

# Quantifying Regional Variation in Ivermectin Use During the Pandemic using Regularized Synthetic Controls

Dinesh Puranam,<sup>1</sup> Rashmi Ranjan Bhuyan,<sup>1</sup> Ivan Belov,<sup>1</sup> Shantanu Dutta,<sup>1</sup> Jeroen van Meijgaard,<sup>2</sup> Reetabrata Mookherjee<sup>2</sup> and Gourab Mukherjee<sup>1</sup> 

<sup>1</sup>Marshall School of Business, University of Southern California, 3670 Trousdale Pkwy, 90089, Los Angeles, CA, United States of America and <sup>2</sup>GoodRx, 2701 Olympic Blvd, 90404, Santa Monica, CA, United States of America

\*Corresponding author: gourab@usc.edu

FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

## Abstract

We analyze weekly U.S. prescription claims data over a 36-month period spanning the COVID-19 pandemic to investigate patterns of overconsumption of the antiparasitic drug Ivermectin (IVM). To quantify the overconsumption of IVM use following the heightened public attention on IVM as a candidate treatment for COVID-19, we adopt a causal framework based on synthetic controls, comparing IVM prescription trends to those of a large set of control medications. We employ a regularized synthetic control (RSC) method using continuous spike-and-slab shrinkage priors to estimate state-level deviations in IVM consumption. This approach offers decision-theoretic guarantees on predictive risk, supporting its application in downstream policy analysis at multiple-time points after the intervention. Its empirical robustness is demonstrated through extensive validation checks. Our findings reveal a modest increase in IVM prescriptions following early reports of its potential therapeutic use, with no significant surge over the subsequent eight months. This was followed by a pronounced increase that coincided with the peak in COVID-19 cases. Strikingly, elevated IVM use persisted even after COVID-19 vaccines became widely available and multiple federal countermeasures were implemented. Our estimation approach captures the heterogeneity in long-term effectiveness of these countermeasures across states. We find that state-level political affiliation significantly explains variation in overconsumption, even after accounting for COVID-19 incidence rates. These findings highlight deep regional disparities in the effectiveness of public health messaging and suggest a need for more targeted and trusted communication strategies.

**Key words:** drug misuse, health-care communications, Ivermectin, overconsumption, persistence, regularized synthetic control, spike-and-slab priors.

## 1. Introduction

During the COVID-19 pandemic, several pharmaceutical treatments attracted widespread public attention and experienced substantial increases in consumption—often exceeding clinically recommended levels (Perlis et al., 2023; Vaduganathan et al., 2020). One such drug was Ivermectin (IVM), an antiparasitic medication primarily used to treat parasitic infections (Hill et al., 2022). Overuse of IVM can lead to adverse health

outcomes, prompting multiple federal regulatory agencies to issue warnings against its misuse (Bray et al., 2020). Nevertheless, IVM consumption remained elevated throughout the pandemic.

While some exploratory or prophylactic use is expected during a frightening public health crisis, the persistence and magnitude of IVM overuse raise critical questions about the effectiveness of medical countermeasures and health communication strategies. Accurately quantifying IVM overconsumption requires a rigorous causal inference framework (Imbens, 2024) that accounts for confounding factors such as seasonality, temporal trends, and concurrent COVID-19-related interventions (e.g., shelter-in-place orders, mask mandates, and travel restrictions), which may have influenced drug purchasing behavior.

In this study, we analyze weekly prescription claims data in the United States over a 36-month period corresponding to the COVID-19 pandemic to examine patterns of IVM overuse. Prescription claims data serve as a widely accepted proxy for drug consumption and have been employed in prior studies to investigate medication overuse (Vaduganathan et al., 2020). To establish a baseline, we compare IVM prescription trends to those of a control group of medications that did not receive public or media attention as potential COVID-19 treatments. We apply synthetic controls (SC) (Abadie and Gardeazabal, 2003; Abadie et al., 2010) to estimate counterfactual prescription levels for IVM, isolating the overconsumption attributable to pandemic-era attention. The resulting average treatment effect on the treated (ATT) captures the incremental change in IVM use due to heightened attention.

We construct state-level ATT estimates to map IVM overuse across the U.S. over time, revealing substantial regional and temporal heterogeneity. These findings point to structural differences in institutional trust and the effectiveness of public health messaging, underscoring the need for more targeted and resilient health communication systems in future public health crises.

### 1.1. IVM Over-consumption: Background and Timeline

Even though Ivermectin (IVM) first received scientific attention in the context of COVID-19 from an Australian study in April 2020, which was subsequently retracted, it was not widely known outside the scientific community at that time. IVM gained significant attention as a potential treatment for COVID-19 after Dr. Pierre Kory’s testimony at a US Senate hearing on December 8, 2020, where he advocated for its use. Subsequently, IVM received considerable attention in conservative media and from celebrities. The initial evidence supporting IVM was eventually discredited. The countermeasures to inform public about IVM misuse were primarily informative (Borges Nascimento et al., 2022), with no imposed restrictions on its prescription for COVID-19 treatment. Further evidence emerged indicating the ineffectiveness of IVM against COVID-19, and several studies that had brought IVM into the spotlight for COVID-19 were discredited by July 2021 (Hill et al., 2022).

**Table 1.** Timeline of Ivermectin (IVM) related events

Event Id	Date	Event
1	04/01/2020	Australian Study Suggests IVM Could be a Treatment for Covid-19
2	12/08/2020	Dr. Kory testifies in favor of IVM in a senate committee
3	01/14/2021	NIH: Not enough data to say IVM works
4	02/04/2021	Merck issues a warning against IVM Use
5	03/05/2021	FDA warns against IVM use
6	07/04/2021	67% of adults in the US receive at least one dose of vaccine
7	07/15/2021	Key studies that drew attention to IVM discredited
8	08/26/2021	CDC issues warning on IVM use
9	12/31/2021	End of analysis period

We present the timeline of IVM-related events in Table 1 and the plot of the weekly IVM prescription claims (indexed) in the US in Figure 1. Although the early study in April 2020 offered some preliminary evidence favoring IVM use against COVID-19, Figure 1 shows that IVM prescription claims increased substantially only in December 2020, coinciding with a U.S. Senate hearing exploring alternative treatment

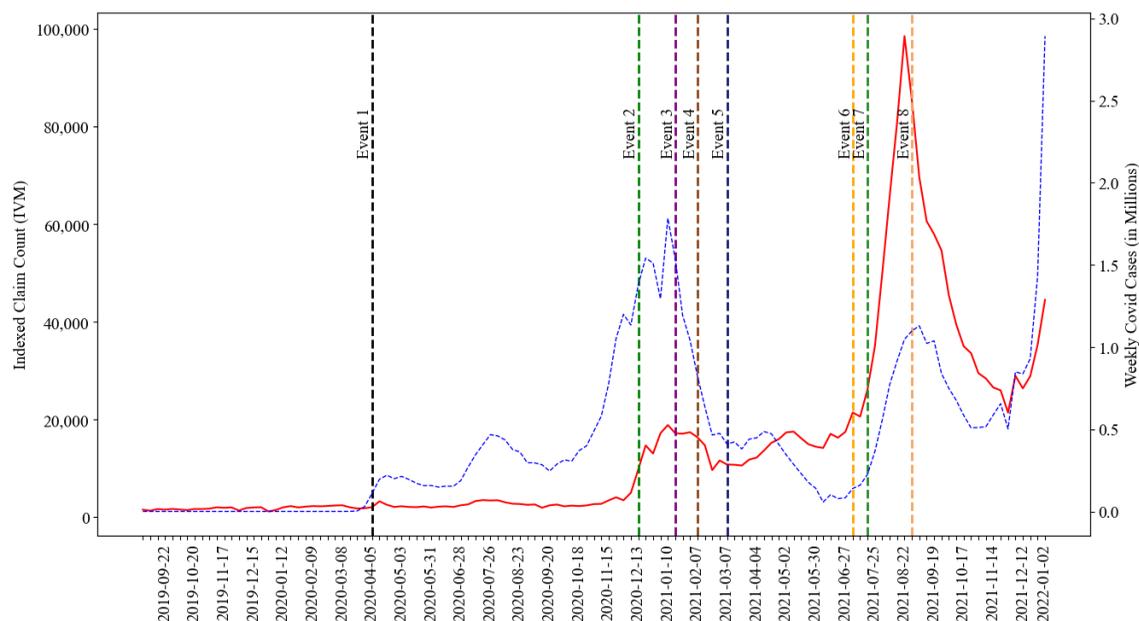


Fig. 1: Plot of total US (indexed) IVM prescription claims (in red) and Covid case counts (in dotted blue) at weekly resolution. The weeks are indexed by their start date on Sundays. The time points marked by dotted vertical lines correspond to major related events listed in Table 1.

strategies (Bonis and Curtis, 2020). The National Institute of Health (NIH) issued a formal statement on January 14, 2021, stating that there was insufficient evidence supporting IVM’s effectiveness against COVID-19 (National Institute of Health, 2021). However, prescription claims began to rise again in August 2021, coinciding with an increase in COVID-19 cases. By that time, 67% of the U.S. population had received at least one dose of the COVID-19 vaccine<sup>1</sup>, and the initial study promoting IVM for COVID-19 had been formally discredited (Hill et al., 2022). Interestingly, IVM claims peaked before COVID-19 cases in 2021, suggesting that some patients might have used IVM as a prophylactic. Additionally, there was an announcement by the Centers for Disease Control and Prevention (CDC) against IVM use on August 26, 2021 (Christensen, 2021). Despite this announcement, we observe an increase in IVM prescription claims in December 2021. We aim to determine the causal impact of these measures on observed changes in IVM prescription claims, while accounting for the effects of confounding factors.

## 1.2. Causal Study of IVM Overuse

We use the synthetic control (SC) method to estimate the net changes in IVM prescription claims due to increased public attention during Covid. We consider a very large control set of medications which provides improved control for unobserved confounders. But, we can not directly apply the widely used SC criterion of Abadie et al. (2010) for estimating the state wise ATTs for IVM as we have more controls than the number of pre-treatment observation time points. In such scenarios, it is recommended to use statistical shrinkage to produce robust SC estimates (Ben-Michael et al., 2021). Detailed literature review on this topic is provided in Sec 3.1. We employ a spike-and-slab shrinkage prior (Malsiner-Walli and Wagner, 2018; Ročková and George, 2018; Antonelli et al., 2019) based Bayesian SC (BSC) method to estimate the state-wise ATT for IVM. We provide optimal predictive risk properties of the ATT estimates. Controlling the predictive risk of ATT estimates is crucial for our application, as we are not only interested in using the ATT as an estimate of the

<sup>1</sup> <https://www.cbsnews.com/news/biden-covid-19-vaccine-goal-missed/>

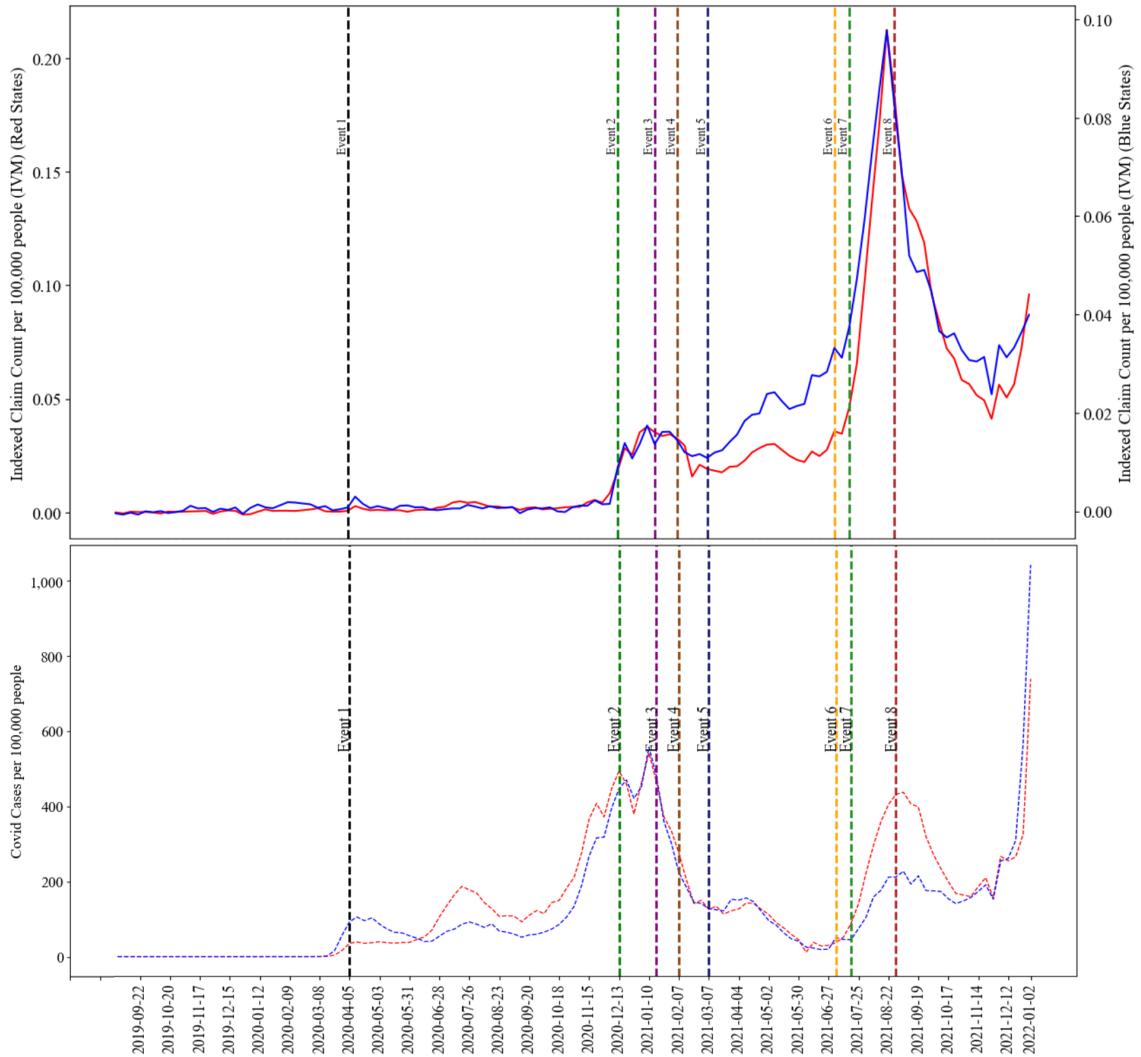


Fig. 2: Plot of weekly per capita IVM prescription claims (indexed) averaged across states that voted *Republican* (in red) and those voted *Democratic* (in blue) in the 2016 US presidential election. The time points marked by dotted vertical lines correspond to major related events listed in Table 1.

overconsumption, but also in conducting further downstream analysis to study the following two important scientific questions.

