Assessing COVID19 Misinformation Effects on Ivermectin Consumption Using Regularized Synthetic Controls

Dinesh Puranam*, Ivan Belov*, Rashmi Ranjan Bhuyan*,
Shantanu Dutta*, Jeroen van Meijgaard†, Reetabrata Mookherjee†
and Gourab Mukherjee*
*University of Southern California and †GoodRx

December 19, 2024

Abstract

Misinformation has the potential to adversely impact public health by significantly influencing the behavior of healthcare professionals. During the COVID-19 pandemic, public healthcare systems experienced substantial disruptions due to various sources of misinformation. Here, we rigorously study the effects of misinformation in a causal framework and assess its persistence after countermeasures were initiated. We analyze weekly prescription claims data in the US over a 36-month period coinciding with the COVID-19 pandemic to thoroughly investigate the impact of misinformation on the use of the medication Ivermectin (IVM). Using continuous spike-and-slab shrinkage prior based regularized synthetic controls (SC), we estimate state-wise elevated claims of IVM due to misinformation. The proposed method enjoys optimal decision-theoretic risk guarantees, and its robustness is empirically validated through extensive checks.

Comparing IVM prescription claims with the SC based counterfactual we witness a modest, short-lived increase in IVM consumption when early interest in it as a potential COVID-19 treatment started. There was no significant surge for the next 8 months. The peak surge in aggregate US consumption was 16 times above the counterfactual and occurred when 67% of the US population had already received one dose of COVID-19 vaccine and several institutionalized counter measures to debunk IVM related misinformation were already implemented. We found that political affiliation of states significantly explains the variations in the impact of misinformation across regions, even when controlling for differences in COVID-19 case numbers. These insights underscore the necessity for targeted intervention strategies to mitigate the enduring effects of misinformation on healthcare decision-making.

Keywords: misinformation, synthetic control, spike-and-slab priors, susceptibility, persistence, prescription drug.

1 Introduction

Misinformation (Fetzer, 2004) presents a significant threat, leading to a false sense of understanding that can have serious consequences for individuals and society as a whole (Guess and Lyons, 2020). The advent of technology has amplified these concerns, spreading misinformation across various fields, including politics, science, education, and healthcare (Ecker et al., 2022, Steen, 2011, West and Bergstrom, 2021). Misinformation can arise from misunderstandings, misreporting, or the misinterpretation of facts, and it can spread rapidly, especially through social media and other digital communication channels (Allcott et al., 2019, Muhammed and Mathew, 2022). In the realm of healthcare, misinformation becomes especially perilous, as it jeopardizes both individual well-being and community health (Chou et al., 2018, Rubin, 2022). Recognizing the gravity of this issue, the United States Surgeon General and the World Health Organization have prioritized combating healthcare misinformation and have actively promoted various debunking efforts and countermeasures (Mendez, 2021). Well-established evidentiary standards help identify factually inaccurate claims (Voss, 2020), and regulatory bodies utilize these standards to assess health claims concerning drugs and therapies. Despite these fact-checking measures, the influence of misinformation often persists in the healthcare system, making it crucial to detect, estimate, and analyze its impact.

Misinformation in healthcare surged dramatically during the COVID-19 pandemic (Kouzy et al., 2020, Naeem and Boulos, 2021). Recent studies document anecdotal and experimental evidence that several medications received increased attention based on misbeliefs propelled by social media campaigns and early-stage studies that were later debunked (Fuhrer and Cova, 2020, Roozenbeek et al., 2020). In this paper, we examine misinformation concerning the medication Ivermectin (IVM) and its effect on IVM consumption

in the USA. During the COVID-19 pandemic, IVM, a medication primarily used to treat parasitic infections in animals and humans, became a focal point of significant misinformation regarding its effectiveness against COVID-19. Short-term studies such as Barnett et al. (2022), Vaduganathan et al. (2020), as well as media reports[1], suggest that the consumption of IVM increased above prior year levels after particular misinformation events.

For an accurate scientific analysis of the phenomenon, it is crucial to provide rigorous statistical analysis (Boukouvalas and Shafer, 2024) and precisely quantify the effects of misinformation on IVM consumption. To assess the causal impact of misinformation on IVM consumption, we adopt a quasi-experimental approach by analyzing prescription claims (Abadie, 2021). Prescription claims serve as a proxy for drug consumption and have been widely employed in previous studies (Vaduganathan et al., 2020). To establish a baseline for IVM prescription claims, we compare them to prescription claims for other drugs which did not face misinformation; they serve as our control group. This approach enables us to account for various confounding factors, including seasonality, trends, and COVID-related measures such as shelter-in-place orders, indoor mask mandates, and travel restrictions.

To causally estimate the incremental change in prescription claims for IVM, we employ the widely used synthetic control (SC) methodology (Abadie et al., 2010, Abadie and Gardeazabal, 2003), using prescription claims of drugs not impacted by misinformation as controls. SC methods have been successfully applied to various contemporary scientific problems to produce efficient counterfactual estimates (see Abadie (2021) and the references therein). Considering misinformation as the treatment in this context, the difference between the observed claims of IVM and the counterfactual estimates represents the incremental change due to misinformation, or the average treatment effect on the treated (ATT). The ATT reflects the 'net effect', i.e., the incremental change in IVM claims result-

[1] `https://www.nytimes.com/2021/08/30/health/covid-ivermectin-prescriptions.html`

ing from misinformation and the subsequent countermeasures implemented at both state and federal levels.

The widely used SC criterion of Abadie et al. (2010) cannot be directly applied for estimating the state wise ATTs for IVM as we have more controls than the number of pre-treatment observation time points. Carvalho et al. (2018), Doudchenko and Imbens (2016) advocate for the use of statistical shrinkage to produce robust SC estimates in such scenarios. The role of shrinkage in robust SC methods has been both theoretically and empirically studied in Abadie and L'hour (2021), Amjad et al. (2018), Ben-Michael et al. (2021, 2022), Kim et al. (2020).

Specifically, Ben-Michael et al. (2021) demonstrated that the predictive risk of shrinkage-based SC estimates is well-controlled, and the necessity of shrinkage for robust SC estimation is further established in Karmakar et al. (2024). 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 misinformation effect, but also in conducting further downstream analysis to understand (a) the persistence of misinformation despite debunking efforts at both federal and state levels in the US, and (b) the degree of susceptibility among different groups in the US.

In this context, we employ a spike-and-slab shrinkage prior (Antonelli et al., 2019, Ishwaran and Rao, 2005, Malsiner-Walli and Wagner, 2018, Ročková and George, 2018) based Bayesian SC (BSC) method to estimate the state-wise ATT. We demonstrate the optimal predictive risk properties of the ATT estimates, thereby validating their suitability for conducting accurate downstream analysis. Next, we describe the background of misinformation surrounding IVM and then present further details on the proposed BSC method and the downstream analyses.

