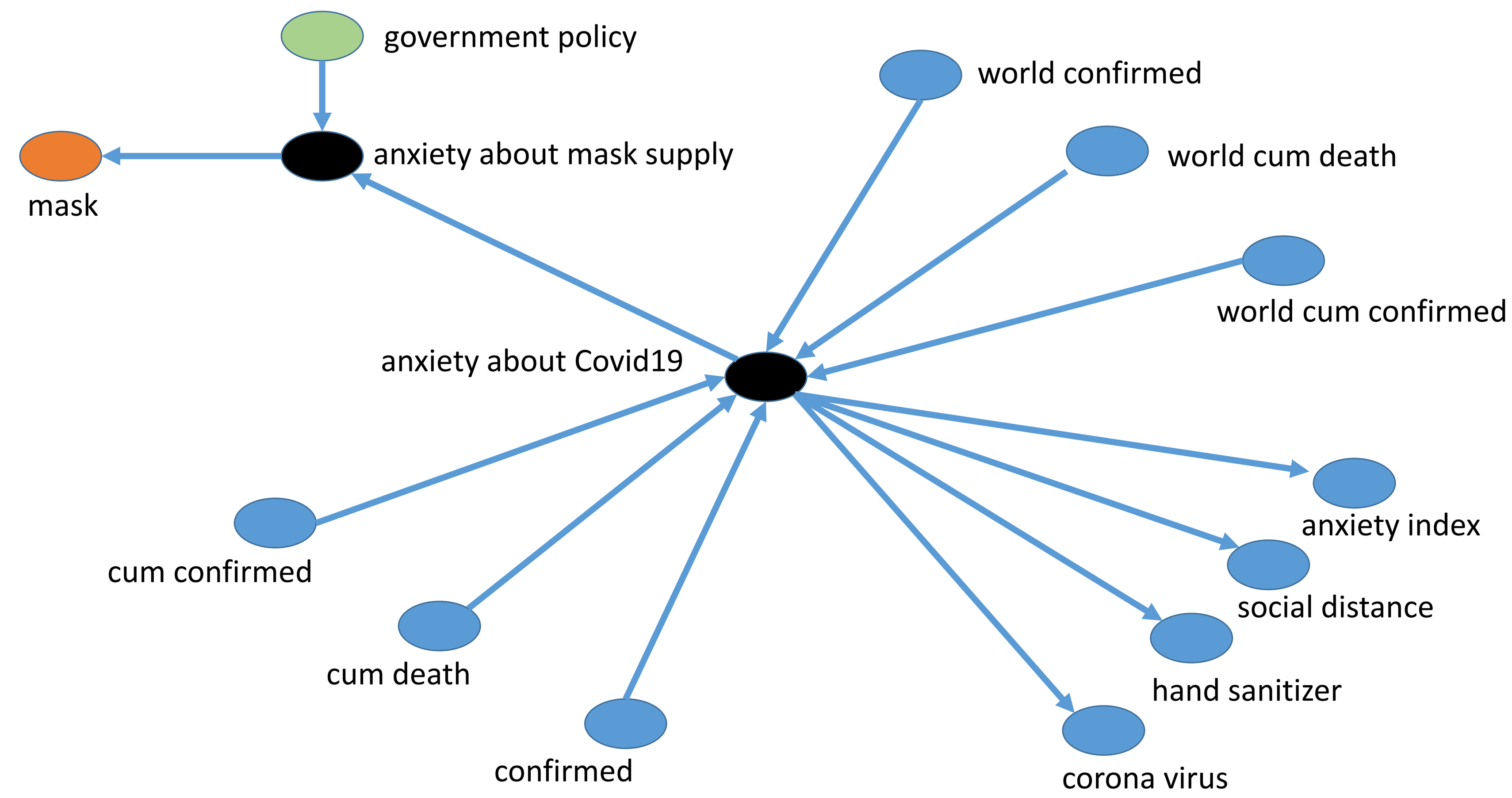

Figure 1. DAG for data analysis
mask : mask search volume (in NAVER) ; **government policy :** 0 before 2/27, 1 after 2/27 ; **cum confirmed :** accumulated number of confirmed cases a day earlier ; **cum death :** accumulated deaths a day before ; **confirmed :** number of confirmed cases a day before ; **corona virus :** coronavirus search volume (in NAVER) ; **hand sanitizer :** hand sanitizer search volume (in NAVER) ; **social distance :** social distance search volume (in NAVER) ; **anxiety index :** Covid-19 Anxiety index (from Software Policy & Research Institute) [4] ; **world cum confirmed :** accumulated confirmed cases worldwide a day ago ; **world cum death :** accumulated deaths worldwide a day ago ; **world confirmed :** number of confirmed cases worldwide a day ago ; **Black nodes are unobserved variables.**

- Variables used for data analysis were gathered from 'NAVER Trend', 'worldometer.info' and 'SPRI(Software Policy & Research Institute)'.
- In brief, relationship between variables through DAG, Covid-19 anxiety index, search volumes and number of confirmed or death cases are associated with mask search volume and independent with government policy(by d-separation, Markov assumption [3]). Therefore, these variables satisfy the condition of control.
- The data analysis result using Bayesian Structural Time Series is shown in **Figure 2**.