**a) Measuring Persistence of IVM overconsumption.** We empirically study the weekly changes in ATTs over twenty months as different countermeasures were announced to understand their impact on the persistence of IVM over-consumption. Informative countermeasures, such as regulatory advisories and debunking efforts, have shown mixed effectiveness in prior studies (Borges Nascimento et al., 2022). While some evidence suggests that the credibility of the source enhances their impact (Fong et al., 2022), other research finds that drug overuse can persist despite such interventions (Ecker et al., 2010, 2022; Lewandowsky et al., 2012).

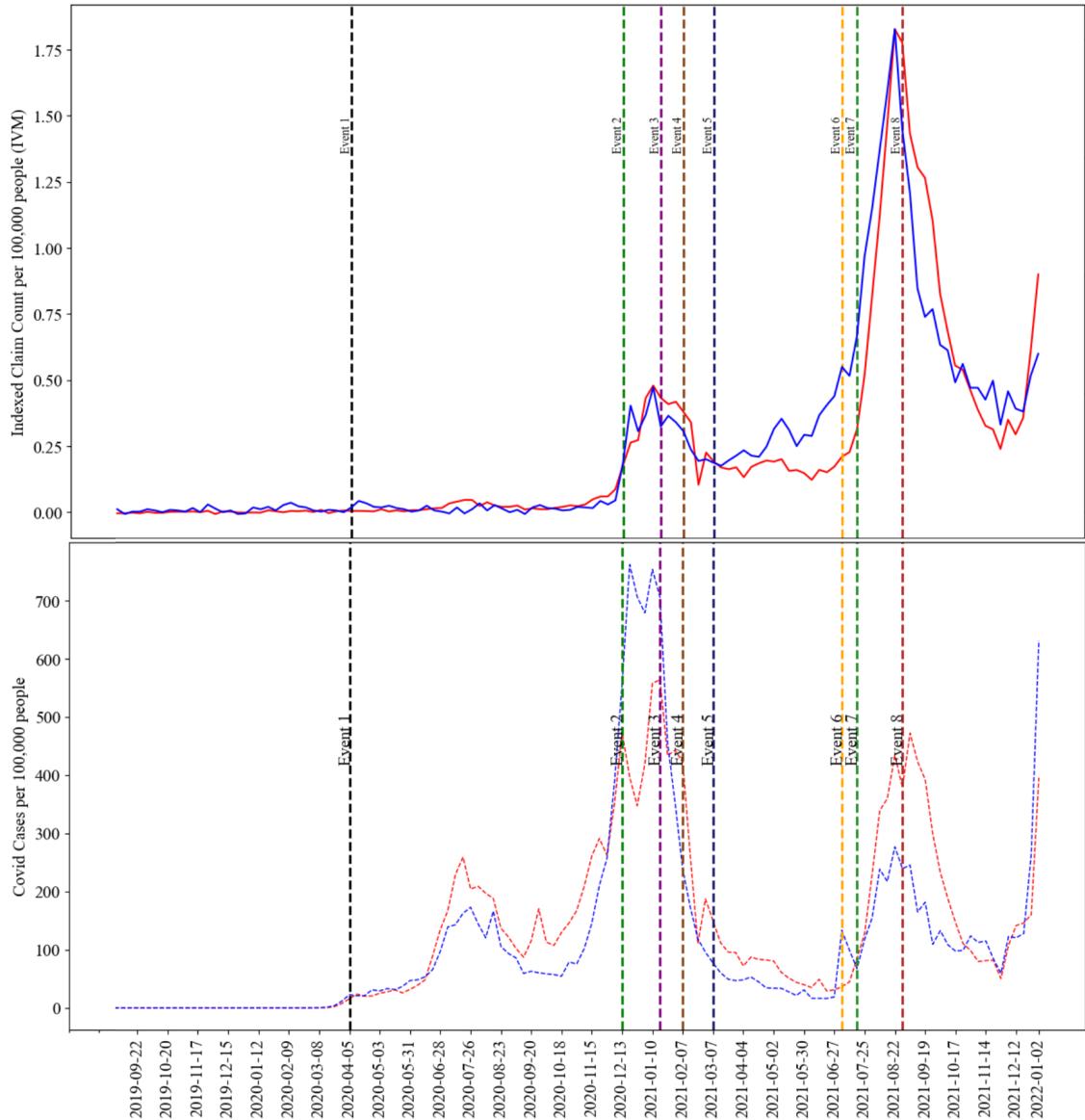


Fig. 3: Plot of weekly per capita IVM prescription claims (top) and Covid case counts (bottom) in California (in blue) and Texas (in red).

**b) Explaining State-Level Variation in Countermeasure Effectiveness.** As IVM gained significant attention as a potential COVID-19 treatment—particularly in conservative media—we examine whether a state’s political affiliation explains variation in its responsiveness to federal countermeasures. Prior descriptive studies have documented a strong correlation between IVM overuse and political affiliation (Barnett et al., 2022). To explore this relationship, we analyze state-level results from the 2016 U.S. presidential election, using a conservativeness index defined as the ratio of the Republican vote share to that of the nearest competitor, following standard practice (Pew Research Center, 2024).

In Figure 2, we plot per capita IVM consumption averaged across states that voted Republican (hereafter, Red states) and those that voted Democratic (Blue states). While Red states show substantially higher average consumption than Blue states (note the differing scales), the timing of consumption peaks is broadly aligned across both groups. However, their alignment with COVID-19 case surges exhibits more complex and state-specific patterns as Figure 3 indicates .

We illustrate this phenomenon using two major (large population) states with contrasting political leanings: Texas, a Red state, and California, a Blue state. See Figure 3. The time trends of IVM prescription claims show similar patterns across these politically diverse states, suggesting consistent underlying factors influencing IVM use. However, Texas consistently maintains higher levels of IVM prescription claims per capita compared to California throughout the observed period, despite having lower COVID-19 cases per 100,000 people. This systematic disparity in per capita claim rates, contrasting with the lower COVID-19 incidence in Texas, raises questions about potential contributing factors beyond disease prevalence.

We provide a formal analysis by using a regression based framework to examine the relationship between the heterogeneity in ATT across states and weeks, and the states responsiveness to countermeasures based on conservativeness.

### 1.3. Our Contributions

We describe the main contributions of our work below:

- We employ a shrinkage prior-based Regularized Synthetic Control (RSC) method to estimate the ATTs, using continuous shrinkage priors based on the logit-normal distribution (Thomson et al., 2019). The proposed RSC approach (see Sec. 3) is computationally efficient and enables straightforward Bayesian inference for ATTs, supporting subsequent analyses of persistence and regional susceptibility.

Within a factor model framework, we show that the ATT estimates generated by the RSC method exhibit well-controlled prediction error (see Sec. 3.1). To establish this property, we build upon theoretical results on the predictive performance of LASSO estimator (Chatterjee, 2013; Tibshirani and Wasserman, 2017) and posterior contraction results for spike-and-slab posteriors (Bhadra et al., 2016; Bhattacharya et al., 2015; Song and Liang, 2023). This predictive risk control ensures that multi-step-ahead ATT estimates can be reliably used for inference in the downstream analyses.

- We report state-level ATTs capturing the net increase in IVM consumption over a 72-week period following early reports of its potential therapeutic use. At the national level, IVM use exhibited a modest uptick following Dr. Kory’s Senate testimony and remained elevated thereafter, with pronounced peaks in July 2021 coinciding with a surge in COVID-19 cases. This temporal alignment supports the interpretation in existing literature (Rockwell et al., 2025; Kautsar et al., 2025) that IVM overconsumption was partly driven by prophylactic use during periods of heightened infection risk. We empirically demonstrate that the reported ATTs are robust to a placebo test, and we rule out alternative explanations such as stockpiling, drug shortages, and differences in media coverage of countermeasures across states in the robustness section.
- We investigate the medium-term effects of IVM overuse by analyzing prescription claims up to twenty months after public interest in its efficacy began. In contrast to prior work—which has primarily focused on short-term, descriptive analyses of prescription trends (Vaduganathan et al., 2020; Barnett et al., 2022)—our study adopts a causal framework to assess sustained impacts. Remarkably, we find that IVM prescriptions remained elevated even after the initial studies promoting its use were discredited. Most notably, IVM prescribing increased by approximately 16 times after COVID-19 vaccines became widely available (see Sec. 4.2).
- We examine whether political beliefs influences the effect of countermeasures. Our findings indicate that regulatory messaging were less effective in politically more conservative states. We also investigate the incremental effect of definitive scientific evidence against the use of IVM on prescription claims. See Sec. 4.3.

## 1.4. Organization of the paper

The paper is organized as follows. In Section 2, we present the details of the data used for the IVM case study. In Section 3, we describe the robust synthetic control methodology employed to estimate the state-wise ATT, reflecting the relative increase in IVM claims compared to control drugs. We also present the subsequent inferential analyses based on the functionals of the estimated ATTs and provide asymptotic guarantees on the accuracy of the ATT estimates. To understand the persistence of overconsumption and the impact of countermeasures, we conduct bulk-level as well as subgroup-level analyses based on state characteristics. In Section 4, we present our empirical results using the data set described in Section 2. We conclude with a discussion in Section 5. All proofs and detailed supporting tables, plots, and results are provided in the appendix.

## 2. Data

We use three different data sources: data on prescription claims, data on COVID-19 cases, and vote shares for the presidential candidates in the 2016 U.S. presidential election.

**Prescription Claims Data.** Our primary data for estimating the ATT for increased relative claims of IVM is obtained from GoodRx, a company that offers a telemedicine platform, as well as a website and mobile app that provide free drug coupons for discounted medications in the United States (US). We obtained state-level weekly prescription claims at the drug name level for the years 2019 (start) to 2021 (end) from GoodRx, covering all 50 states. Prescription claims fulfilled in a hospital setting are not included in our data. In other words, all our prescriptions require healthcare experts to agree to the course of treatment outside of a hospital setting, which is outside the emergency use authorization (EUA) orders issued by the FDA during the pandemic.

In Figure 4, we plot the average IVM prescription claims per capita for each states for four different time interval. We see not only high temporal variability but also significant differences across the US states. The IVM claims increased after Event 2 in Table 1. This increased claims continued even after event 8, i.e. after CDC’s warning on IVM usage.

We have weekly claims data recorded at the drug name level, from which we identify IVM prescription claims and tabulate claims for other drugs. We broadly divide the drugs into (a) COVID-19 drugs based on NIH’s treatment advisories between April 21, 2020, and December 30, 2021, and (b) non-COVID-19 drugs. For the group of non-covid drugs, we aggregate claims by GPI2 code for each week. Thus, our control group of medications consists of COVID-19 drugs and GPI2 code categories. We exclude Hydroxychloroquine drugs from our control group as they were also subject to widespread media and public interest during the pandemic (Perlis et al., 2023). We also excluded drugs and GPI2 categories with low volumes in the pre-treatment window and finally considered 81 COVID-19 drugs and 92 GPI2 categories. For presentation ease, hereon we term all 173 of them as control group drugs, keeping their aggregation based on category codes implicit.

As our data spans January 2019 to December 2021, it allows us to observe prescription claim behaviors well after scientific evidence showed that IVM is not a treatment for COVID-19 and that viable vaccine alternatives exist. To address data privacy concerns, the claims are indexed before conducting the analysis. The indexing is constructed by multiplying an arbitrary scaling factor for all values before any modeling. This masks the actual magnitude but preserves relative magnitudes, which are sufficient statistics for all our analyses. We report indexed claims in all the plots presented in the paper.

**Secondary Data: COVID-19 Case Counts and Political Affiliations.** To measure political affiliation of each state, we use state-level results from the 2016 U.S. presidential election from New York Times (2016). To control for the potential impact of countermeasures or policies on IVM prescriptions, we use the event timelines reported in Table 1. For COVID-19 case counts, we rely on the weekly time series data published by New York Times (2023).

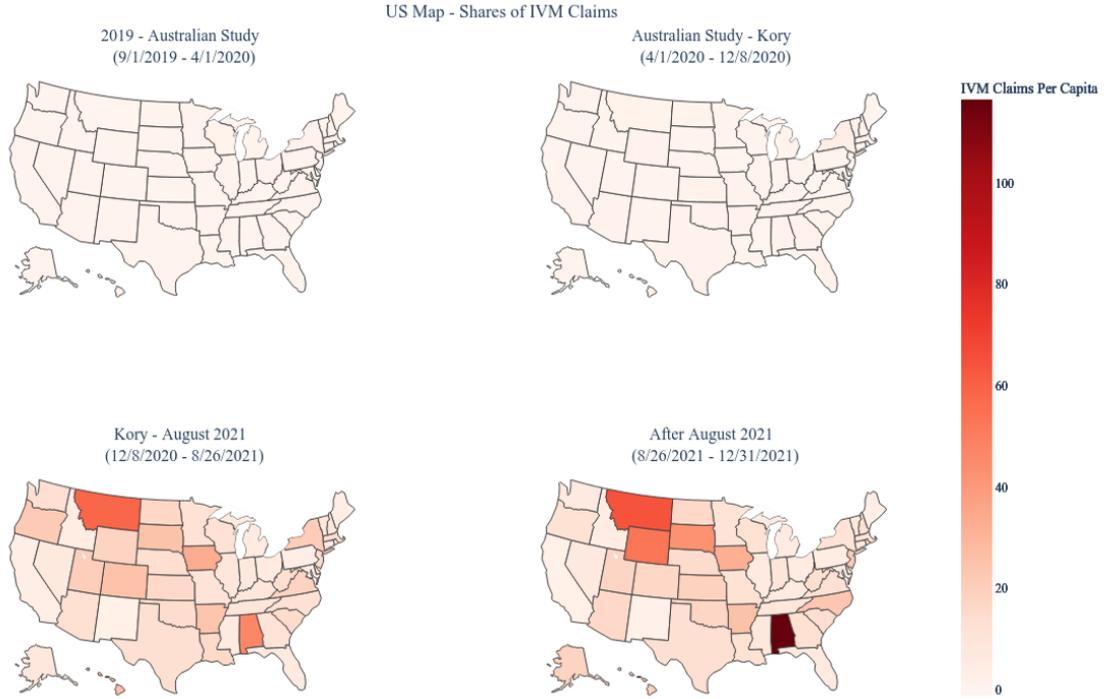


Fig. 4: Maps of average weekly IVM prescription claims for US states across four time intervals based on Table 1. From top-left, row wise we have maps for (a) Before Event 1 (b) Event 1 to Event 2 (c) Event 2 to Event 8 (d) Event 8 to Event 9.