1.1 IVM Misinformation: Background and Prescription Claims

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 weak and eventually discredited. The countermeasures for IVM 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

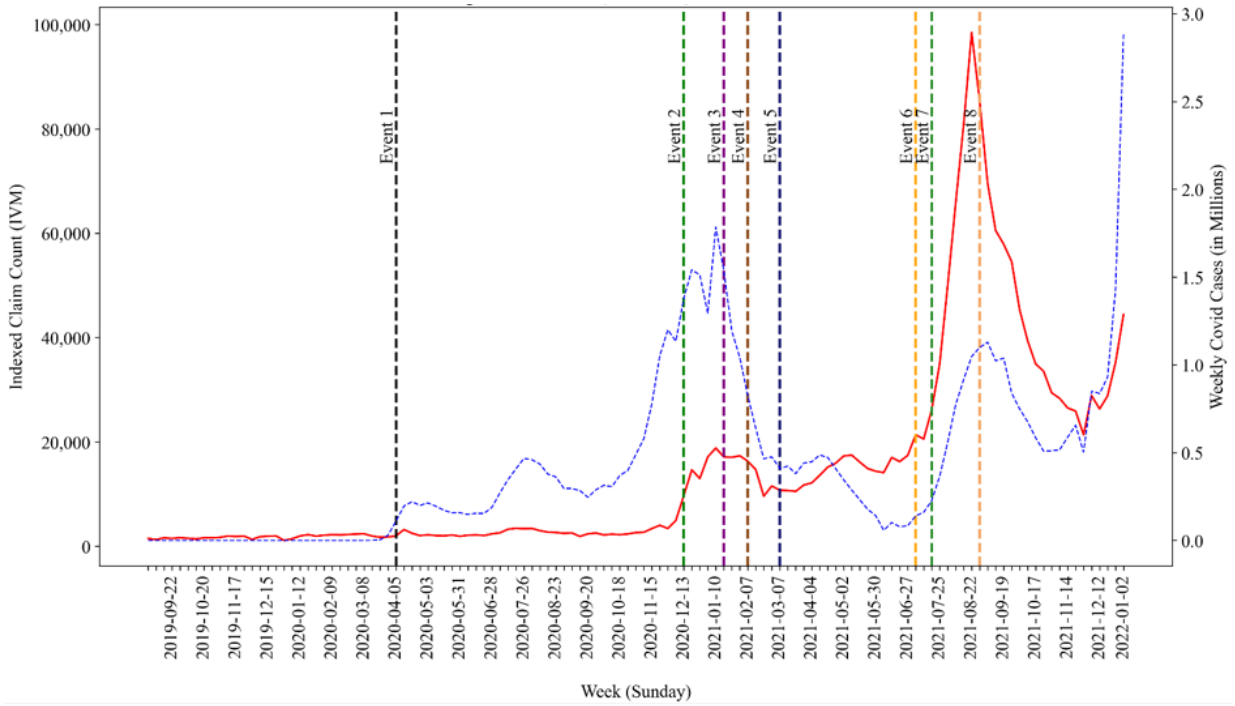

Figure 1: Plot of total US (indexed) IVM prescription claims (in red) and Covid case counts (in dotted blue) at weekly resolution. The weeks are indexed by their start date on Sundays. The time points marked by dotted vertical lines correspond to major related events listed in Table 1.

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 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[2], and the initial

[2]https://www.cbsnews.com/news/biden-covid-19-vaccine-goal-missed/

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 regulations on observed changes in IVM prescription claims, while accounting for the effects of confounding factors.

1.2 Causal Impact of IVM Misinformation

We use the synthetic control (SC) method to estimate the net changes in IVM prescription claims due to misinformation. By employing shrinkage priors, we derive robust average treatment effect on the treated (ATT) estimates based on control drugs. We then compare the patterns observed in the ATTs over the next 18 months with the findings of Barnett et al. (2022), Vaduganathan et al. (2020), who conducted very short-term studies. Using the estimated ATTs, we address two important scientific questions: (a) the persistence of misinformation after countermeasures, and (b) the degree of susceptibility among states to misinformation. These aspects are detailed below.

Measuring Persistence of Misinformation Effects. We focus on whether the effects of misinformation on IVM consumption persisted after countermeasures were deployed. For IVM, the primary countermeasures were informative measures (Borges Nascimento et al., 2022). The existing literature lacks consensus on the role of informative countermeasures. For example, Fong et al. (2022) demonstrates through experiments that informative countermeasures, especially those from regulators, can mitigate misinforma-

tion. An important finding is that the source of countermeasures can moderate their effectiveness. In contrast, other studies suggest that misinformation can persist even after informative countermeasures are implemented (Ecker et al., 2022, 2010, Lewandowsky et al., 2012). In some domains, such as politics, misinformation can persist or even increase after informative countermeasures are in place (Nyhan, 2021, Nyhan and Reifler, 2010). We empirically study the changes in ATTs as different countermeasures were announced to understand their impact on the persistence of misinformation.

Explaining Susceptibility to Misinformation. Individuals are more susceptible to influence from those they perceive as similar or belonging to their social group (Mackie et al., 1990, Pennycook and Rand, 2019, Roozenbeek et al., 2020, Scherer et al., 2021). As IVM gained substantial attention as a potential treatment for COVID-19 in conservative media, we focus on studying whether the political affiliation of a state can explain its degree of susceptibility to misinformation. To examine the impact of susceptibility on the effects of misinformation and countermeasures, we analyze state-level results from the 2016 presidential election, focusing on each state's conservativeness index. Conservativeness generally refers to a political and social ideology that values tradition, cultural norms rooted in history and religion, and is typically associated with right-leaning policies that emphasize national security and fiscal conservatism. Consistent with common practice (Pew Research Center, 2024), we define a state's conservativeness index as the ratio of the Republican vote share to that of the closest competitor in that state, as is commonly done. In Figure 2, we present the per capita IVM consumption averaged across states that voted for the Republican party (henceforth called *Red* states) and the Democratic party (we call them *Blue* states) in 2016. Note that while the republican voting states has much higher average than the democratic voting states (the scales of the two series are different), the location of the peaks in the two series roughly coincide but their association with Covid