(prior setting : $R^2 = 0.5, v_\epsilon = 0.01, \pi_\kappa = 0.5, b_\gamma = 0, \kappa = 1, \omega = 0.5$)

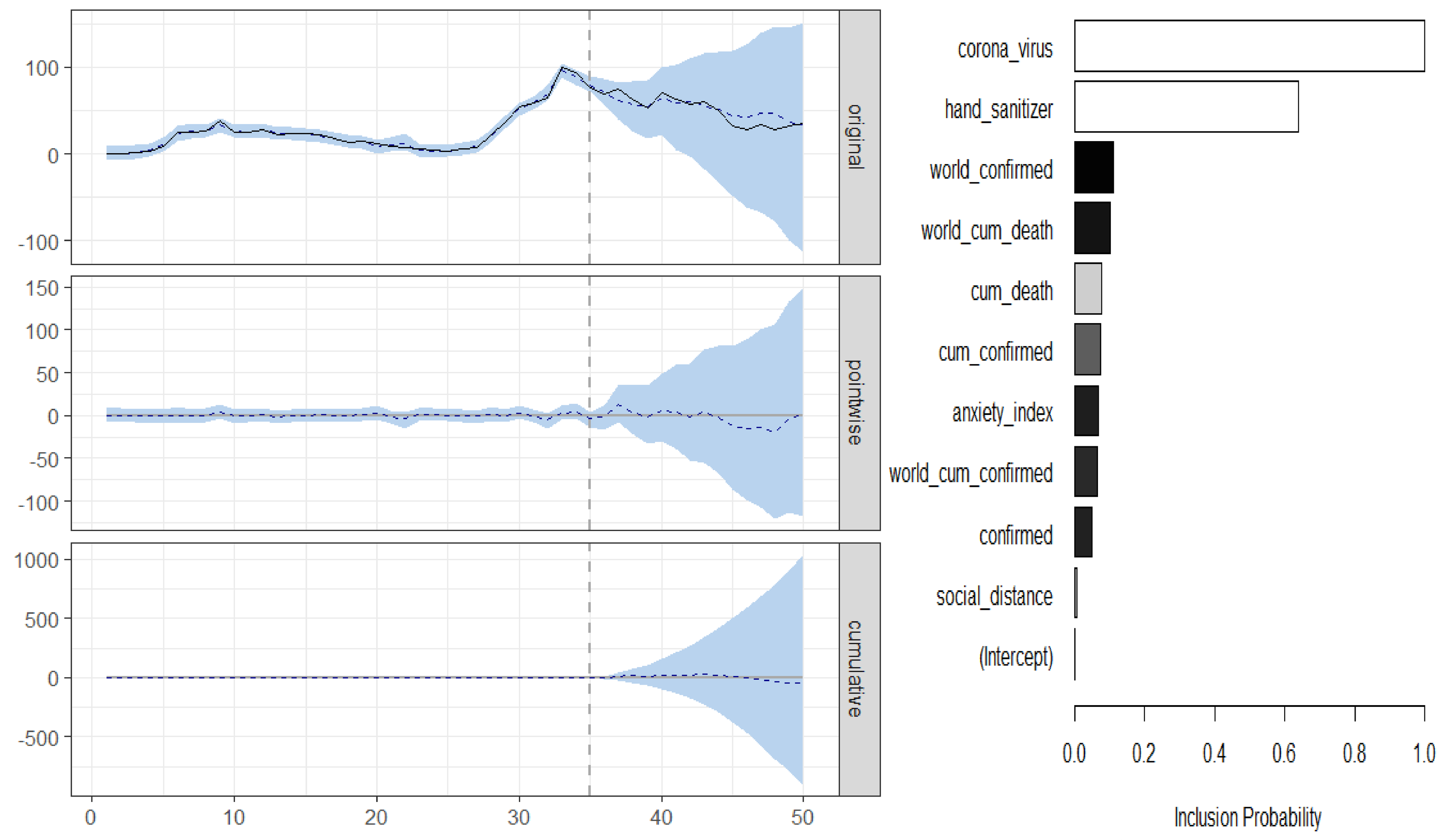

Figure 2. Result of Data analysis. time series plot of causal impact of government policy(left) and posterior inclusion probability plot of BSTS model(right).

- The left plot of **Figure 2**, the vertical dotted line indicates the time when the policy was enforced. And from top to bottom, the plot shows a time series plot of the mask search volume(solid line is a real value, dotted line is a predicted value and shaded area is 95% credible interval), causal impact of the mask policy and cumulative causal impact of the mask policy, respectively.
- In the right plot of **Figure 2**, coronavirus search volume contributed the most to the prediction of mask search volume followed by hand sanitizer search volume.
- In the left plot of **Figure 2**, It seems like that the causal impact of the mask policy exists after March 6th when the 2nd policy was announced, 8 days after the 1st policy was enforced.
- It means that the 1st mask policy hardly decreased the public's anxiety about the mask supply, but the 2nd mask policy, the five-day rotation system, decreased the anxiety a little.
- However, the impact would not be considered statistically significant.(Point estimate of Relative cumulative impact is -5%, but 95% Credible Interval is [-114%, 132%]).

References

- [1] Steven L. Scott and Hal Varian (2013). Predicting the Present with Bayesian Structural Time Series. *International Journal of Mathematical Modelling and Numerical Optimisation*, **5,** pp. 4-23
- [2] Kay H. Brodersen, 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), 247-274.
- [3] Miguel A. Hernán, James M. Robins (2019), Causal Inference: What If. *CRC Press.*
- [4]이승환(2020). 코로나19 불안지수: 감성분석과 의미., *월간 SW중심사회 6월호*

SEOKMIN KIM : bear900410@gmail.com, MINSOO KIM : pos02043@gmail.com, KYOUNGJAE LEE : leekjstat@inha.ac.kr
회의 ID : 885 8041 4381
초대 링크 : https://us02web.zoom.us/j/88580414381?pwd=Zk9HRnliU1REanJ5MmhEYXZoNjVPZz09