### 3. Methodology

#### 3.1. Estimating State-level ATTs using SC

Let  $X_{cst}$  denotes the prescription claims of control drug  $c = 1, \dots, C$  at state  $s = 1, \dots, S$  and time  $t = 1, \dots, N$  weeks. The treatment date is April 1, 2020, the date when the first study on IVM was published online. The treatment occurs between  $n$  and  $n + 1$  weeks, so we have  $n$  pre-treatment weeks and  $(N - n)$  post-treatment weeks. Let  $Y_{st}$  denotes the observed IVM prescription claims at state  $s = 1, \dots, S$  and time  $t = 1, \dots, n$  weeks. After the intervention, we denote the observed IVM prescription claims by  $Y_{st}^{(1)}$ . For all  $t > n$  and  $s = 1, \dots, S$ , by  $\hat{Y}_{st}^{(0)}$  we estimate the counterfactual  $Y_{st}^{(0)}$  - IVM prescription claims had there been no intervention. Then, the ATT estimate for state  $s$  and time  $t$  is:  $\widehat{ATT}_{st} = Y_{st}^{(1)} - \hat{Y}_{st}^{(0)}$ . To estimate  $\hat{Y}_{st}^{(0)}$ , we employ synthetic control methodology (Abadie and Gardeazabal, 2003; Abadie et al., 2010). We utilize a large control set comprising both COVID-19-related drugs and medications categorized at the GPI-2 level. The inclusion of a rich set of control units allows us to implicitly adjust for multiple confounding factors, including platform-specific effects such as GoodRx market penetration, seasonality, long-term trends, and pandemic-related policies (e.g., shelter-in-place orders, mask mandates, and travel restrictions). We assume these COVID-19-related measures affect all drug claims similarly, except in cases where patients may reasonably forgo a prescription. Notably, by incorporating COVID-19 drugs into the control pool, we also account for shifts in the underlying intensity of the pandemic. We next consider a factor model structure (Abadie et al., 2010) on the prescription claims.

### 3.2. Factor Model

For any  $1 \leq s \leq S$  and  $1 \leq t \leq n$ , assume the observed IVM prescription claims are generated from a low-dimensional factor model:

$$Y_{st} = \sum_{k=1}^K \phi_{sk} \mu_{kst} + \varepsilon_{st}. \quad (1)$$

Consider counterfactual IVM prescription claims from the same model:

$$Y_{st}^{(0)} = \sum_{k=1}^K \phi_{sk} \mu_{kst} + \varepsilon_{st} \text{ for } t = n+1, \dots, N. \quad (2)$$

The claims of the control units for  $1 \leq s \leq S$  and  $1 \leq t \leq N$  obey:

$$X_{cst} = \sum_{k=1}^K \psi_{csk} \mu_{kst} + \varepsilon_{cst}, \quad (3)$$

where, the coefficient  $\psi_{cs} = (\psi_{csk} : 1 \leq k \leq K)$  varies across units and states but is invariant across time whereas the factor  $\boldsymbol{\mu}_{st} = (\mu_{kst} : 1 \leq k \leq K)$  is invariant across units but varies across time and state and are bounded by  $m$ . The noise terms  $\varepsilon_{cst}$  are independent with  $E(\varepsilon_{cst}) = 0$ ,  $E(\varepsilon_{cst}^2) = \nu^2$  and  $E(\varepsilon_{c_1 s_1 t_1} \cdot \varepsilon_{c_2 s_2 t_2}) = 0$  whenever  $c_1 \neq c_2$  or  $s_1 \neq s_2$  or  $t_1 \neq t_2$ .

### 3.3. LNP based regularized SC

As  $C \gg n$  here, the standard SC criterion of Abadie et al. (2010) cannot be directly applied, as it may lead to overfitting (Carvalho et al., 2018; Doudchenko and Imbens, 2016), resulting in high prediction error for the estimated ATT. Since we rely on multi-step-ahead ATT estimates for inference, it is essential to ensure adequate control over their predictive accuracy. To address this, Doudchenko and Imbens (2016) and Carvalho et al. (2018) recommend incorporating statistical shrinkage to produce robust SC estimates in high-dimensional settings.

The importance of shrinkage in synthetic control methods has been both theoretically and empirically validated in recent studies (Amjad et al., 2018; Abadie and L'hour, 2021; Ben-Michael et al., 2022, 2021; Kim et al., 2020). In particular, Ben-Michael et al. (2021) demonstrate that shrinkage-based SC estimators can achieve well-controlled predictive risk, and Karmakar et al. (2024) further establish the necessity of shrinkage for robust SC inference in high-dimensional regimes. To this end, we employ a Bayesian Synthetic Control (BSC) method using spike-and-slab shrinkage priors (Ishwaran and Rao, 2005; Malsiner-Walli and Wagner, 2018; Ročková and George, 2018; Antonelli et al., 2019) to estimate state-level ATTs.

We consider estimating  $Y_{st}$  in (1) via a linear combination of the outcomes of control drugs:  $\widehat{Y}_{st}(\boldsymbol{\beta}_s) = \boldsymbol{\beta}_s' \mathbf{x}_{st} = \sum_{c=1}^C \beta_{cs} X_{cst}$ , where  $\boldsymbol{\beta}_s \in \mathbb{R}^C$  for all  $s$ . We examine the difference between  $Y_{st}$  and its linear estimate  $\widehat{Y}_{st}(\boldsymbol{\beta}_s)$  under the aforementioned factor model across all time periods  $t = 1, \dots, N$ . This difference can be decomposed into the sum of two distinct components: one that depends on the model parameters and one that does not. Specifically, for any  $\mathbf{b} \in \mathbb{R}^C$ , define the residual  $R_{sk}(\mathbf{b}) = \phi_{sk} - \sum_{c=1}^C b_c \Psi_{csk}$ , and let  $R_s(\mathbf{b}) = \{R_{sk}(\mathbf{b}) : k = 1, \dots, K\}$ . The decomposition result is stated below. To simplify the presentation and to state the results uniformly for all  $t = 1, \dots, n$ , we omit the superscript on the counterfactual in (2) in the display below.

**Lemma 1** *In (1)-(3), for any  $\mathbf{b} \in \mathbb{R}^C$ ,  $Y_{st}$  can be decomposed as:*

$$Y_{st} = \widehat{Y}_{st}(\mathbf{b}) + \sum_{k=1}^K R_{sk}(\mathbf{b}) \mu_{kst} + \rho(\mathbf{b}) \zeta_{st}, \text{ for all } s \text{ and } t \leq n, \quad (4)$$

where,  $\rho(\mathbf{b}) = \nu(1 + \|\mathbf{b}\|_2^2)^{1/2}$  and  $\zeta_{st}$  has standard normal distribution and independent among each other for all  $(s, t)$  pairs.

If the linear coefficients  $\mathbf{b}$  can be optimized such that both residual components in equation (4) are small, then the counterfactual outcome can be well approximated. To this end, we consider state-specific linear estimators  $\beta_s$ . For tractability, we ignore the first residual term in (4) and the dependence of the scale of the second term on  $\|\beta_s\|$ , and instead consider a simplified misspecified model:  $Y_{st} = \widehat{Y}_{st}(\beta_s) + \sigma z_{st}$ , where  $z_{st}$  are i.i.d. standard normal noise. We construct our regularized SC estimates based on this misspecified model observed for all  $s$  but  $t = 1, \dots, n$ . In the following section, we show that despite the model misspecification, the resulting estimators accurately approximate the counterfactual outcomes generated from the factor model in (1)-(3).

Next, consider shrinkage priors on the  $\beta_s = \{\beta_{cs} : c = 1, \dots, C\}$  coefficients. Following Thomson et al. (2019) we consider an exchangeable hierarchical structure: for  $c = 1, \dots, C$  and  $s = 1, \dots, S$ :

$$[\beta_{cs} | \lambda_{cs}] \stackrel{\text{i.i.d.}}{\sim} N(0, \tau \lambda_{cs}^2) \text{ and } \lambda_{cs} \stackrel{\text{i.i.d.}}{\sim} \text{LogitNormal}(m, v)$$

where,  $m \in \mathbb{R}$  and  $\tau, v > 0$ . The above prior structure on  $\beta_s$  is equivalent to:

$$[\beta_{cs} | z_{cs}] \stackrel{\text{i.i.d.}}{\sim} N(0, \tau \exp(2z_{cs})(1 + \exp(z_{cs}))^{-2}) \text{ and } z_{cs} \stackrel{\text{i.i.d.}}{\sim} N(m, v).$$

Note that, the shrinkage of  $|\beta_{cs}|$  towards 0 is monotone in  $z_{cs}$ . As  $z_{cs} \downarrow -\infty$ ,  $\beta_{cs} \rightarrow 0$  where as  $z_{cs} \uparrow \infty$ ,  $\beta_{cs} \rightarrow N(0, \tau)$ , where,  $\tau$  is the global shrinkage factor. This Logit-Normal prior (LNP) structure is a global-local continuous shrinkage prior (Bhattacharya et al., 2015; Bhadra et al., 2016; Miller and Harrison, 2018) where  $\tau$  controls the overall level of shrinkage and the local shrinkage parameters  $\lambda_{cs}$  applies individual shrinkage to each parameter separately. LNP allows for varying levels of shrinkage across different parameters, enabling the model to retain important control units with minimal shrinkage while shrinking less important ones more aggressively (Ročková and George, 2018).

Akin to the horse-shoe priors (Bhadra et al., 2019), LNP is a fully continuous approximation to the mixed spike-and-slab prior which has an atom at origin and a continuous slab (Castillo and Misner, 2018). Unlike the mixed prior, it is much easier to sample from the LNP based posterior though the hard distinction between zero and non-zero parameters is no longer retained (Malsiner-Walli and Wagner, 2018). However, as LNP substitutes the discrete Bernoulli distribution in the mixed prior with a logit-normal distribution, with appropriate parameter settings it can resemble a U-shaped distribution over  $(0, 1)$ , concentrating most of its mass near the endpoints. We prefer LNP over the more commonly used Beta distribution (Ročková and George, 2018) as it can be derived as a transformation of standard normal random variables, significantly enhancing the convergence of the sampler (Thomson et al., 2019).

For any fixed  $\sigma, \tau, m, v$ , the logarithm of the posterior distribution of  $B = \{\beta_s : s = 1, \dots, S\}$  is given by  $\pi(B|Y, X, Z) = \sum_{s=1}^S \pi(\beta_s | Y_s, X_s, Z_s)$ , where,  $\pi(\beta_s | Y_s, X_s, Z_s)$  equals

$$\frac{1}{2\sigma^2} \sum_{t=1}^n \left( Y_{st} - \beta_{0s} - \sum_{c=1}^C \beta_{cs} X_{cst} \right)^2 + \frac{1}{2\tau} \sum_{c=1}^C (1 + \exp(-z_{cs}))^2 \beta_{cs}^2 + \frac{1}{2v} \sum_{c=1}^C (z_{cs} - m)^2.$$

We use an inverse-gamma prior on  $\sigma$ , and a half-Cauchy prior on  $\tau$ . We work with centered variables that directly evaluate  $\beta_{0s}$ . Using code written in the PyMC 4.0 probabilistic programming language (Salvatier et al., 2016), we sample from the posterior distribution and define the counterfactual estimate as  $\widehat{Y}_{st}[\mathbf{m}] := \widehat{Y}_{st}^{(0)}(\widehat{\beta}_s[\mathbf{m}]) = \mathbf{x}'_{st} \widehat{\beta}_s[\mathbf{m}]$ , where  $\widehat{\beta}_s[\mathbf{m}]$  denotes the posterior mean of  $\beta_s$ . The code for implementing the proposed regularized synthetic control estimates is publicly available from the GitHub repository <https://github.com/rrbhuyan/ivermectin>.

### 3.4. Risk Properties of regularized SC

For linear SC estimators of the form  $\widehat{Y}_{st}^{(0)}(\mathbf{b}) = \mathbf{b}' \mathbf{x}_{st}$ , it follows from appendix B of Abadie et al. (2010) that for any fixed  $\mathbf{b} \in \mathbb{R}^C$  and for all  $s$  and  $t > n$ , the error of linear SC estimates in the factor model of (1)-(3)

is upper bounded as:

$$|Y_{st}^{(0)} - \widehat{Y}_{st}^{(0)}(\mathbf{b})| \leq \|\boldsymbol{\mu}_{st}\|_2 \lambda_s \left\{ \text{Im}_s(\mathbf{b}) + \nu \rho(\mathbf{b}) \frac{\|\mathbf{W}_s\|_2}{\sqrt{n}} \right\} + \nu \rho(\mathbf{b}) |\widetilde{W}_{st}|, \quad (5)$$

where,  $\rho(\mathbf{b}) = (1 + \|\mathbf{b}\|_2^2)^{1/2}$  involves the  $L_2$  norm of  $\mathbf{b}$ ,  $\text{Im}_s(\mathbf{b}) = [n^{-1} \sum_{t=1}^n (y_{st} - \mathbf{b}' \mathbf{x}_{st})^2]^{1/2}$  is the average pre-treatment imbalance (see Theorem 1 of Ben-Michael et al. (2021)) and  $\lambda_s$  is the  $K^{\text{th}}$  eigen value of  $M_s = n^{-1} \sum_{t=1}^n \boldsymbol{\mu}'_{st} \boldsymbol{\mu}_{st}$ ;  $\mathbf{Z}_s$  is a vector of dimension  $n$  with  $\mathbf{W}_s \stackrel{d}{=} N_n(0, I)$  and  $\widetilde{Z}_{st}$  is white noise. Also,  $\{\mathbf{W}_s : s = 1, \dots, S\}$  and  $\{\widetilde{W}_{st} : s = 1, \dots, S; t = n+1, \dots, N\}$  are independent with each other.

For linear estimators of this form, unless  $\|\mathbf{b}\|_2$  is properly constrained, the error in equation (5) can become unbounded. Given that the number of control units is very large and  $C \gg n$ , the use of unregularized estimators can easily lead to this issue. However, imposing a constraint on  $\|\mathbf{b}\|_2$  indiscriminately may also result in large imbalance that will be reflected in  $\text{Im}_s(\mathbf{b})$ . To address this, the LNP prior induces sparsity in the state-specific coefficient vectors  $\mathbf{b}_s$ , thereby achieving a balance between controlling the estimation error and maintaining low imbalance.

In our factor model structure in (1)–(3), we impose a sparsity assumption. Specifically, we assume that although there is a very large number ( $C$ ) of control units, for any treated unit  $s$ , the coefficient vector  $\phi_s = (\phi_{s1}, \dots, \phi_{sK})'$  can be well approximated by a linear combination  $\Psi_s$  involving only  $\ell_s$  control units, where  $\ell_s \ll C$  for all  $s$ . The set of contributing control units may vary across treated units, and the identity of these effective control units is unknown a priori. We show that, with high probability, the proposed regularized synthetic control (SC) method based on the LNP (latent noise perturbation) framework accurately selects this subset of relevant control units. Before proceeding, we formally state the sparse factor model assumption. Sparsity assumptions of this type have been widely adopted in the synthetic control literature (see Abadie (2021); Chen and Li (2024) for a comprehensive review).