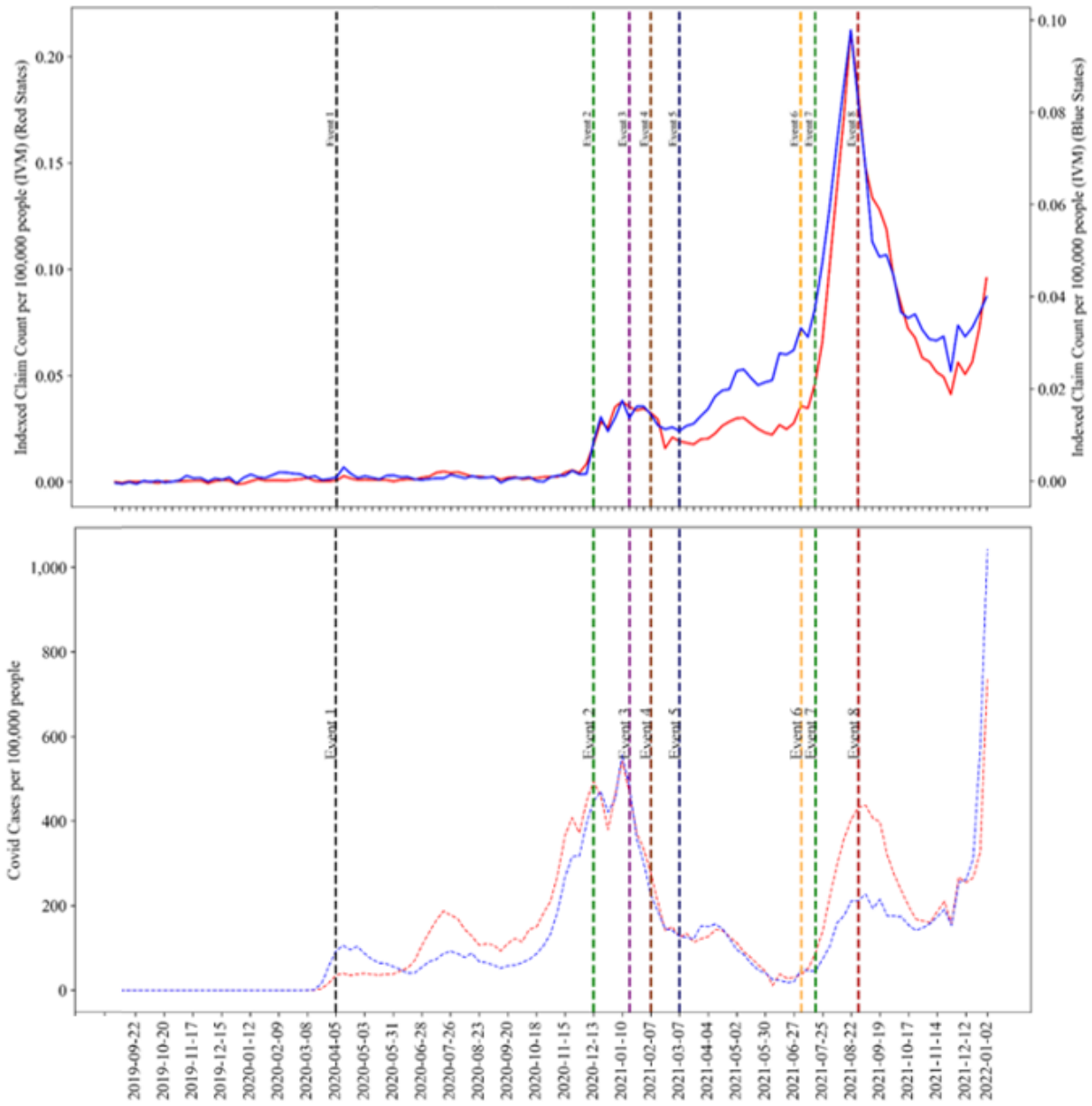

Figure 2: Plot of weekly per capita IVM prescription claims (indexed) averaged across states that voted *Republican* (in red) and those voted *Democratic* (in blue) in the 2016 US presidential election. The time points marked by dotted vertical lines correspond to major related events listed in Table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

case counts is more nuanced.

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

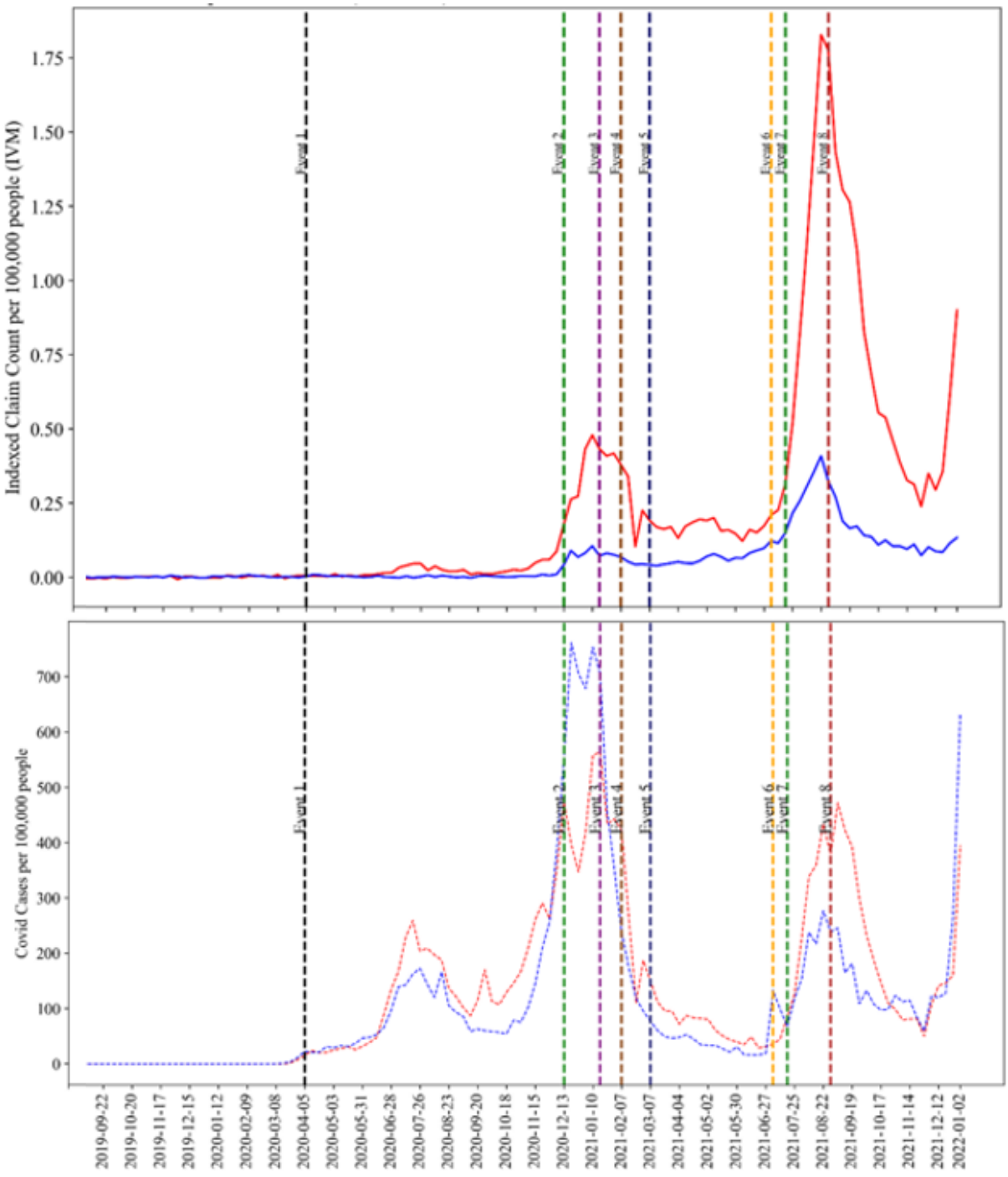

Figure 3: Plot of weekly per capita IVM prescription claims (top) and Covid counts (bottom) in California (in blue) and Texas (in red).

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 Ivermectin 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, such as political attitudes, demographic differences, or state-specific healthcare policies, warranting further investigation into the reasons behind the higher IVM utilization in Texas.

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 degree of susceptibility based on conservativeness. Demographic characteristics such as race, age, income, and education could also be considered as interesting metrics for this study. They are strongly correlated with political affiliation. We focus on the conservativeness metric throughout this paper.

Brief Summary. We describe the **main contributions** of our work below.

- We use shrinkage prior-based regularized synthetic control (RSC) method to estimate the average treatment effects on the treated (ATTs). Specifically, we employ logit-normal distribution-based continuous shrinkage priors (Thomson et al., 2019). The proposed RSC method (see Sec 3) is computationally efficient and facilitates straightforward Bayesian inference for ATTs, aiding in the subsequent persistence and susceptibility analysis. We demonstrate that the ATT estimates produced by the proposed method exhibit well-controlled prediction error (see Sec 3.1). Consequently, multi-step ahead ATT estimates can be consistently utilized for our inference.

- We report state-level average treatment effects on the treated (ATTs) estimates of increased IVM consumption (net effect) due to misinformation for 72 subsequent

weeks. The US aggregate net effect showed a significant surge following Dr. Kory's Senate hearing and persisted thereafter, with major peaks in July 2021 coinciding with a spike in COVID-19 cases in the US. 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.

- Our research focuses on the persistence of misinformation effects on a medication that requires a prescription, a context where the stakes are very high. Existing research on countermeasures against misinformation primarily focuses on consumer products and over-the-counter medications (Fong et al., 2022, Rao and Wang, 2017). We also examine the medium-term effects of misinformation, observing prescription claims up to twenty months after misinformation regarding IVM began. In contrast, existing studies (Barnett et al., 2022, Vaduganathan et al., 2020) on misinformation have only studied short-term impacts on prescription drugs. We found that even after the studies that provoked initial interest in IVM as a treatment option were discredited, IVM continued to be prescribed at high levels. Notably, IVM prescriptions increased by an astounding 1617.64% after COVID-19 vaccines became widely available. See Sec. 4.2.

- We examine whether susceptibility to conservative politics moderates the effect of countermeasures. Our findings indicate that available evidence 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.

1.3 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 misinformation 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 supplement.

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 after event 8, i.e. after CDC's warning on IVM usage.

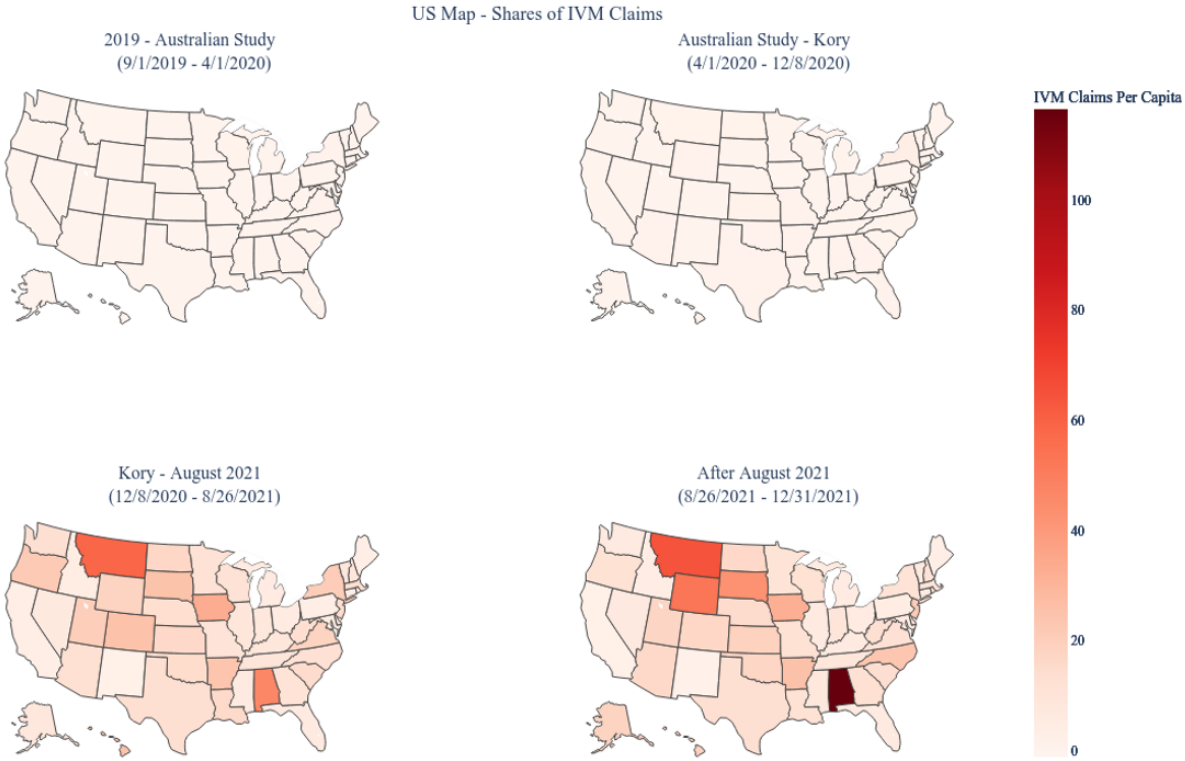

Figure 4: Maps of average weekly IVM prescription claims for US states across four time intervals based on Table 1. From top-left, row wise we have maps for (a) Before Event 1 (b) Event 1 to Event 2 (c) Event 2 to Event 8 (d) Event 8 to Event 9.

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 having GPI2 identifiers. From the second group, we aggregate claims for each GPI2 code. Thus, our control group of medications

consists of COVID-19 drugs and GPI2 code categories. We exclude Chloroquine drugs from our control group as they were also subject to misinformation 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).

3 Methodology

3.1 Regularized SC: Estimating State-level ATTs

We apply regularized synthetic control to estimate counterfactual number of claims for IVM in case the misinformation campaign had not taken place and compare it to the actual

number of claims. Let X_{cst} denotes the prescription claims of control drug $c = 1, \ldots, C$ at state $s = 1, \ldots, S$ and time $t = 1, \ldots, N$ 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 $s = 1, \ldots, S$ and time $t = 1, \ldots, n$ weeks. After the intervention, we denote the observed IVM prescription claims by $Y_{st}^{(1)}$. For all $t > n$ and $s = 1, \ldots, S$, 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: $\widehat{\mathrm{ATT}}_{st} = Y_{st}^{(1)} - \hat{Y}_{st}^{(0)}$.

To estimate $\hat{Y}_{st}^{(0)}$, we consider approximating the observed pre-treatment IVM claims $\{Y_{st} : s = 1, \ldots, S; t = 1, \ldots, n\}$ by linear combinations of the controls as,

$$Y_{st} = \beta_{0s} + \sum_{c=1}^{C} \beta_{cs} X_{cst} + \epsilon_{st}, \tag{1}$$

where ϵ_{st} are independent and identically distributed (i.i.d.) from a normal distribution with mean 0 and variance σ^2. If $C < n$, the traditional (unregularized) SC estimates would be $\hat{Y}_{st}^{(0)}[\mathtt{un}] = \boldsymbol{x}_{st}' \hat{\boldsymbol{\beta}}_s^{(\mathtt{un})}$ where $\hat{\boldsymbol{\beta}}_s^{(\mathtt{un})}$ is the ordinary least squares estimates in (1) and $\boldsymbol{x}_{st} = (1, x_{1st}, \ldots, x_{Cst})'$. However, here we use a large set of control units and $C \gg n$. By utilizing a large set of control units that includes Covid-19 drugs and GPI2 level medications, we implicitly control for several confounding factors. These factors include platform-specific effects such as market penetration for GoodRx, seasonality, trends, and Covid-related measures such as shelter-in-place orders, indoor mask mandates, and travel restrictions. We make the reasonable assumption that these Covid measures affect all drug claims similarly, except in cases where patients can reasonably be safe without the prescription medication. Importantly, by using Covid-19 drugs in the control pool, we control for shifts in the pandemic's intensity.