**Assumption 1.** We consider an asymptotic regime where  $n \rightarrow \infty$  and  $\liminf_{n \rightarrow \infty} C/n > 1$ . For each  $s \in \{1, \dots, S\}$ , we assume that there exists  $\boldsymbol{\beta}_s^* \in \mathbb{R}^C$  such that:

- (a) Let  $\delta_n := \max_k |R_{sk}(\boldsymbol{\beta}_s^*)|$ . Then,  $n\delta_n \rightarrow 0$  as  $n \rightarrow \infty$ .
- (b) Let  $\alpha_n := \|\boldsymbol{\beta}_s^*\|_\infty$  and  $\ell_n := \|\boldsymbol{\beta}_s^*\|_0$ , where  $\|\cdot\|_0$  denotes the number of nonzero entries and  $\|\cdot\|_\infty$  denotes the supremum norm. Then,  $\lim_{n \rightarrow \infty} n^{-1} \ell_n (\log C)^3 = 0$  and  $\lim_{n \rightarrow \infty} \alpha_n/n = 0$ .

Note that, since  $C \gg n$ , ignoring benign pathological cases we have  $\inf_{\mathbf{b} \in \mathbb{R}^C} R_{sk}(\mathbf{b}) = 0$  for all  $k$  and  $s$ . Assumption 1 guarantees the existence of at least one unknown  $\mathbf{b}_s^* \in \mathbb{R}^C$  with controlled  $L_0$  and  $L_\infty$  norms such that  $R_{sk}(\mathbf{b}_s^*)$  is asymptotically negligible. Moreover, if  $\phi_{sk}$  and  $\Psi_{csk}$  in equations (1)–(3) are uniformly bounded and the matrix  $\boldsymbol{\psi}_{.s}$  has full rank, then Assumption 1 automatically holds, as  $R_{sk}(\mathbf{b}_s^*)$  would be exactly zero.

In the context of (5), Assumption 1 ensures that both the imbalance  $\text{Im}_s(\mathbf{b}_s^*)$  and the variance  $\|\mathbf{b}_s^*\|_2$  remain controlled as  $n \rightarrow \infty$ . In such a set-up, we can set the hyper-parameters of LNP prior based on  $n$  and  $C$  only (and without using any information of the sparsity level) such that its posterior concentrates around  $\boldsymbol{\beta}_s^*$  and so, the resultant SC predictor  $\widehat{Y}_{st}^{(0)}[\mathbf{m}]$  have controlled predictive risk. To arrive at this result, based on calculations in Chatterjee (2013), we first show that the SC predictor  $\widehat{Y}_{st}^{(0)}[\mathbf{mo}] := \widehat{Y}_{st}(\widehat{\boldsymbol{\beta}}_s[\mathbf{mo}]) = \mathbf{x}'_{st} \widehat{\boldsymbol{\beta}}_s[\mathbf{mo}]$ , based on the posterior mode  $\widehat{\boldsymbol{\beta}}_s[\mathbf{mo}]$  of the LNP inherits the slow rate associated with the LASSO predictor (Tibshirani and Wasserman, 2017).

**Theorem 2** *In (1)–(3), under Assumption 1, as  $n \rightarrow \infty$ , there exists an LNP prior sequence calibrated only on  $n$  and  $C$ , such that its posterior mode based SC predictions  $\widehat{Y}_{st}^{(0)}[\mathbf{mo}]$  satisfy:*

$$E \left( Y_{st}^{(0)} - \widehat{Y}_{st}^{(0)}[\mathbf{mo}] \right)^2 = O(\mathcal{R}_n^{(1)} + \mathcal{R}_n^{(2)} + \mathcal{R}_n^{(3)} + \mathcal{R}_n^{(4)}),$$

for  $1 \leq s \leq S$  and  $n+1 \leq t \leq N$ , where the error terms on the right side are,  $\mathcal{R}_n^{(1)} = \|\boldsymbol{\beta}_s^*\|_1^2 (\log C)^{3/2} n^{-1/2}$ ,  $\mathcal{R}_n^{(2)} = \|\boldsymbol{\beta}_s^*\|_1 \delta_n \|\boldsymbol{\mu}_{st}\|_1 (\log n)^{1/2}$ ,  $\mathcal{R}_n^{(3)} = \delta_n^2 \|\boldsymbol{\mu}_{st}\|_1^2$  and  $\mathcal{R}_n^{(4)} = \nu^2 (1 + \|\boldsymbol{\beta}_s^*\|_2^2)$ .

The proof of Theorem 2 is provided in the appendix. The first term on the right-hand side of the bound reflects the slow convergence rate of the LASSO predictor. This slower rate is justified, as we make no

assumptions regarding the correlations among the control variables, which may be quite high for certain controls. The second and the third term capture the prediction error due to the model misspecification. The last term  $\mathcal{R}_n^{(4)}$  is due to future variability and is the unavoidable part of the prediction error (Ben-Michael et al., 2021). Moreover, if the time-varying components  $\mu$  in the factor model are uniformly bounded, then under Assumption 1, the first three terms in the bound become asymptotically negligible. Consequently, the prediction error is essentially  $\mathcal{R}_n^{(4)}$ , which represents the theoretically achievable lower bound using synthetic control methods with perfect oracle knowledge of  $\beta_s^*$ .

The above result can be extended for the LNP posterior mean based SC estimate  $\widehat{Y}_{st}^{(0)}[m]$ . In the appendix, using the results in (Song and Liang, 2023), we show that  $\{\widehat{Y}_{st}^{(0)}[m] : n+1 \leq t \leq T\}$  concentrates around  $\{x'_{st}\beta_s^* : n+1 \leq t \leq T\}$  with high probability. It provides asymptotic control on the predictive risk of the proposed regularized SC estimates over post treatment time-points. We subsequently use the statewise  $\widehat{\text{ATT}}_{st}$  for subsequent analysis.

**Theorem 3** *Under Assumption 1 and the factor model structure in equations (1)–(3), the posterior predictive density induced by the LNP prior sequence in Theorem 2 for predicting the counterfactual based on  $\mathcal{D}_n(s) = \{Y_{st}, X_{cst} : 1 \leq t \leq n, 0 \leq c \leq C\}$  concentrates around  $x'_{st}\beta_s^*$  at the following rate:*

$$P\left(\log\left\{\pi\left(\|y - x'_{st}\beta_s^*\| \geq c_1\epsilon_n \mid \mathcal{D}_n(s)\right)\right\} \geq -c_2n\epsilon_n^2\right) \leq \exp(-c_3n\epsilon_n^2),$$

where  $\epsilon_n = c_4\ell_n^{1/2}(\log C)^{3/2}n^{-1/2}$  and  $c_1, c_2, c_3, c_4$  are positive constants. Consequently, the trimmed posterior mean  $\widehat{Y}_{st}^{(0)}[m]$  concentrates around the true counterfactual  $Y_{st}^{(0)}$  as:

$$P\left(\|\widehat{Y}_{st}^{(0)}[m] - Y_{st}^{(0)}\| \leq (c_1\epsilon_n + \delta_n\|\mu_{st}\|_1 + c_5\rho(\beta_s^*))\right) \geq 1 - e^{-c_3n\epsilon_n^2} \text{ for all } t \geq n.$$

### 3.5. Persistence Analysis of IVM Overconsumption

In addition to addressing the  $C \gg n$  problem, the proposed regularized synthetic control (RSC) method overcomes several issues of the difference-in-differences model (Imbens and Rubin, 2015). The difference-in-differences approach would require us to pick one control drug that satisfies all the necessary assumptions, including the parallel trend assumption, from the entire set of potential controls or treat all controls as equally important, which significantly reduces the flexibility of the model. An advantage of the proposed LNP based Bayesian RSC approach over frequentist SC methods is straightforward statistical inference, facilitating hypothesis testing.

We use the aggregated state-wise ATT estimates to plot the overall US effects over time as several counter measures were implemented after  $n$ . We also look at the effects for different sub-group of the states. Results are robust to an alternative prior, the horseshoe (Kim et al., 2020), although in some instances the horseshoe prior-based model did not converge. The critical and untestable assumption is that the relationship between the focal drug and the control drugs/GPI2 categories in the pre-treatment period would continue to hold after the treatment.

### 3.6. Analysis of State-Level Countermeasure Effectiveness

We consider the variation in state-level support for the Republican party in the 2016 elections to assess how state-level ATTs vary with differences in political beliefs. To do so, we use state-level ATT estimates  $\widehat{\text{ATT}}_{st}$  for  $s = 1, \dots, S$  and  $t = n+1, \dots, T$ . Our approach to analyzing the heterogeneity in state-week ATTs is similar to meta-analysis (Fuhrer and Cova, 2020). Specifically, the state-level synthetic control approach provides us with both a point estimate of the incremental effect and the posterior distribution of the incremental effect for each state-week combination (Esposito et al., 2013). The standard deviation of each posterior distribution serves as an estimate of measurement error. To account for measurement error, we incorporate the square of this measurement error as weights in a weighted least squares regression. We estimate the following equation:

$$\text{ATT Per Capita}_{st} = \mu_0 + \mu_1 \cdot \text{Week}_t + \mu_2 \cdot \text{Consv}_i + \mu_3 \cdot \text{Covid cases per capita}_{st} + \epsilon_{st}. \quad (6)$$

The regression model analyzes the relationship between incremental IVM prescription claims per capita (ATT Per Capita) and various factors across states and time. It includes variables for time trends (Week), state conservatism (Consv), and COVID-19 cases per capita, allowing for the examination of how these factors correlate with IVM usage while controlling for potential confounding variables. Our focal variable of interest in this regression is a state’s conservativeness (Consv).

We estimate the model in (6) over three distinct time periods. The first period,  $T_1$ , spans from April 1, 2020 to December 7, 2020. This period commences with the emergence of Ivermectin (IVM) in the context of Covid-19, marked by a study that was subsequently retracted, and concludes just prior to Dr. Pierre Kory’s testimony at the Senate hearing. The second period,  $T_2$ , extends from December 8, 2020 to August 25, 2021, encompassing the duration between Dr. Kory’s testimony advocating for IVM’s use and the day before the CDC’s warning. The third period,  $T_3$ , begins on August 26, 2021, coinciding with the CDC’s warning against IVM use for Covid-19, and continues until December 31, 2021. This temporal segmentation allows us to examine the evolving impact of key events on IVM prescription patterns and public perception throughout the course of the pandemic.

## 4. Results

### 4.1. IVM Overconsumption patterns over 20 months

We standardized all the indexed claims by the mean and variances till the intervention  $n$ . We set treatment date to April 1, 2020, the date when the first study on IVM was published online. In figure 5, we plot the statewise ATT estimates aggregated across the three interesting time intervals reported in figure 4. Our analysis reveals consistent temporal patterns in IVM prescription claims across states, with notable differences in magnitude between politically diverse regions. Prior to the emergence of IVM in the COVID-19 pandemic as a potential treatment option, incremental claims per capita were negligible in all states, as expected. However, post-pandemic, we observe a significant increase in claims, with red states exhibiting nearly four times higher rates than blue states. Comparing the ATT plots in figure 5 with the raw claims in figure 4, we see fundamental differences. For example, comparing between the penultimate and last maps in both figures, we see that the ATTs of Texas, Florida and even California are comparatively much pronounced than raw claims maps in figure 4. This differences show the importance of conducting a causal study based on control drugs rather than reporting the differences based on raw IVM claims only. Next, we study these ATTs as counter measures are introduced.

In Figure 6, we plot the weekly ATT estimates for IVM prescription claims aggregated across all US states. We see that the two major peaks coincide with the peaks of increased COVID19 case counts in the US. The latter peak is much pronounced than the former peak and occurred after key studies that drew attention to IVM usage for COVID19 was scientifically discredited. Next, we study the fluctuation in the ATTs as countermeasures was introduced.

### 4.2. Persistence Analysis

Considering the time varying treatment effects in the ATT plot of figure 6 we observe a small, short-lived increase in prescription claims in relation to the counterfactual on April 1, 2020, but the major increase does not occur till December 8, 2020, when Dr. Kory testified for the senate committee. This increase sustains till the end of 2021, with a substantial increase in 2021 coincident with an increase in COVID cases. Thus, we observe a persistent effect of IVM overconsumption, despite federal informative countermeasures. Unlike restrictive countermeasures that involves banning consumption, informative countermeasures are actions designed to reduce the spread and impact of misleading information by enhancing public knowledge, promoting critical thinking, and improving access to accurate information. All IVM countermeasures were informative in nature.

In table 2, we summarize the information in the weekly ATTs plots of figure 6 by considering three interesting time periods from the span of the study. Note that, as these ATT estimates are based on indexed claims, the estimates unless compared to a base is not interpretative. For each of the time-period, we report the average increase in total US ATTs as relative change over the counterfactual, i.e., for any time-interval

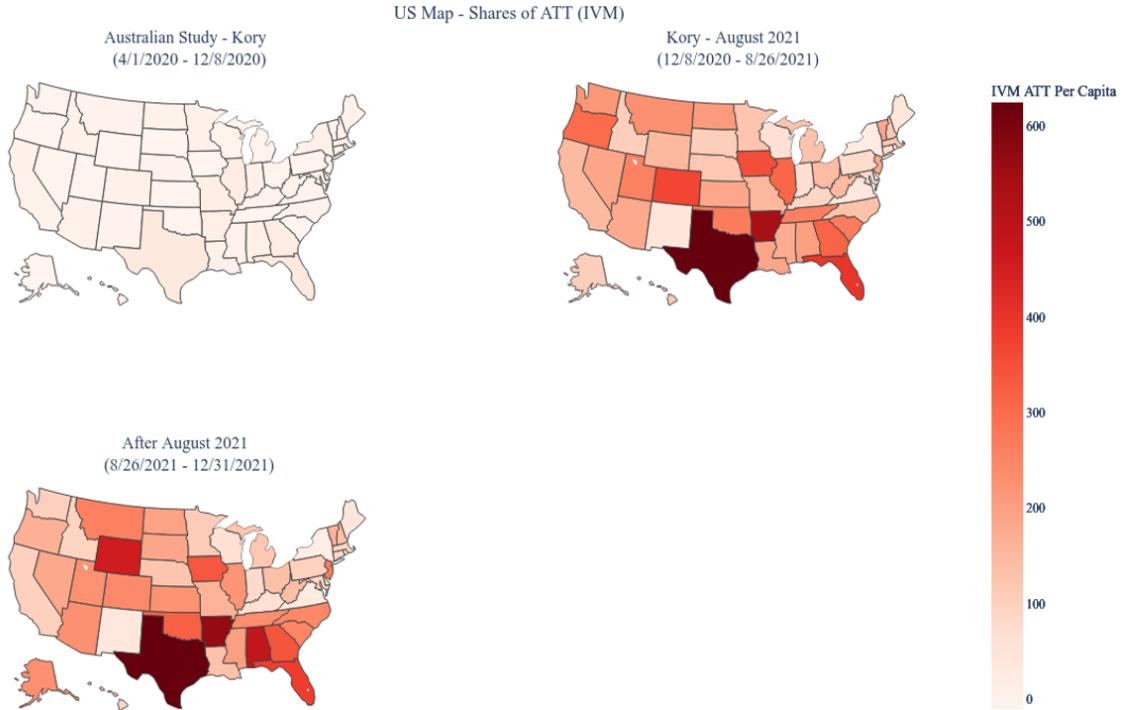


Fig. 5: From top-left row-wise, we have maps of ATT estimates for IVM averaged across the time intervals (a) Event 1 to Event 2 (b) Event 2 to Event 8 (c) Event 8 to Event 9. Events listed in Table 1.