As $C \gg n$, we run into overfitting issues (Carvalho et al., 2018, Doudchenko and Imbens, 2016) with the unregularized SC estimates in (1), which can lead to very high prediction error of the resultant ATT estimates. As we use multi-step ahead ATT estimates for inference it is extremely important to have adequate control on their prediction error. For the purpose, we consider shrinkage priors on the $\boldsymbol{\beta}_s = \{\beta_{cs} : c = 0, \ldots, C\}$ coefficients. Following Thomson et al. (2019) we consider an exchangeable hierarchical structure: for $c = 1, \ldots, C$ and $s = 1, , S$:

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

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

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

Note that, the shrinkage of $|\beta_{cs}|$ towards 0 is monotone in z_{cs}. As $z_{cs} \downarrow -\infty$, $\beta_{cs} \to 0$ where as as $z_{cs} \uparrow \infty$, $\beta_{cs} \to N(0, \tau)$, where, τ is the global shrinkage factor. This Logit-Normal prior (LNP) structure is a global-local continuous shrinkage prior (Bhadra et al., 2016, Bhattacharya et al., 2015, Miller and Harrison, 2018) where τ controls the overall level of shrinkage and the local shrinkage parameters λ_{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 Mismer, 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 σ, τ, m, v, the logarithm of the posterior distribution of $B = \{\boldsymbol{\beta}_s : s = 1, \ldots, S\}$ is given by $\pi(B|Y, X, Z) = \sum_{s=1}^{S} \pi(\boldsymbol{b}_s|Y_s, X_s, Z_s)$, where, $\pi(\boldsymbol{b}_s|Y_s, X_s, Z_s)$ 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 consider a non-informative prior on β_{0s}, an Inverse-Gamma prior on σ and a half-cauchy prior on τ. Using codes written in `PyMC 4.0` programming language (Salvatier et al., 2016) we sample from the posterior and consider the estimate $\hat{Y}_{st}^{(0)} = \boldsymbol{x}'_{st}\hat{\boldsymbol{\beta}}_s$ where $\hat{\boldsymbol{\beta}}_s$ is the posterior mean. The codes for implementing all the results in this paper is provided in a supplementary zipped folder.

3.1.1 Risk Properties of regularized SC

The LNP induces sparsity in the SC estimates and controls the prediction error of the ATT estimates at future time-points. We rigorously prove it for a factor model structure (Abadie et al., 2010) on the prescription claims. 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: $Y_{st} = \sum_{k=1}^{K} \phi_{sk}\, \mu_{kst} + \varepsilon_{st}$. Consider counterfactual IVM prescription claims from the same model: $Y_{st}^{(0)} = \sum_{k=1}^{K} \phi_{sk}\, \mu_{kst} + \varepsilon_{st}$ for $t = n+1, \ldots, N$. The claims of the control units for $1 \leq s \leq S$ and $1 \leq t \leq N$ obey: $X_{cst} = \sum_{k=1}^{K} \psi_{csk}\, \mu_{kst} + \varepsilon_{cst}$, where, the coefficient $\boldsymbol{\psi}_{cs} = (\psi_{csk} : 1 \leq k \leq K)$ 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 ε_{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 $t_1 \neq t_2$.

On this standard factor model structure used in Abadie et al. (2010), we impose another assumption which states that though we have a very large number (C) of control units, for any state s we can well approximate the coefficients $\boldsymbol{\phi}_s = (\phi_{s1}, \ldots, \phi_{sK})'$ of the treated unit by a linear combination $\boldsymbol{\Psi}_{sc}$ of ℓ control units where $\ell \ll C$. The combination can vary across states and we do not know the effective control units beforehand. We show that with very high probability the LNP based proposed regularized SC method selects the subset of control units accurately.

For any $\boldsymbol{b} \in \mathbb{R}^C$ define, $R_{sk}(\boldsymbol{b}) = \phi_{sk} - \sum_{c=1}^{C} b_c \Psi_{csk}$ and $R_s(\boldsymbol{b}) = \{R_{sk}(\boldsymbol{b}) : k = 1, \ldots, K\}$. Consider the subset $\mathbb{B}(\ell, \alpha)$ of $\mathbb{R}^C$ as $\mathbb{B}(\ell, \alpha) = \{\boldsymbol{x} \in \mathbb{R}^C : ||x||_0 \leq \ell \text{ and } ||x||_\infty \leq \alpha\}$ where the L_0 norm is defined as the number of non-zero coordinates and L_∞ is the supremum norm. Define, $\delta_s(\ell, \alpha) = \min_{\boldsymbol{b} \in \mathbb{B}(\ell, \alpha)} ||R_s(\boldsymbol{b})||_2$.

Assumption 1. Next, we assume that $\boldsymbol{\phi}_s$ can be well-approximated by a sparse combination of control units. Consider an asymptotic regime where $n \to \infty$ and $\liminf_{n\to\infty} C/n > 1$. Assume that there exists $\ell_n \in \mathbb{N}$ and $\alpha_n \geq 0$ satisfying $\lim_{n\to\infty} n^{-1} l_n (\log C)^3 = 0$ and $\lim_{n\to\infty} \alpha_n / n = 0$ such that $\limsup_{n\to\infty} \sup_{1 \leq s \leq S} \delta_s(\ell_n, \alpha_n) \leq \delta$.

Under assumption 1, we show that for all s and $n \gg K$, the parameters $\boldsymbol{\phi}_s = (\phi_{s1}, \ldots, \phi_{sK})'$ and $\{\mu_t : 1 \leq t \leq N\}$ can be well learnt by our proposed LNP based method leading to good SC based estimates of $Y_{st}^{(0)}$. This would subsequently induce uniformly (over time and state) valid controls on the predictive risk of the ATT estimates.

Henceforth, assume $b_{0s} = 0$ as we consider centered covariates. For linear SC estimators of the form $\hat{Y}_{st}^{(0)}(\boldsymbol{b}) = \boldsymbol{b}' \boldsymbol{x}_{st}$, it follows from appendix B of Abadie et al. (2010) that for any

fixed $\boldsymbol{b} \in \mathbb{R}^C$ and for all s and $t > n$, the error of linear SC estimates is upper bounded as:

$$\left|Y_{st}^{(0)} - \hat{Y}_{st}^{(0)}(\boldsymbol{b})\right| \leq ||\mu_{st}||_2 \, \lambda_s \left\{ \mathrm{Im}_s(\boldsymbol{b}) + \sigma\rho(\boldsymbol{b}) \, \frac{||\boldsymbol{Z}_s||_2}{\sqrt{n}} \right\} + \sigma\rho(\boldsymbol{b})|\tilde{Z}_{st}| \tag{2}$$

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

For such linear estimators unless $||\boldsymbol{b}||_2$ is constrained the error in (2) can be unbounded. However, constraining $||\boldsymbol{b}||_2$ might also lead to very large imbalance $\mathrm{Im}_s(\boldsymbol{b})$. Assumption 1 guarantees that there exists at least one unknown $\boldsymbol{b}_s^* \in \mathbb{R}^C$ for which $\mathrm{Im}_s(\boldsymbol{b}_s^*)$ and $||\boldsymbol{b}_s^*||_0$ are controlled as $n \to \infty$. 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 $\hat{Y}_{st}^{(0)}$ have controlled predictive risk. To arrive at this result, based on calculations in Chatterjee (2013), we first show that the SC predictor $\hat{Y}_{st}^{(0)}[\texttt{mo}] = \boldsymbol{x}_{st}'\hat{\boldsymbol{\beta}}_s[\texttt{mo}]$, based on the posterior mode $\hat{\boldsymbol{\beta}}_s[\texttt{mo}]$ of the LNP inherits the slow rate associated with the LASSO predictor.

Theorem 1. *Under Assumption 1 as $n \to \infty$, there exists an LNP prior sequence calibrated only on n and C, such that its posterior mode based SC predictions $\hat{Y}_{st}^{(0)}[\texttt{mo}]$ satisfy:*

$$E\left(Y_{st}^{(0)} - \hat{Y}_{st}^{(0)}[\texttt{mo}]\right)^2 = O\left(\delta^2 m^2 + (\log C)^{3/2} \, n^{-1/2}\right) ,$$

for $1 \leq s \leq S$ and $n+1 \leq t \leq N$ and any $u > 0$.

The proof of Theorem 1 is provided in the supplementary materials. Note that, since we are not assuming any control over the correlations among the control variables (which may be quite high for some controls), the slow rates are justifiable. The above result can

be extended for the LNP posterior mean based SC estimate $\hat{Y}_{st}^{(0)}$. In the supplement, using the results in (Song and Liang, 2023), we show that $\{\hat{Y}_{st}^{(0)} : n+1 \leq t \leq T\}$ concentrates around $\{\boldsymbol{x}'_{st}\boldsymbol{b}_s^* : n+1 \leq t \leq T\}$ 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 2. *Under Assumption 1 as* $n \to \infty$ *and* $\mathcal{D}_n(s) = \{Y_{st}, X_{cst} | 1 \leq t \leq n, 0 \leq c \leq C\}$*, the predictive density* $\pi(\boldsymbol{x}'_{st}\boldsymbol{\beta}_s | \mathcal{D}_n(s))$ *from the LNP prior sequence of Theorem 1 concentrates around* $\boldsymbol{x}'_{st}\boldsymbol{b}_s^*$ *as:*

$$P\left(\log\{\pi\big(\big\|\boldsymbol{x}'_{st}\boldsymbol{\beta}_s - \boldsymbol{x}'_{st}\boldsymbol{b}_s^*\big\| \geq c_1\epsilon_n \big| \mathcal{D}_n(s)\big)\} \geq -c_2\, n\, \epsilon_n^2\right) \leq \exp(-c_3\, n\, \epsilon_n^2)\,,$$

where, $\epsilon_n = c_4 l_n^{1/2} (\log C)^{3/2} n^{-1/2}$ *and* c_1, c_2, c_3, c_4 *are constants. Consequently, the trimmed mean predictor* $\hat{Y}_{st}^{(0)}$ *satisfies* $P\left(\|\hat{Y}_{st}^{(0)} - \boldsymbol{x}'_{st}\boldsymbol{\beta}_s^*\| \leq c_1\epsilon_n\right) \geq 1 - e^{-c_3 n \epsilon_n^2}$ *for all* $t \geq n$.

3.2 Misinformation Persistency Analysis

In addition to addressing the $C \gg n$ problem in (1), 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.

3.3 Misinformation Susceptibility Analysis

We consider the variation in state-level support for the Republican party in the 2016 elections to assess how state-level ATTs vary with the degree of susceptibility. To do so, we use state-level ATT estimates $\widehat{\text{ATT}}_{st}$ for $s = 1, \ldots, S$ and $t = n+1, \ldots, T$. 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 (Esposo 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:

$$\text{ATT Per Capita}_{st} = \mu_0 + \mu_1 \cdot \text{Week}_t + \mu_2 \cdot \text{Consv}_i + \mu_3 \cdot \text{Covid cases per capita}_{st} + \epsilon_{st} \ . \tag{3}$$

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 (3) 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 consumption increase due to misinformation effects

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.

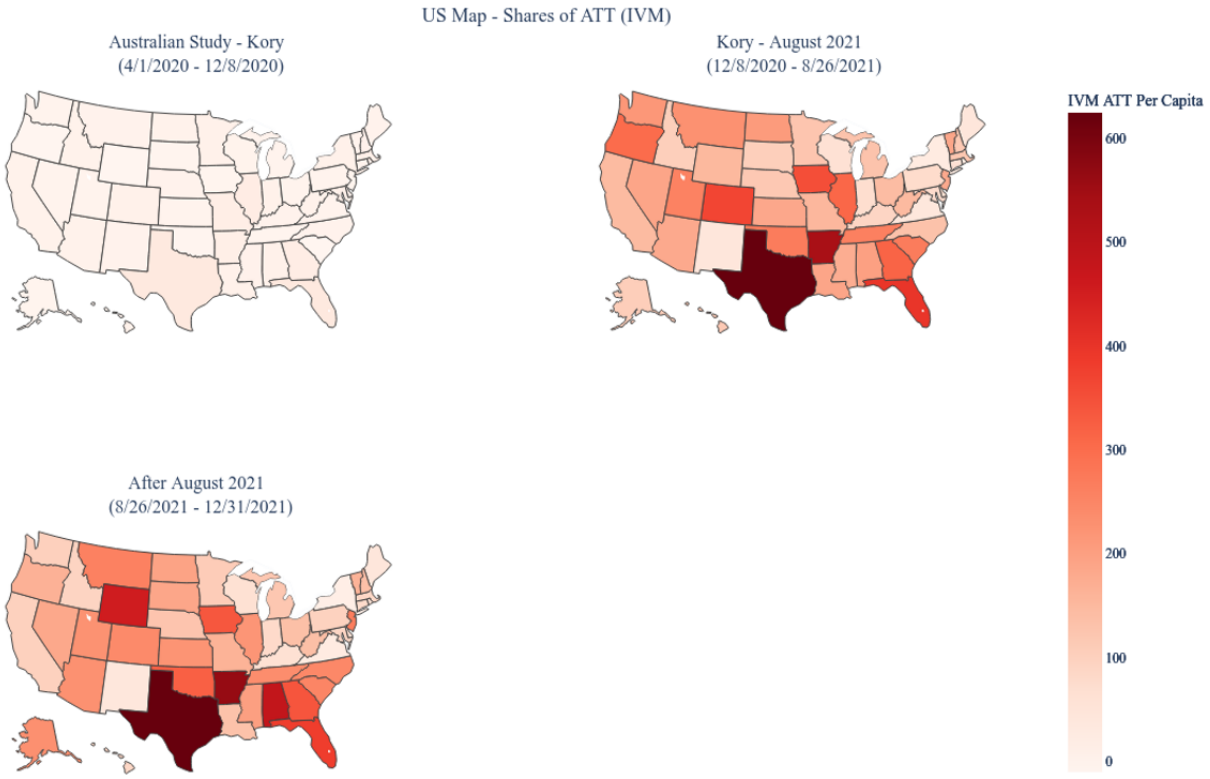

Figure 5: From top-left row-wise, we have maps of ATT estimates for IVM averaged across the time invervals (a) Event 1 to Event 2 (b) Event 2 to Event 8 (c) Event 8 to Event 9. Events listed in Table 1.

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

to misinformation was introduced.