$\mathcal{T}$ , we report  $(\sum_{s \in \mathcal{S}, t \in \mathcal{T}} Y_{st}(1) - \hat{Y}_{st}(0)) / (\sum_{s \in \mathcal{S}, t \in \mathcal{T}} \hat{Y}_{st}(0))$ . For approximately a month after Dr/ Kory's testimony, the relative increase in IVM claims were around 6 times. However, we observe that the highest increase on average which was 16 times occurred after 67% of the US population had received at least one dose of the COVID19 vaccine. Also, at that point there has been scientific evidence refuting the initial studies that suggested IVM usage for COVID19. This shows that the effect of overconsumption persisted even after counter measures and vaccination.

**Table 2.** Change in Incremental IVM Claims Over Time

Interval	Significance	ATT Increase
12/08/2020 – 01/14/2021	Immediately after Dr. Kory's Senate Testimony	6.15 times
01/21/2021 – 07/04/2021	After NIH Refutes IVM usage for COVID19	4.52 times
07/04/2021 – 12/31/2021	After Vaccination reaches 67% of US population	16.18 times

#### 4.3. State-wise variability in regulatory countermeasure efficacy

We next calculate the weekly ATT per capita across the Blue and the Red states. This is done by adding the ATTs across all the Blue states and then dividing by their cumulative population size. Similar calculation is done for Red states. We plot the weekly time series in Figure 7. In figure 8, we plot the ATT per capita for California and Texas as representative examples of Blue and Red states. Note the difference in scale for the y-axes for Red states (on the left) and Blue states (on the right). The ATT per 100,000 people zero for both Red and Blue states, prior to IVM gaining attention. Subsequently the increases in claims are systematically

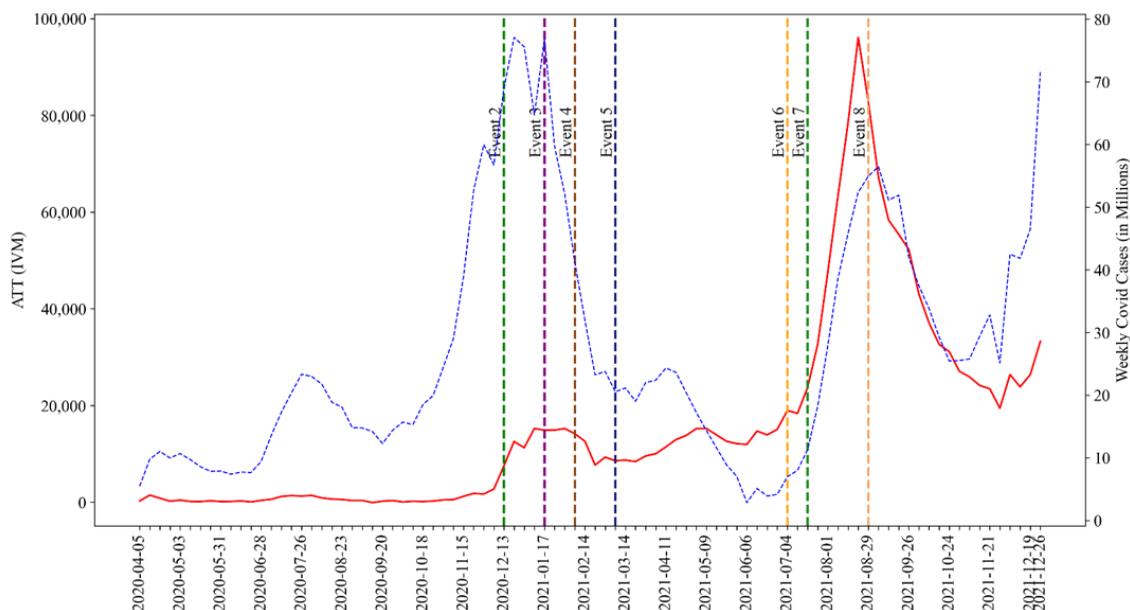


Fig. 6: Plot of weekly ATT estimates for IVM prescription claims aggregated across all US states (in red) and of total US weekly Covid case counts (in blue). The time points marked by dotted vertical lines correspond to major related events listed in Table 1.

higher in Red states than in Blue states. The disparity we see between the Red and Blue series in figure 7 is even more pronounced in figure 8. Both states show similar temporal trends, starting with zero incremental claims per capita before the IVM gained public attention. However, after Dr. Kory's testimony on December 8th, 2020, Texas demonstrates a markedly higher increase in IVM prescription claims compared to California, despite having lower COVID-19 case rates per capita (see Figure 4 bottom subplot).

To systematically analyze these discrepancies and understand the association between countermeasure efficacy and political affiliation, we divide the study period into three distinct intervals. These intervals correspond to those considered in Table 2. We use state-level ATT-derived per capita estimates of overconsumption of IVM claims as our response variable, which measures Ivermectin prescribing beyond expected baseline levels. For each of the three intervals, we conduct a regression analysis based on the framework described in (6). Weighted least squares estimates are employed to down-weight states with low claims. The results are reported in Table 3.

Our analysis reveals that Covid case counts have a significant positive effect on increasing incremental IVM claims across all three intervals. Conservatism, however, was not a significant predictor at the 5% level in two intervals. It was significant and positively associated with incremental IVM claims during the last interval. This finding is noteworthy because, by July 15, 2021, the initial studies suggesting that IVM could help with Covid-19 had been discredited. Moreover, the CDC issued a statement on August 26, 2021, discrediting the effectiveness of IVM, which led to a reduction in prescription claims. Thus, the significant positive association between conservatism and incremental IVM claims in the third column of Table 3, along with the exploratory findings in Figures 7 and 8 showing a subsequent increase in prescription claims towards the end of 2021, suggests that even though definitive evidence discrediting IVM was effective, its impact was short-lived and insufficient to return IVM prescriptions to the counterfactual baseline.

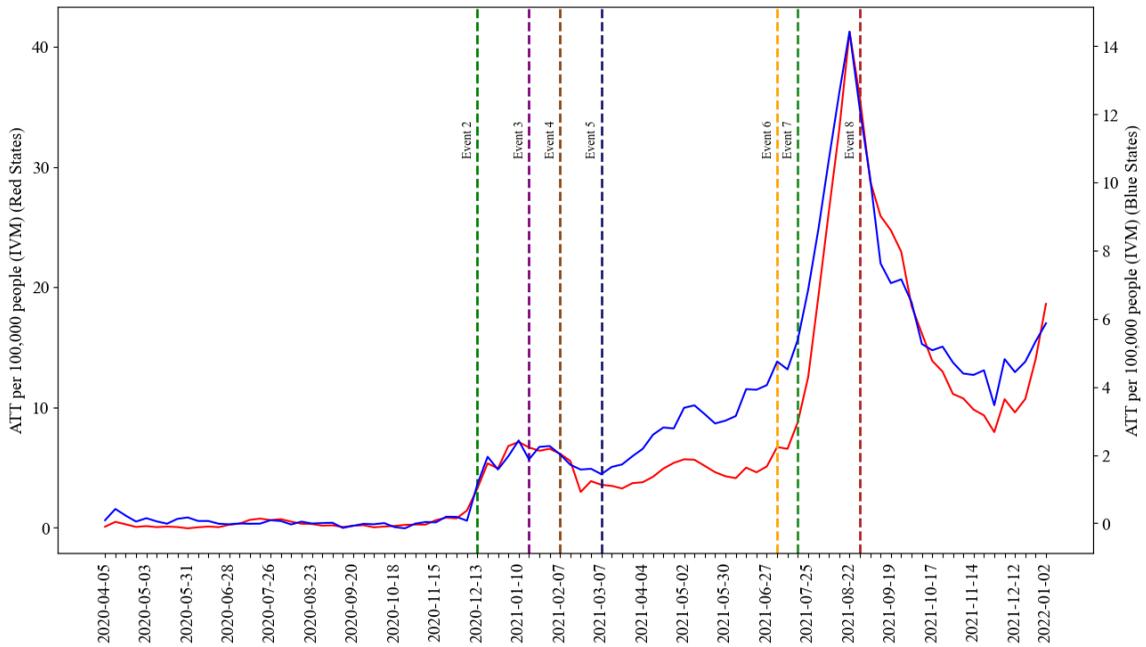


Fig. 7: Plot of weekly ATT estimates for IVM prescription claims aggregated across Blue and Red states. The time points marked by dotted vertical lines correspond to major related events listed in Table 1.

**Table 3.** Variability in countermeasure efficacy. Results from regression analysis based on (6) for three time periods are presented. Note: \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

	<b>Immediate 12/8/2020-1/7/2021</b>	<b>After NIH Refutes IVM 1/14/2021-7/4/2021</b>	<b>After Vaccination 7/4/2021-12/31/2021</b>
Constant	-1.085 (0.773)	0.868 (0.588)	1.628* (0.742)
Cases Per Capita	0.007*** (0.002)	0.007*** (0.002)	0.022*** (0.002)
Conservativeness	0.600 (0.400)	0.039 (0.129)	2.199*** (0.473)
Observations	150	1,100	1,250
R <sup>2</sup>	0.127	0.106	0.318
Adjusted R <sup>2</sup>	0.103	0.087	0.303
Residual Std. Error	3.367 (df = 145)	2.616 (df = 1076)	8.899 (df = 1223)
F Statistic	5.269*** (df = 4, 145)	5.531*** (df = 23, 1076)	21.927*** (df = 26, 1223)

#### 4.4. Robustness Checks

We conduct robustness checks in two main categories: data, model-related checks and tests for alternative explanations of our observed effects. The results are provided in the appendix.

## 5. Discussion

This study is the first to examine the medium-term effects of Ivermectin (IVM) overuse and the efficacy of federal informative countermeasures within a causal inference framework. Our findings reveal that IVM overuse persisted well beyond the initial reports of its potential therapeutic use, despite the implementation of regulatory advisories and public health messaging. While countermeasures yielded some benefits, their overall impact was limited, and substantial residual overuse remained.

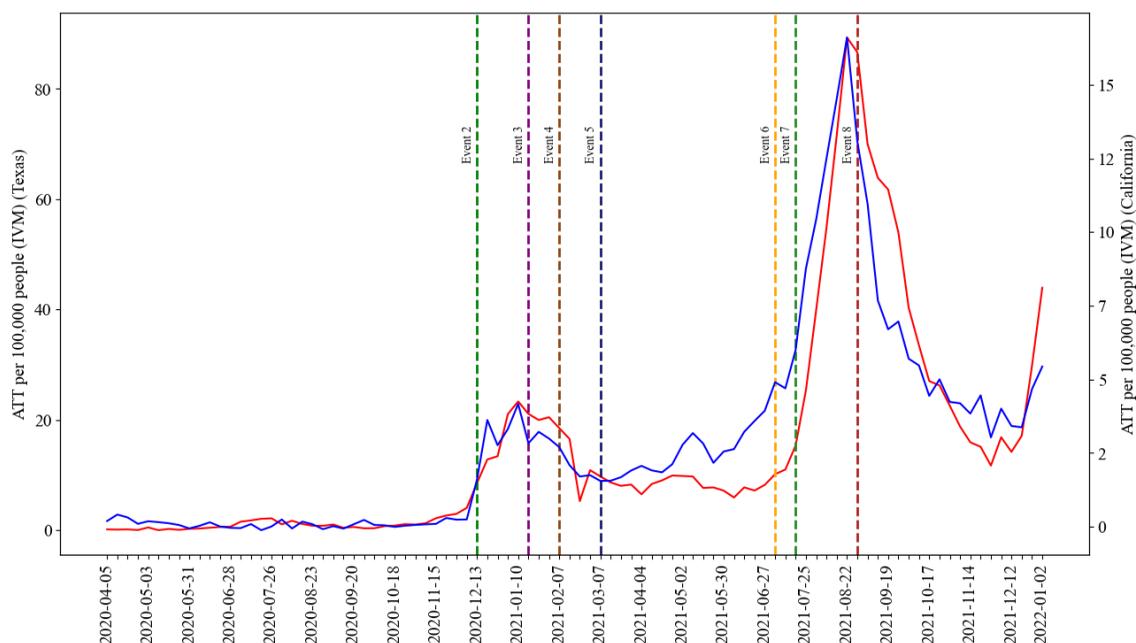


Fig. 8: Plot of weekly ATT estimates for IVM prescription claims for California (in blue) and Texas (in red). The time points marked by dotted vertical lines correspond to major related events listed in Table 1.

We observe significant heterogeneity in the effectiveness of countermeasures across states. States with higher levels of political conservativeness experienced greater per capita overconsumption and exhibited lower responsiveness to federal advisories. Additionally, the dissemination of definitive scientific evidence discrediting IVM’s efficacy appeared to have only a modest impact on curbing its overuse.

These findings highlight the limitations of federal informative countermeasures during the pandemic and underscore the need for more targeted, localized health communication strategies. The substantial variation in state-level responses suggests that one-size-fits-all messaging is insufficient. Tailoring messages to specific subpopulations—based on behavioral, cultural, and political contexts—can improve salience and effectiveness. Integrating data analytics and behavioral insights into communication design may enhance message uptake and public compliance. Effective health communication will require coordinated efforts across disciplines, including public policy, healthcare, behavioral science, and communication, to develop adaptive and resilient strategies for future public health crises.

## A. Robustness Checks

Our robustness checks broadly fall into two types: data/model related and alternative explanations for our observed effects.

**Single Platform Data.** Our data is sourced from a single platform, GoodRx. We view our analysis as an index of claim patterns in the market. As a data check, IVM claims at the national level from another data supplier, who wishes to be anonymous. The anonymous data supplier is a hub for an extensive pharmacy network, routing real-time Rx claims transactions between pharmacies and pharmacy benefit managers (PBMs). This connectivity offers broad coverage of pharmacies, prescribing providers, and PBMs sourced as a part of the business platform. Coverage includes specialty pharmacy claims, general distribution prescription drugs, and durable medical equipment claims filled via pharmacy. The data supply is 1st party and deliverable on a  $t+1$  basis from the claims submission date. The Pearson correlation between GoodRx data and the data supplier data for IVM claims is 0.9.