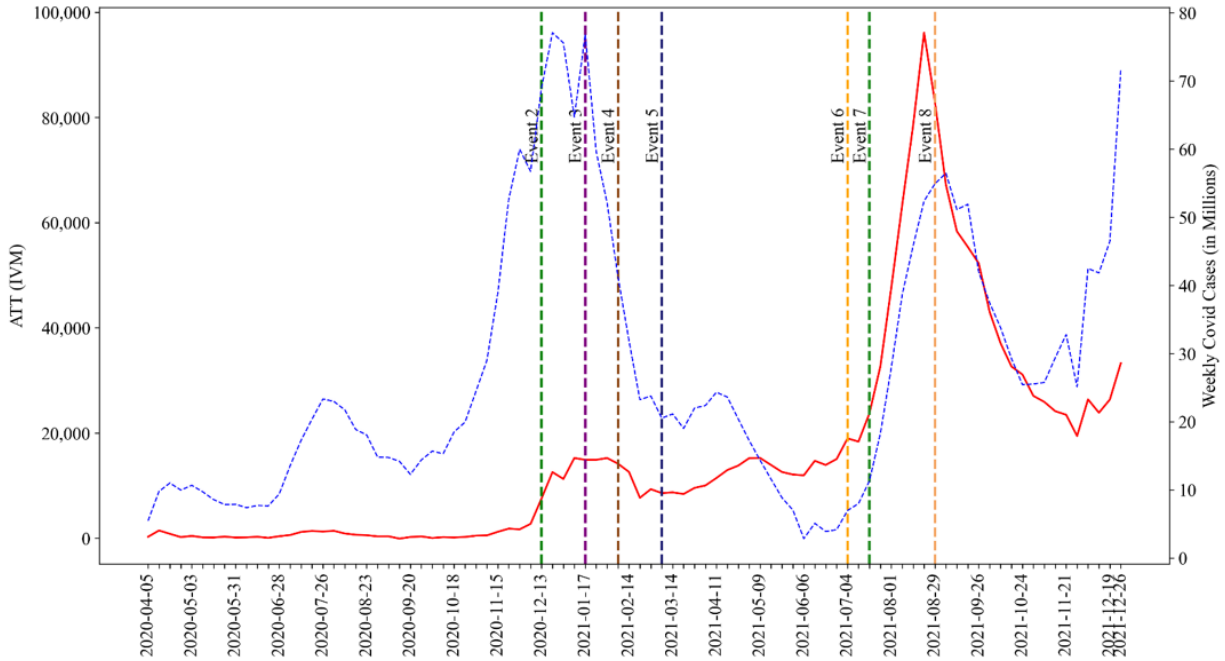

Figure 6: Plot of weekly ATT estimates for IVM prescription claims aggregated across all US states (in red) and of total US weekly Covid case counts (in blue). The time points marked by dotted vertical lines correspond to major related events listed in Table 1.

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 misinformation for IVM, despite federal informative countermeasures. Unlike restrictive countermeasures that involves banning consumption, informative countermeasures to misinformation 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 $\mathcal{T}$, 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 misinformation 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 Susceptibility Analysis

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 weakly 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. The disparity we see between the Read 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).

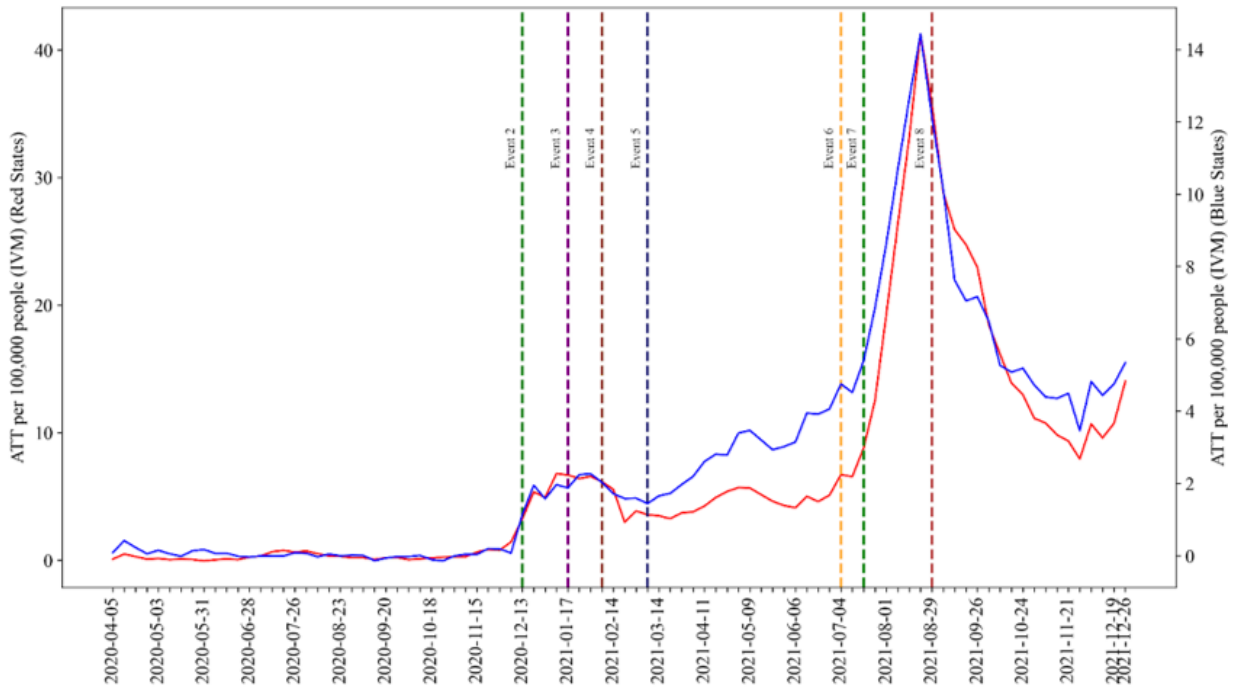

Figure 7: Plot of weekly ATT estimates for IVM prescription claims aggregated across Blue and Red states. The time points marked by dotted vertical lines correspond to major related events listed in Table 1.

To systematically analyze these discrepancies and understand the association between susceptibility to misinformation and political affiliation, we divide the study period into

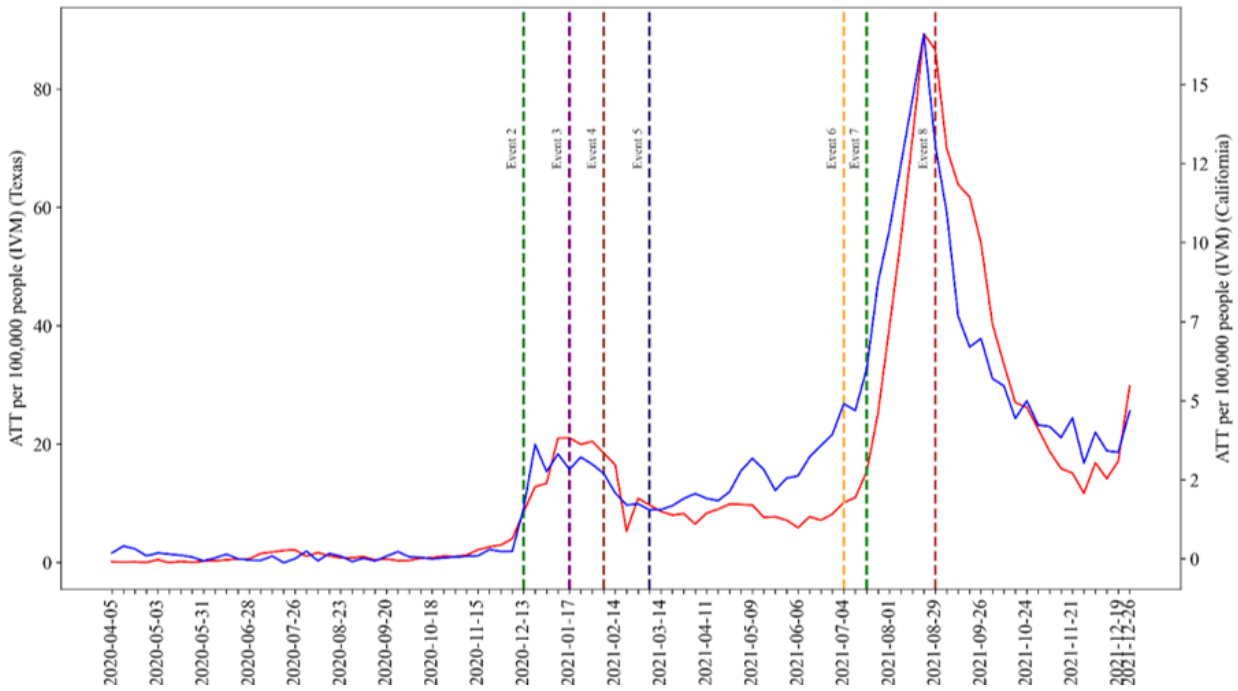

Figure 8: Plot of weekly ATT estimates for IVM prescription claims for California (in blue) and Texas (in red). The time points marked by dotted vertical lines correspond to major related events listed in Table 1.

three distinct intervals. These intervals correspond to those considered in Table 2. We use state-wise weekly indexed incremental IVM claims per capita, adjusted based on controls and derived from ATT estimates, as the response variable. For each of the three intervals, we conduct a regression analysis based on the framework described in (3). 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.