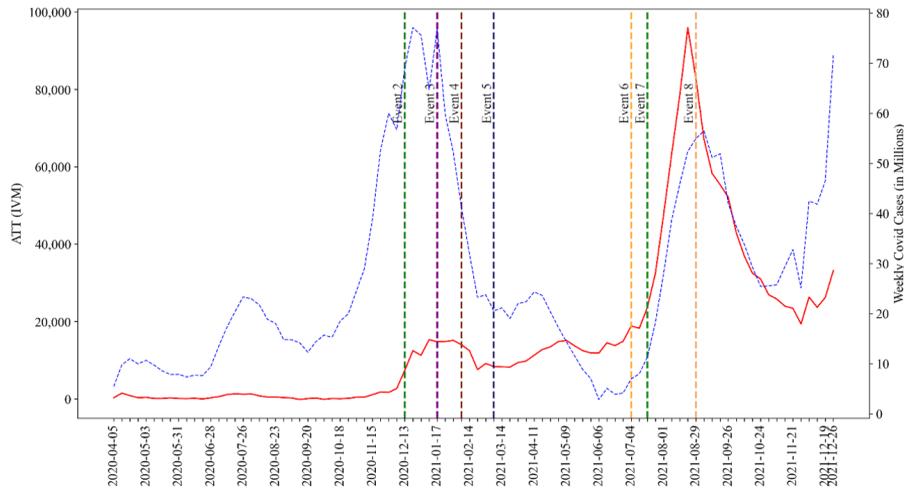


Fig. 9: Based on an alternative control set we plot the ATT estimates for IVM prescription claims aggregated across all US states (in red) and of total US weekly Covid case counts (in blue).

**Choice of Controls.** We estimated models with other drugs that were used to treat Covid-19 and available at pharmacies along with aggregate GPI2 categories as controls. We identified drugs used for Covid-19 treatment from the NIH’s treatment advisories between April 21, 2020 and December 30, 2021 (inclusive) for Covid-19. We check if the results substantively remain the same if we use different control sets. We tried (a) top 200 drugs by volume (b) use a random sampling of 200 drugs, and (c) use top 100 drugs by volume and aggregate GPI2 categories as controls. We see very similar results in all the three cases. To avoid repetition, we present the results from case (c) only. We plot the ATT estimates based in case (c) in figure 9. We see that with this alternative specification the treatment effect estimates are even larger. The regression results for this alternative specification are presented in Table 4. Comparing with table 3 we see that these results are also substantively the same.

**Table 4.** Moderating Effect of Susceptibility based on an alternative control set. Results from regression analysis based on (6) for three time periods are presented. Note: \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

	Immediate 12/8/2020-1/7/2021	After NIH Refutes IVM 1/14/2021-7/4/2021	After Vaccination 7/4/2021-12/31/2021
Constant	-1.242 (0.766)	0.342 (0.611)	1.432 (0.755)
Cases Per Capita	0.007*** (0.002)	0.008*** (0.002)	0.014*** (0.002)
Conservativeness	0.742 (0.429)	0.098 (0.130)	2.895*** (0.446)
Observations	150	1,100	1,350
R <sup>2</sup>	0.126	0.110	0.289
Adjusted R <sup>2</sup>	0.102	0.091	0.274
Residual Std. Error	3.442 (df = 145)	2.657 (df = 1076)	8.920 (df = 1321)
F Statistic	5.235*** (df = 4, 145)	5.772*** (df = 23, 1076)	19.217*** (df = 28, 1321)

**Placebo Test.** As a falsification test, we assume that the treatment occurred at the beginning of 2020, i.e., we move back the treatment date by three months from April 1, 2020, the date when the first study suggesting IVM might work was published, to January 12, 2020. We re-estimate the counterfactuals by SC and present the results in Figure 10. We observe no incremental IVM claims prior to April 1, 2020. This suggests that our model specification is appropriate and indicates that there was no stockpiling in anticipation of shortages before April 1, 2020.

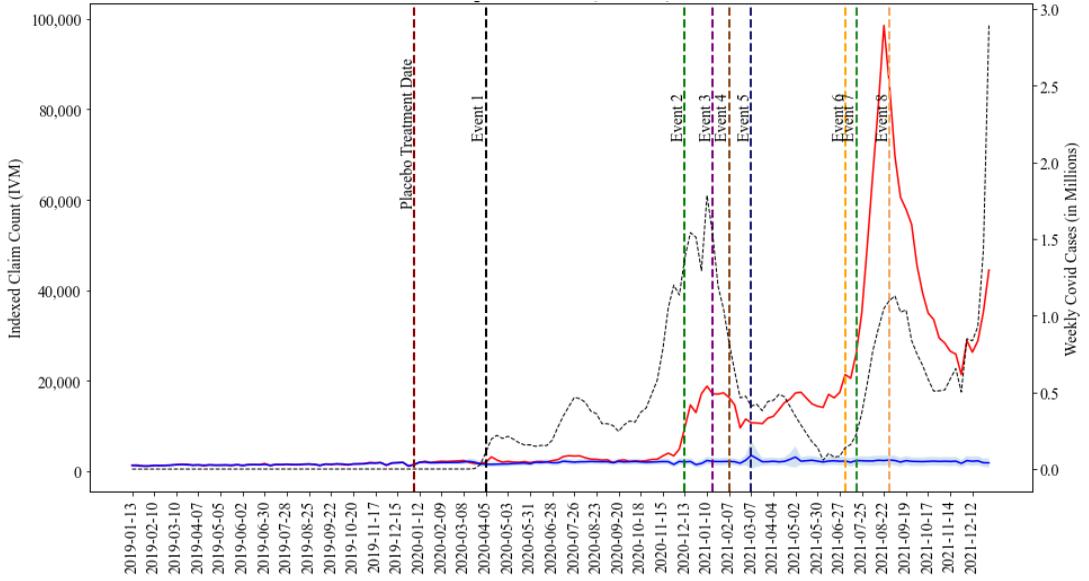


Fig. 10: In red and blue we have the actual and predicted IVM claims when the treatment date was preponed by 12 weeks for a placebo test. In dotted black line we have Covid case counts.

**Covid Cases.** Inclusion of covid death rates in place of case rates does not substantially alter our results.

**Differential Media Coverage of Countermeasures across states.** It is quite possible that the media covered countermeasures differently across states. This could mitigate the effect of countermeasures. However, our study focuses on prescription claims - where physicians have to sign off on the prescription. Ideally, experts like physicians should be able to directly access updated information outside of media sources.

**Substitutes for IVM.** There are no over-the-counter (OTC) substitutes for IVM and IVM purchase requires a prescription. Therefore, differential access to over-the-counter substitutes does not explain our results. Some consumers used animal formulations of IVM, but this is not pertinent to our study as such consumption does not include physician decisions.

## B. Proofs of the results in Section 3.1

### B.1. Proof of Lemma 1

Based on our factor model, we have  $Y_{st} = \sum_{k=1}^K \phi_{sk} \mu_{kst} + \varepsilon_{st}$ , and similarly the controls are  $X_{cst} = \sum_{k=1}^K \psi_{csk} \mu_{kst} + \varepsilon_{cst}$ . The residual loadings for any  $\mathbf{b} \in \mathbb{R}^c$  is defined as

$$R_{sk}(\mathbf{b}) = \phi_{sk} - \sum_{c=1}^C b_c \psi_{csk}.$$

Since  $\phi_{sk} = R_{sk}(\mathbf{b}) + \sum_{c=1}^C b_c \psi_{csk}$ , we can rewrite  $Y_{st}$  as

$$\begin{aligned} Y_{st} &= \sum_{k=1}^K \left( R_{sk}(\mathbf{b}) + \sum_{c=1}^C b_c \psi_{csk} \right) \mu_{kst} + \varepsilon_{st} \\ &= \sum_{k=1}^K R_{sk}(\mathbf{b}) \mu_{kst} + \sum_{c=1}^C b_c \sum_{k=1}^K \psi_{csk} \mu_{kst} + \varepsilon_{st}. \end{aligned}$$

Next, by using the factor models for the control units, we can replace  $\sum_{k=1}^K \psi_{csk} \mu_{kst}$  with  $X_{cst} - \varepsilon_{cst}$ , to simplify as follows,

$$\begin{aligned} Y_{st} &= \sum_{k=1}^K R_{sk}(\mathbf{b}) \mu_{skt} + \sum_{c=1}^C b_c (X_{cst} - \varepsilon_{cst}) + \varepsilon_{st} \\ &= \sum_{k=1}^K R_{sk}(\mathbf{b}) \mu_{kst} + \hat{Y}_{st}(\mathbf{b}) - \sum_{c=1}^C b_c \varepsilon_{cst} + \varepsilon_{st}. \end{aligned}$$

The final result follows by noting that  $\varepsilon_{st} - \sum_{c=1}^C b_c \varepsilon_{cst} \stackrel{d}{=} \rho(\mathbf{b}) \zeta_{st}$ , where  $\rho(\mathbf{b}) = \nu(1 + \|\mathbf{b}\|_2^2)^{1/2}$  and  $\zeta_{st}$  is standard normal.

## B.2. Proof of Theorem 2

Consider the setup for a particular state  $s$ . We remove subscripts pertaining to state  $s$  for ease of notation. In this case we model the pre-treatment IVM claims  $\{Y_t : t = 1, \dots, n\}$ , with the misspecified model for a single state as

$$Y_t = \sum_{c=1}^C \beta_c X_{ct} + \epsilon_t; \quad [\beta_c | \lambda_c] \stackrel{\text{i.i.d.}}{\sim} N(0, \tau \lambda_c^2) \text{ and } \lambda_c \stackrel{\text{i.i.d.}}{\sim} \text{LogitNormal}(m, v),$$

For fixed  $\sigma, \tau$  and a fixed sequence of  $\lambda_c$ , the negative log likelihood of posterior distribution of  $\beta_c$  without the prior structure on  $\lambda_c$  is

$$l(\beta | Y, X, \lambda) \propto \frac{1}{2\sigma^2} \sum_{t=1}^n (Y_t - \sum_{c=1}^C \beta_c X_{ct})^2 + \frac{1}{2\tau} \sum_{c=1}^C \frac{\beta_c^2}{\lambda_c^2}.$$

The posterior mode  $\hat{\beta}$  is the minimizer of the above loss function (negative log-likelihood),  $\hat{\beta} = \arg \min l(\beta | Y, X, \lambda)$ . The optimization problem of minimizing the loss can be rewritten as a constrained problem for some  $K$  as

$$\min_{\beta} \quad \frac{\tau}{\sigma^2} \sum_{t=1}^n (Y_t - \sum_{c=1}^C \beta_c X_{ct})^2 \quad \text{s.t.} \quad \sum_{c=1}^C \frac{\beta_c^2}{\lambda_c^2} \leq K. \quad (7)$$

We next provide the construction to choose prior hyperparameters for  $\lambda_c$ .

**Lemma 4** *Let  $X \sim \text{LogitNormal}(-tf(t), t^2)$ , where  $f(t)$  is the solution to the equation*

$$\frac{f(t)^2}{2} + \log f(t) = (1 + u) \log t,$$

*then as  $t \rightarrow \infty$ , for some  $u > 0$ , there exists constants  $k_1, k_2$  such that*

$$P[X \in (0, t^{-1/2})] = 1 - k_1 t^{-(1+u)} \quad P[X \in (1 - t^{-1/2}, 1)] = k_2 t^{-(1+u)}$$

Using the above lemma, we set the prior mean  $-Cf(C)$  and variance  $C^2$  for  $\lambda_c$ . The limiting distribution closely resembles a Bernoulli distribution. With this choice of parameters  $\lambda_c \leq C^{-1/2}$  with probability  $1 - k_1 C^{-(1+u)}$  and the entire sequence of  $\lambda_c$  is bounded by  $1/\sqrt{C}$  with probability  $(1 - k_1/C^{1+u})^C \approx 1 - C^{-u}$  for some  $u > 0$ . Thus,  $\lambda_c \geq 1/\sqrt{C}$  for only constants  $s$   $\lambda_c$ 's as  $C$  grows. Without loss of generality assume

that  $\lambda_1, \lambda_2, \dots, \lambda_s$  are greater than  $1/\sqrt{C}$  and  $\lambda_{s+1}, \lambda_{s+2}, \dots, \lambda_C \leq 1/\sqrt{C}$ . Consider  $2s\sqrt{\sum_{c=1}^C \beta_c^2/\lambda_c^2}$ ,

$$\begin{aligned} 2s\sqrt{\sum_{c=1}^C \beta_c^2/\lambda_c^2} &= 2s\sqrt{\sum_{c=1}^s \beta_c^2/\lambda_c^2 + \sum_{c=s+1}^C \beta_c^2/\lambda_c^2} \geq s\sqrt{\sum_{c=1}^s \beta_c^2/\lambda_c^2} + s\sqrt{\sum_{c=s+1}^C \beta_c^2/\lambda_c^2} \\ &\geq s\sqrt{\sum_{c=1}^s \beta_c^2} + \sqrt{C \sum_{c=s+1}^C \beta_c^2} \geq \|\beta\|_1. \end{aligned}$$

This means that  $\sum_{c=1}^C \beta_c^2/\lambda_c^2 \leq K$  implies that  $\|\beta\|_1 \leq 2s\sqrt{K}$ . Thus consider the following optimization problem instead of (7),

$$\min_{\beta} \sum_{t=1}^n \frac{\tau}{\sigma^2} (Y_t - \sum_{c=1}^C \beta_c X_{ct})^2 \quad \text{s.t.} \quad \|\beta\|_1 \leq 2s\sqrt{K}, \quad (8)$$

where,  $\tilde{Y}_t = \sqrt{\tau/\sigma^2} Y_t$  and  $\tilde{X}_{ct} = \sqrt{\tau/\sigma^2} X_{ct}$ , the problem is thus normalized for the prior parameters. While the objective function of the optimization problems in both (7) and (8) are the equivalent, since  $\sum_{c=1}^C \beta_c^2/\lambda_c^2 \leq K$  implies that  $\|\beta\|_1 \leq 2s\sqrt{K}$ , thus, the feasible set of (8) is a subset of the feasible set of (7). Minimising over a smaller region can only increase (or leave unchanged) the optimal value, so the infimum attained in (8) is at least as large as that for (7). Because our ultimate goal is to derive an upper bound on the prediction error, we concentrate on solving the tighter problem (8). The optimization problem in (8) resembles the problem defined in Chatterjee (2013) to show the bound on the prediction error. We provide a version of the result from Chatterjee (2013), which we use to provide our result

**Result** (Chatterjee (2013)) *Consider the regression setup  $Y_t = \sum_{c=1}^C \beta_c^* X_{ct} + \epsilon_t$  where  $\|\beta^*\|_1 \leq K_\beta$  and the errors have variance  $\sigma^2$ . Let MSPE stand for the mean squared prediction error, defined as  $E(\tilde{Y} - \hat{Y})^2$  where  $\hat{Y}$  is the prediction based on  $\hat{\beta}$  (restricted to  $\|\hat{\beta}\|_1 \leq K$ ) and  $\tilde{Y}$  is the best possible prediction based on the true parameter  $\beta^*$ , then*

$$E(\tilde{Y} - \hat{Y})^2 = \tilde{K} M \sigma \sqrt{\frac{\log(p)}{n}} + 2\tilde{K}^2 M^2 \sqrt{\frac{\log(p)}{n}},$$

where  $\max |X_{ct}| \leq M$  and  $\tilde{K} := (K + K_\beta)^2$ .

In our case, as described in Lemma 1, we have

$$Y_t = \sum_{c=1}^C \beta_c^* X_{ct} + \sum_{k=1}^K R_k(\beta^*) \mu_{kt} + \rho(\beta^*) \zeta_t. \quad (9)$$

Since we have formulated our optimization problem in (8) using the normalized  $\tilde{Y}_t$ , the variance in our setting is thus  $\tau\rho(\beta^*)/\sigma^2$ . Also  $M$  is the maximum of  $n$  of  $X_{ct}$ 's which is of the order  $\sqrt{\log n}$ . The model misspecification term  $\sum_{k=1}^K R_k(\beta^*) \mu_{kt}$  also adds another error of  $M\tilde{K} \sum_{k=1}^K R_k(\beta^*) \mu_{kt}$  into our existing error. Combining all, we get for our case, given fixed  $\tau$  and  $\sigma$ ,

$$\begin{aligned} E\left[\left(\tilde{Y}_t^{(0)} - \hat{Y}_t\right)^2 \mid \tau, \sigma\right] &= \tilde{K} M \sqrt{\frac{\tau\rho(\beta^*) \log C}{\sigma^2 n}} + 2\tilde{K}^2 M^2 \sqrt{\frac{\log C}{n}} + M\tilde{K} \sum_{k=1}^K R_k(\beta^*) \mu_{kt} \\ &= \tilde{K} M \left( \sqrt{\frac{\tau\rho(\beta^*) \log C}{\sigma^2 n}} + 2\tilde{K} M \sqrt{\frac{\log C}{n}} + \sum_{k=1}^K R_k(\beta^*) \mu_{kt} \right) \end{aligned}$$

If  $\sigma \sim \text{Inverse-Gamma}(a_0, b_0)$  and  $\tau \sim \text{Half-Cauchy}(c_0, \gamma)$ , then  $E[1/\sigma] = a_0/b_0$  and  $E[\sqrt{\tau}] = c_0 + \sqrt{2/\gamma}$ . Since  $\sigma$  and  $\tau$  are independent of each other, we can simplify  $E(\tilde{Y}_t^{(0)} - \hat{Y}_t)^2$  as,

$$\begin{aligned}
& E \left[ E \left[ \left( \tilde{Y}_t^{(0)} - \hat{Y}_t \right)^2 \mid \tau, \sigma \right] \right] \\
&= \tilde{K}M \left( \frac{a_0}{b_0} (c_0 + \sqrt{2/\gamma}) \sqrt{\frac{\rho(\beta^*) \log C}{n}} + 2\tilde{K}M \sqrt{\frac{\log C}{n}} + \sum_{k=1}^K R_k(\beta^*) \mu_{kt} \right) \\
&\leq \tilde{K} \sqrt{\log n} \left( c_1 \sqrt{\frac{\rho(\beta^*) \log C}{n}} + 2\tilde{K} \sqrt{\log n} \sqrt{\frac{\log C}{n}} + \delta_n \|\mu_k\|_1 \right) \\
&= O \left( \|\beta^*\|_1 \sqrt{\frac{\rho(\beta^*) \log n \log C}{n}} + \|\beta^*\|_1^2 \sqrt{\frac{\log C (\log n)^2}{n}} + \sqrt{\log n} \|\beta^*\|_1 \delta_n \|\mu_k\|_1 \right) \\
&= O \left( \|\beta^*\|_1^2 \sqrt{\frac{(\log C)^3}{n}} + \sqrt{\log n} \|\beta^*\|_1 \delta_n \|\mu_k\|_1 \right),
\end{aligned}$$

where the last inequalities follows since  $C \geq n$ ,  $\sqrt{\rho(\beta^*)} \stackrel{O}{=} \|\beta^*\|_1$  and  $\log n$  dominates  $\sqrt{\log n}$ . Note that this simplification only gives us the error between  $\hat{Y}_t$  and  $\tilde{Y}_t^{(0)}$ , while we want to finally compute the error between  $\hat{Y}_t$  and  $Y_t^{(0)}$ . Thus, next we consider  $E(Y_t^{(0)} - \tilde{Y}_t^{(0)})^2$  which constitutes the random error on the final prediction.

Note that based on our factor model,  $Y_t^{(0)} - \tilde{Y}_t^{(0)} = \sum_{k=1}^K R_k(\beta^*) \mu_{kt} + \rho(\beta^*) \zeta_t \leq \delta_n \|\mu_k\|_1 + \rho(\beta^*) \zeta_t$ . Thus,  $E(Y_t^{(0)} - \tilde{Y}_t^{(0)})^2 \leq \delta_n^2 \|\mu_k\|_1^2 + \rho(\beta^*)^2$ . Combined with the simplification of  $E(\tilde{Y}_t^{(0)} - \hat{Y}_t)^2$ , this proves our final bound since

$$E \left( Y_t^{(0)} - \hat{Y}_t \right)^2 \leq E \left( \tilde{Y}_t^{(0)} - \hat{Y}_t \right)^2 + E \left( Y_t^{(0)} - \tilde{Y}_t^{(0)} \right)^2.$$

### B.3. Proof of Theorem 3

We first describe some properties pertaining to our choice of prior in our model coefficients  $\beta_c$ . The coefficients are mean zero Gaussian distributed with the variances being controlled by  $\tau$  and  $\lambda_c^2$ . While  $\tau$  is distributed with a Half Cauchy prior,  $\lambda_c$ 's are Logit Normals with hyperparameter as defined Lemma 4.

**Lemma 5** Consider the full prior on  $\beta_c$  described with  $g(\beta_c)$  where,  $[\beta_c | \tau, \lambda_c] \stackrel{i.i.d.}{\sim} N(0, \tau \lambda_c^2)$ ,  $\lambda_c \stackrel{i.i.d.}{\sim} \text{LogitNormal}(-Cf(C), C^2)$ ,  $\tau \sim \text{HalfCauchy}(c_0, \gamma)$ . Then, for  $a_n = \sqrt{\frac{4\gamma(1+u) \log C}{n C^2}}$ ,

$$1 - \int_{-a_n}^{a_n} g(x) dx \leq C^{-(1+u)} \text{ and } -\log \left( \inf_{x \in (-n, n)} g(x) \right) = O(\log C)$$

In addition to the properties on  $g(\beta_c)$ , in our case, we also have,

- 1.The scaled covariates  $X_{ct}/\sqrt{\log n C}$  are bounded by  $M$ .
- 2.The dimensionality is high,  $C \geq n$  and  $l_n \log C \leq n$ .
3. $\max\{|\beta_c^*/\sigma^*|\} \leq c_3 n$  for some fixed  $c_3$ .

The first statement holds since  $X_{ct}$ 's have  $nC$  Gaussian elements whose maximum is of the order  $\sqrt{\log n C}$  and the other two statements follow from Assumption 1. Lemma 5 together with the above three assumptions are similar to the conditions required for Theorem 2.2 in Song and Liang (2023). The only difference being the first assumption, where Song and Liang (2023) assumes bounded covariates rather than the bound on scaled covariates.

Define  $A_n = \{ \text{At least } \tilde{C} \text{ entries of } |\beta/\sigma| \text{ is larger than } a_n \} \cup \{ (\|x'_{st} \beta_s - x'_{st} \mathbf{b}_s^*\| \geq c_1 \sigma^* \epsilon_n) \cup \{ \sigma^2/\sigma^{*2} > (1 + \epsilon_n)/(1 - \epsilon_n) \text{ or } \sigma^2/\sigma^{*2} < (1 - \epsilon_n)/(1 + \epsilon_n) \} \}$ ,  $B_n = \{ \text{At least } \tilde{C} \text{ entries of } |\beta/\sigma| \text{ is larger than } a_n \}$  and  $C_n = A_n \setminus B_n$ , where  $\tilde{C}$  is defined exactly as in Song and Liang (2023),  $\tilde{C} = \lfloor \min\{\hat{c}_3, \tilde{c}_3\} n \epsilon_n^2 / (2 \log C) \rfloor$ .

Following the three-step technique as in Theorem 2.2 in Song and Liang (2023) with the following testing functions completes the proof. Since the proof is similar, the details are omitted. Consider the following two testing functions,

$$\begin{aligned}\phi'_n &= \max_{\{\xi \geq \xi^*, |\xi| \leq \bar{C} + l_n\}} \mathbb{1}\{|\mathbf{y}^T(I - H_\xi)\mathbf{y}/(n - |\xi|)\sigma^{*2} - 1| \geq \epsilon_n\}, \text{ and} \\ \tilde{\phi}_n &= \max_{\{\xi \geq \xi^*, |\xi| \leq \bar{C} + l_n\}} \mathbb{1}\{\|X_\xi(X_\xi^T X_\xi)^{-1} X_\xi^T \mathbf{y} - X_\xi \boldsymbol{\beta}_\xi^*\| \geq c_1 \sigma^* \epsilon_n\},\end{aligned}$$

where  $H_\xi = X_\xi (X_\xi^T X_\xi)^{-1} X_\xi^T$ .

With these testing functions, we finally get the first main result since  $\mathbf{y} := \mathbf{x}'_{st} \boldsymbol{\beta}_s$ ,

$$P\left(\{\pi(\|\mathbf{x}'_{st} \boldsymbol{\beta}_s - \mathbf{x}'_{st} \boldsymbol{\beta}_s^*\| \geq c_1 \epsilon_n | \mathcal{D}_n(s))\} \geq e^{-c_2 n \epsilon_n^2}\right) \leq \exp(-c_3 n \epsilon_n^2). \quad (10)$$

Now based on (4) in Lemma 1, we know that for any time  $t > n$ ,

$$Y_{st}^{(0)} = \mathbf{x}'_{st} \boldsymbol{\beta}_s^* + \sum_{k=1}^K R_{sk}(\boldsymbol{\beta}_s^*) \mu_{kst} + \rho(\boldsymbol{\beta}_s^*) \zeta_{st}.$$

With the decomposition of the counterfactual, it is easy to note that if  $\|\mathbf{x}'_{st} \boldsymbol{\beta}_s - \mathbf{x}'_{st} \boldsymbol{\beta}_s^*\| \leq c_1 \epsilon_n$ , then

$$\begin{aligned}\|\mathbf{y} - Y_{st}^{(0)}\| &\leq \|\mathbf{y} - \mathbf{x}'_{st} \boldsymbol{\beta}_s^*\| + \left\| \sum_{k=1}^K R_{sk}(\boldsymbol{\beta}_s^*) \mu_{kst} \right\| + \|\rho(\boldsymbol{\beta}_s^*) \zeta_{st}\| \\ &\leq c_1 \epsilon_n + \delta_n \|\boldsymbol{\mu}_{st}\|_1 + \rho(\boldsymbol{\beta}_s^*) |\zeta_{st}|.\end{aligned} \quad (11)$$

Consider the set,  $\mathcal{S} := \{s : \pi(\|\mathbf{y} - \mathbf{x}'_{st} \boldsymbol{\beta}_s^*\| \geq c_1 \epsilon_n | \mathcal{D}_n(s)) \leq e^{-c_2 n \epsilon_n^2}\}$ , which gives us  $P(\mathcal{S}) \geq 1 - \exp(-c_3 n \epsilon_n^2)$ . Based on the definition, under the set  $\mathcal{S}$ , we have  $\pi(\|\mathbf{y} - \mathbf{x}'_{st} \boldsymbol{\beta}_s^*\| \geq c_1 \epsilon_n) \leq e^{-c_2 n \epsilon_n^2}$  which is equivalent to  $\pi(\|\mathbf{y} - \mathbf{x}'_{st} \boldsymbol{\beta}_s^*\| \leq c_1 \epsilon_n) \geq 1 - e^{-c_2 n \epsilon_n^2}$ . As shown above in (11), since  $\|\mathbf{y} - \mathbf{x}'_{st} \boldsymbol{\beta}_s^*\| \leq c_1 \epsilon_n$  implies  $\|\mathbf{y} - Y_{st}^{(0)}\| \leq c_1 \epsilon_n + \delta_n \|\boldsymbol{\mu}_{st}\|_1 + \rho(\boldsymbol{\beta}_s^*) |\zeta_{st}|$ , under  $\mathcal{S}$ , we have

$$\pi(\|\mathbf{y} - Y_{st}^{(0)}\| \leq c_1 \epsilon_n + \delta_n \|\boldsymbol{\mu}_{st}\|_1 + \rho(\boldsymbol{\beta}_s^*) |\zeta_{st}|) \geq 1 - e^{-c_2 n \epsilon_n^2}.$$

Finally, since  $\zeta_{st}$  is standard Gaussian, hence  $P(|\zeta_{st}| \leq c_5) \geq 1 - 2 \exp(-c_5^2/2)$ . Combining this with the above bound on the posterior distribution finally gives us that under  $\mathcal{S}$  and subsequently with probability  $1 - \exp(-c_3 n \epsilon_n^2)$ ,

$$\pi(\|\mathbf{y} - Y_{st}^{(0)}\| \geq c_1 \epsilon_n + \delta_n \|\boldsymbol{\mu}_{st}\|_1 + \rho(\boldsymbol{\beta}_s^*) c_5 | \mathcal{D}_n(s)) \leq \exp(-c_2 n \epsilon_n^2) + 2 \exp(-c_5^2/2),$$

The trimmed mean result follows directly by considering the  $2(\exp(-c_2 n \epsilon_n^2) + 2 \exp(-c_5^2/2))$  trimming of the posterior distribution in the set  $\mathcal{S}$ .

#### B.4. Proof of Lemma 4

We have  $X \sim \text{LogitNormal}(-tf(t), t^2)$ , and  $\exp(f(t)^2/2)f(t) = t^{(1+u)}$ .

$$\begin{aligned}P[X \in (0, t^{-1/2})] &= P[\log X \in (-\infty, -0.5 \log t)] \\ &\approx P[\text{logit}(X) \in (-\infty, -0.5 \log t)] \\ &= P\left[\frac{\text{logit}(X) + tf(t)}{t} \in \left(-\infty, -\frac{\log t}{2t} + f(t)\right)\right] \\ &= \Phi\left(-\frac{\log t}{2t} + f(t)\right) \approx \Phi(f(t))\end{aligned}$$

The second inequality holds, since for large  $t$ , in the set  $\Theta = (0, t^{-1/2})$ ,  $\theta \in \Theta$   $\text{logit}(\theta) = \log(\theta) - \log(1 - \theta) \approx \log(\theta)$  as  $\log(1 - \theta)$  is close to 0 since  $\theta$  as  $t^{-1/2}$  is small. The last step follows as  $\text{logit}(X) \sim \mathcal{N}(-tf(t), t^2)$  and consequently  $t^{-1}(\text{logit}(X) + tf(t))$  is standard normal.

Using Mill's inequality,  $1 - \Phi(f(t)) = O(\exp(-f(t)^2/2)/f(t)) = O(t^{-(1+u)})$  or  $1 - \Phi(f(t)) = k_1 t^{-(1+u)}$  for some constant  $k_1$  which proves the required.

The second one follows similarly as  $P[X \in (1 - t^{-1/2}, 1)] = \Phi\left(\frac{\log t}{2t} + f(t)\right) \approx \Phi(f(t))$ .

### B.5. Proof of Lemma 5

For any set of hyperparameters  $\tau$ ,  $\lambda_c$ , the  $\beta_c$  is zero mean Gaussian distribution. Thus  $\inf_{x \in (-n, n)} g(x)$  is  $g(n)$  since the marginal distribution is symmetric around zero and keeps decreasing as  $x$  increases.

Consider the set  $\mathcal{A} = \{\tau \geq \gamma C^2 + c_0\} \cap \{\lambda_c \geq 1 - C^{-1/2}\}$ . In this set  $\mathcal{A}$ , variance  $\tau \lambda_c^2 \geq (\gamma C^2 + c_0)(1 - C^{-1/2})^2 \geq \gamma C^2/2$  for  $C > 1$ . Under this set  $\mathcal{A}$ ,  $g(n) \geq (C\sqrt{\pi})^{-1} C \exp(-n^2/\gamma C^2)$ . Also the probability  $P(\mathcal{A}) = P(\tau \geq \gamma C^2 + c_0)P(\lambda_c \geq 1 - C^{-1/2}) \approx (2/\pi C^2)C^{-(1+u)} = 2C^{-(3+u)}/\pi$  since  $\tau$  and  $\lambda_c$  are independent.

Using the fact that  $P(\mathcal{A}) \approx 2C^{-(3+u)}$  and  $g(n|\mathcal{A}) \geq (C\sqrt{\pi})^{-1} \exp(-n^2/\gamma C^2)$ , thus the marginal distribution  $g$  is such that

$$\begin{aligned} g(n) &\geq P(\mathcal{A}g(n|\mathcal{A})) \approx \frac{2}{\sqrt{\pi}} C^{-(4+u)} \exp(-n^2/\gamma C^2) \\ \implies -\log(g(n)) &\leq \frac{n^2}{\gamma C^2} + (4+u) \log C - \log(2/\sqrt{\pi}) = O(\log C), \end{aligned}$$

which follows since  $C \geq n$ .

For the second inequality consider the set  $\mathcal{B} = \{\tau \leq c_0 + \gamma C^{1+u}\} \cap \{\lambda_c \leq C^{-(4+u)/2}\}$ . The probability of this set,  $P(\mathcal{B}) = P(\tau \leq c_0 + \gamma C^{1+u})P(\lambda_c \leq C^{-(4+u)/2})$ . Consider  $P(\lambda_c \leq C^{-(4+u)/2})$ ,

$$\begin{aligned} P[\lambda_c \leq C^{-(4+u)/2}] &= P[\log \lambda_c \leq -(4+u) \log C/2] \approx P[\text{logit}(\lambda_c) \leq -(4+u) \log C/2] \\ &= P\left[\frac{\text{logit}(\lambda_c) + C f(C)}{C} \leq -\frac{(4+u) \log C}{2C} + f(C)\right] \\ &= \Phi\left(-\frac{(4+u) \log C}{2C} + f(C)\right) \approx \Phi(f(C)) = 1 - c_0 C^{-(1+u)}, \end{aligned}$$

for some constant  $c_0$  and, where the last inequality follows from Mill's inequality and  $\exp(f(C)^2/2)f(C) = t^{(1+u)}$ . Also,  $P(\tau \leq c_0 + \gamma C^{1+u}) \approx 1 - 2C^{-(1+u)}/\pi$ . Thus,

$$\begin{aligned} P(\mathcal{B}) &= P(\tau \leq c_0 + \gamma C^{1+u})P(\lambda_c \leq C^{-(4+u)/2}) \\ &= (1 - c_0 C^{-(1+u)})(1 - 2C^{-(1+u)}/\pi) \approx 1 - (2/\pi + c_0)C^{-(1+u)}. \end{aligned}$$

Also, in this set  $\mathcal{B}$ , the marginal variance of  $\beta_c$  is  $\tau \lambda_c^2 \leq (c_0 + \gamma C^{1+u})C^{-(4+u)} \leq 2\gamma C^{-3}$ . Thus, if  $(\tau, \lambda_c) \in \mathcal{B}$ , then  $\int_{a_n}^{\infty} g(x)dx \leq \int_{a_n}^{\infty} f(x)dx$ , where  $f$  is density of Gaussian distribution with variance  $2\gamma C^{-3}$ .

$$\int_{a_n}^{\infty} g(x)dx \leq \int_{a_n}^{\infty} f(x)dx = O(\exp(-a_n^2/2\sigma^2)/(a_n/\sigma))$$

where  $\sigma = \tau \lambda_c^2$ . Since  $a_n^2/\tau \lambda_c^2 \geq (4\gamma(1+u) \log C/nC^2)(2\gamma C^{-3})^{-1} \geq 2(1+u) \log C$ , as  $C \geq n$ . Thus  $O(\exp(-a_n^2/2\sigma^2)/(a_n/\sigma)) = O(C^{-(1+u)}/\sqrt{\log C}) \leq C^{-(1+u)}$  for large enough  $C$ , which completes the proof.

## 6. Competing interests

No competing interest is declared.

## 7. Author contributions statement

D.P. and I.B. conducted the data analysis. R.R.B. and G.M. developed the statistical methodology and provided theoretical proofs related to the synthetic control approach. S.D., J.M., and R.M. were responsible for data extraction and preprocessing. D.P., R.R.B., S.D., and G.M. contributed to the writing of the manuscript.

## 8. Acknowledgments

The authors acknowledge the Center for Advanced Research Computing (CARC) [<https://carc.usc.edu>] at the University of Southern California for providing computing resources that have contributed to the research results reported within this publication.

## References

- A. Abadie. Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2):391–425, 2021.
- A. Abadie and J. Gardeazabal. The economic costs of conflict: A case study of the basque country. *American economic review*, 93(1):113–132, 2003.
- A. Abadie and J. L’hour. A penalized synthetic control estimator for disaggregated data. *Journal of the American Statistical Association*, 116(536):1817–1834, 2021.
- A. Abadie, A. Diamond, and J. Hainmueller. 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, 2010.
- M. Amjad, D. Shah, and D. Shen. Robust synthetic control. *Journal of Machine Learning Research*, 19(22):1–51, 2018.
- J. Antonelli, G. Parmigiani, and F. Dominici. High-dimensional confounding adjustment using continuous spike and slab priors. *Bayesian analysis*, 14(3):805, 2019.
- M. L. Barnett, M. Gaye, A. B. Jena, and A. Mehrotra. Association of county-level prescriptions for hydroxychloroquine and ivermectin with county-level political voting patterns in the 2020 us presidential election. *JAMA Internal Medicine*, 182(4):452, 2022.
- E. Ben-Michael, A. Feller, and J. Rothstein. The augmented synthetic control method. *Journal of the American Statistical Association*, 116(536):1789–1803, 2021.
- E. Ben-Michael, A. Feller, and J. Rothstein. Synthetic controls with staggered adoption. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(2):351–381, 2022.
- A. Bhadra, J. Datta, N. G. Polson, and B. Willard. Default bayesian analysis with global-local shrinkage priors. *Biometrika*, 103(4):955–969, 2016.
- A. Bhadra, J. Datta, Y. Li, N. G. Polson, and B. Willard. Prediction risk for the horseshoe regression. *Journal of Machine Learning Research*, 20(46):1–43, 2019.
- A. Bhattacharya, D. Pati, N. S. Pillai, and D. B. Dunson. Dirichlet-laplace priors for optimal shrinkage. *Journal of the American Statistical Association*, 110(512):1479–1490, 2015.
- L. Bonis and M. Curtis. Senate hearing to discuss therapies to fight covid-19 turns political. <https://local12.com/health/medical-edge-reports/senate-hearing-to-discuss-therapies-to-fight-covid-19-turns-political-cincinnati>, 2020. Accessed: 2024-05-11.
- I. J. Borges Nascimento, A. B. Pizarro, J. Almeida, N. Azzopardi-Muscat, M. A. Gonçalves, M. Björklund, and D. Novillo-Ortiz. Infodemics and health misinformation: a systematic review of reviews. *Bulletin of the World Health Organization*, 100(9):544–561, 2022.
- M. Bray, C. Rayner, F. Noël, D. Jans, and K. Wagstaff. Ivermectin and covid-19: a report in antiviral research, widespread interest, an fda warning, two letters to the editor and the authors’ responses. *Antiviral research*, 178:104805, 2020.
- C. Carvalho, R. Masini, and M. C. Medeiros. Arco: An artificial counterfactual approach for high-dimensional panel time-series data. *Journal of econometrics*, 207(2):352–380, 2018.
- I. Castillo and R. Mismar. Empirical bayes analysis of spike and slab posterior distributions. *Electronic Journal of Statistics*, 2018.
- S. Chatterjee. Assumptionless consistency of the lasso. *arXiv preprint arXiv:1303.5817*, 2013.
- Q. Chen and W. Li. On the sparsity of synthetic control method. *Journal of Applied Economics*, 27(1):2361184, 2024.
- J. Christensen. Cdc warns against use of anti-parasitic drug ivermectin for covid-19, as calls to poison control centers increase. <https://www.cnn.com/2021/08/26/health/ivermectin-covid-19-warning-injuries/index.html>, 2021. Accessed: 2024-05-11.
- N. Doudchenko and G. W. Imbens. Balancing, regression, difference-in-differences and synthetic control methods: A synthesis. Technical report, National Bureau of Economic Research, 2016.
- U. K. H. Ecker, S. Lewandowsky, and D. T. W. Tang. Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition*, 38(8):1087–1100, 2010.

- U. K. H. Ecker, S. Lewandowsky, J. Cook, P. Schmid, L. K. Fazio, N. Brashier, P. Kendeou, E. K. Vraga, and M. A. Amazeen. The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1(1):13–29, 2022.
- S. R. Esposito, M. J. Hornsey, and J. R. Spoor. Shooting the messenger: Outsiders critical of your group are rejected regardless of argument quality. *British Journal of Social Psychology*, 52(2):386–395, 2013.
- J. Fong, T. Guo, and A. Rao. Debunking misinformation about consumer products: Effects on beliefs and purchase behavior. *Journal of Marketing Research*, 2022. 002224372211470.
- J. Fuhrer and F. Cova. “quick and dirty”: Intuitive cognitive style predicts trust in didier raoult and his hydroxychloroquine-based treatment against covid-19. *Judgment and Decision Making*, 15(6):889–908, 2020.
- A. Hill, M. Mirchandani, L. Ellis, and V. Pilkington. Ivermectin for the prevention of covid-19: addressing potential bias and medical fraud. *Journal of Antimicrobial Chemotherapy*, 77(5):1413–1416, 2022.
- G. W. Imbens. Causal inference in the social sciences. *Annual Review of Statistics and Its Application*, 11, 2024.
- G. W. Imbens and D. B. Rubin. *Causal inference in statistics, social, and biomedical sciences*. Cambridge university press, 2015.
- H. Ishwaran and J. S. Rao. Spike and slab variable selection: frequentist and bayesian strategies. *Annals of Statistics*, 2005.
- B. Karmakar, G. Mukherjee, and W. Kar. Using penalized synthetic controls on truncated data: A case study on effect of marijuana legalization on direct payments to physicians by opioid manufacturers. *Journal of the American Statistical Association*, 2024.
- A. P. Kautsar, R. K. Sinuraya, J. van der Schans, M. J. Postma, and A. A. Suwantika. Exploring public sentiment on the repurposing of ivermectin for covid-19 treatment: Cross-sectional study using twitter data. *JMIR Formative Research*, 9(1):e50536, 2025.
- S. Kim, C. Lee, and S. Gupta. Bayesian synthetic control methods. *Journal of Marketing Research*, 57(5):831–852, 2020.
- S. Lewandowsky, U. K. H. Ecker, C. M. Seifert, N. Schwarz, and J. Cook. Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3):106–131, 2012.
- G. Malsiner-Walli and H. Wagner. Comparing spike and slab priors for bayesian variable selection. *arXiv preprint arXiv:1812.07259*, 2018.
- J. W. Miller and M. T. Harrison. Shrink globally, act locally: Sparse bayesian regularization and prediction. *Journal of the American Statistical Association*, 113(524):1691–1703, 2018.
- R. H. Perlis, K. L. Trujillo, J. Green, A. Safarpour, J. N. Druckman, M. Santillana, K. Ognyanova, and D. Lazer. Misinformation, trust, and use of ivermectin and hydroxychloroquine for covid-19. In *JAMA Health Forum*, volume 4, pages e233257–e233257. American Medical Association, 2023.
- V. Ročková and E. I. George. The spike-and-slab lasso. *Journal of the American Statistical Association*, 113(521):431–444, 2018.
- M. S. Rockwell, S. Vangala, M. Hadfield, J. Cantor, D. Skinner, M. Craff, A. M. Fendrick, C. L. Damberg, K. Kahn, and J. N. Mafi. Demographic variation in us outpatient hydroxychloroquine and ivermectin use during the covid-19 pandemic: Article examines demographic variation in us outpatient hydroxychloroquine and ivermectin use during the covid-19 pandemic. *Health Affairs*, 44(3):246–255, 2025.
- V. Ročková and E. I. George. Spike-and-slab meets lasso: A review of the spike-and-slab lasso. *Statistical Science*, 33(4):485–514, 2018.
- J. Salvatier, T. V. Wiecki, and C. Fonnesbeck. Probabilistic programming in python using pymc3. *PeerJ Computer Science*, 2:e55, 2016.
- Q. Song and F. Liang. Nearly optimal bayesian shrinkage for high-dimensional regression. *Science China Mathematics*, 66(2):409–442, 2023.
- National Institute of Health. The covid-19 treatment guidelines panel’s statement on the use of ivermectin for the treatment of covid-19. <https://files.covid19treatmentguidelines.nih.gov/guidelines/archive/statement-on-ivermectin-01-14-2021.pdf>, 2021. Accessed: 2024-05-11.
- New York Times. 2016 presidential election results. <https://www.nytimes.com/elections/2016/results/president>, 2016. Accessed: 2024-05-11.
- New York Times. Coronavirus (covid-19) data in the united states. <https://www.nytimes.com/interactive/2021/us/covid-cases.html>, 2023. Accessed: 2024-05-11.
- Pew Research Center. The partisanship and ideology of american voters, 2024. URL <https://www.pewresearch.org/politics/2024/04/09/the-partisanship-and-ideology-of-american-voters/>. Accessed: 2024-08-12.
- W. Thomson, S. Jabbari, A. E. Taylor, W. Arlt, and D. J. Smith. Simultaneous parameter estimation and variable selection via the logit-normal continuous analogue of the spike-and-slab prior. *Journal of The Royal Society Interface*, 16(150):20180572, 2019.

- 
- R. Tibshirani and L. Wasserman. Sparsity, the lasso, and friends. *Lecture notes from "Statistical Machine Learning," Carnegie Mellon University, Spring, 2017.*
- M. Vaduganathan, J. van Meijgaard, M. R. Mehra, J. Joseph, C. J. O'Donnell, and H. J. Warraich. Prescription fill patterns for commonly used drugs during the covid-19 pandemic in the united states. *JAMA*, 323(24):2524, 2020.