Table 3: Moderating Effect of Susceptibility. Results from regression analysis based on (3) 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	-1.085 (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^2	0.127	0.106	0.318
Adjusted R^2	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 Check

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 supplement.

5 Discussion

This study is the first to examine the medium term effects of misinformation and its countermeasures on physician prescribing behaviors in a causal framework. The findings in this paper indicate that the effects of misinformation persist in the medium term, despite the implementation of informative countermeasures. While countermeasures do provide benefits, the residual impact of misinformation remains substantial. Furthermore, susceptibility to misinformation plays a significant role in its impact. States with higher levels of conservativeness exhibited larger treatment effects per capita. Moreover, countermeasures were found to be less effective in these states. We also find that the emergence of definitive evidence has at best a modest effect in reducing the misinformation effect.

This study underscores the potentially persistent influence of misinformation on the public and experts, and the role of political affiliation in moderating countermeasures. These findings are all the more relevant as a recent survey by the Federation of State Medical Boards showed that 2 out of 3 medical boards reported increased complaints about misinformation spread in 2021 (Chaudhry, 2021). Addressing these challenges will require multidisciplinary efforts from policymakers, healthcare professionals, communication experts, and researchers to develop robust strategies for combating misinformation and promoting evidence-based practices (Cacciatore, 2021).

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature 59*(2), 391–425.

Abadie, A., A. Diamond, and J. Hainmueller (2010). 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.

Abadie, A. and J. Gardeazabal (2003). The economic costs of conflict: A case study of the basque country. *American economic review 93*(1), 113–132.

Abadie, A. and J. L'hour (2021). A penalized synthetic control estimator for disaggregated data. *Journal of the American Statistical Association 116*(536), 1817–1834.

Allcott, H., M. Gentzkow, and C. Yu (2019). Trends in the diffusion of misinformation on social media. *Research & Politics 6*(2), 2053168019848554.

Amjad, M., D. Shah, and D. Shen (2018). Robust synthetic control. *Journal of Machine Learning Research 19*(22), 1–51.

Antonelli, J., G. Parmigiani, and F. Dominici (2019). High-dimensional confounding adjustment using continuous spike and slab priors. *Bayesian analysis 14*(3), 805.

Barnett, M. L., M. Gaye, A. B. Jena, and A. Mehrotra (2022). 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.

Ben-Michael, E., A. Feller, and J. Rothstein (2021). The augmented synthetic control method. *Journal of the American Statistical Association 116*(536), 1789–1803.

Ben-Michael, E., A. Feller, and J. Rothstein (2022). Synthetic controls with staggered adoption. *Journal of the Royal Statistical Society Series B: Statistical Methodology 84*(2), 351–381.

Bhadra, A., J. Datta, Y. Li, N. G. Polson, and B. Willard (2019). Prediction risk for the horseshoe regression. *Journal of Machine Learning Research 20*(46), 1–43.

Bhadra, A., J. Datta, N. G. Polson, and B. Willard (2016). Default bayesian analysis with global-local shrinkage priors. *Biometrika 103*(4), 955–969.

Bhattacharya, A., D. Pati, N. S. Pillai, and D. B. Dunson (2015). Dirichlet-laplace priors for optimal shrinkage. *Journal of the American Statistical Association 110*(512), 1479–1490.

Bonis, L. and M. Curtis (2020). 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`. Accessed: 2024-05-11.

Borges Nascimento, I. J., A. B. Pizarro, J. Almeida, N. Azzopardi-Muscat, M. A. Gonçalves, M. Björklund, and D. Novillo-Ortiz (2022). Infodemics and health misinformation: a systematic review of reviews. *Bulletin of the World Health Organization 100*(9), 544–561.

Boukouvalas, Z. and A. Shafer (2024). Role of statistics in detecting misinformation: A review of the state of the art, open issues, and future research directions. *Annual Review of Statistics and Its Application 11*.

Cacciatore, M. A. (2021). Misinformation and public opinion of science and health: Approaches, findings, and future directions. *Proceedings of the National Academy of Sciences 118*(15), e1912437117.

Carvalho, C., R. Masini, and M. C. Medeiros (2018). Arco: An artificial counterfactual approach for high-dimensional panel time-series data. *Journal of econometrics 207*(2), 352–380.

Castillo, I. and R. Mismer (2018). Empirical bayes analysis of spike and slab posterior distributions. *Electronic Journal of Statistics*.

Chatterjee, S. (2013). Assumptionless consistency of the lasso. *arXiv preprint arXiv:1303.5817*.

Chaudhry, H. (2021). Fsmb — two-thirds of state medical boards see increase in covid-19 disinformation complaints. accessed August 19, 2023.

Chou, W.-Y. S., A. Oh, and W. M. P. Klein (2018). Addressing health-related misinformation on social media. *JAMA 320*(23), 2417.

Christensen, J. (2021). 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`. Accessed: 2024-05-11.

Doudchenko, N. and G. W. Imbens (2016). Balancing, regression, difference-in-differences and synthetic control methods: A synthesis. Technical report, National Bureau of Economic Research.

Ecker, U. K. H., S. Lewandowsky, J. Cook, P. Schmid, L. K. Fazio, N. Brashier, P. Kendeou, E. K. Vraga, and M. A. Amazeen (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology 1*(1), 13–29.

Ecker, U. K. H., S. Lewandowsky, and D. T. W. Tang (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition 38*(8), 1087–1100.

Esposo, S. R., M. J. Hornsey, and J. R. Spoor (2013). Shooting the messenger: Outsiders critical of your group are rejected regardless of argument quality. *British Journal of Social Psychology 52*(2), 386–395.

Fetzer, J. H. (2004). Information: Does it have to be true? *Minds and Machines 14*, 223–229.

Fong, J., T. Guo, and A. Rao (2022). Debunking misinformation about consumer products: Effects on beliefs and purchase behavior. *Journal of Marketing Research*. 002224372211470.

Fuhrer, J. and F. Cova (2020). "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.

Guess, A. M. and B. A. Lyons (2020). Misinformation, disinformation, and online propaganda. *Social media and democracy: The state of the field, prospects for reform 10*.

Hill, A., M. Mirchandani, L. Ellis, and V. Pilkington (2022). Ivermectin for the prevention of covid-19: addressing potential bias and medical fraud. *Journal of Antimicrobial Chemotherapy 77*(5), 1413–1416.

Imbens, G. W. and D. B. Rubin (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge university press.

Ishwaran, H. and J. S. Rao (2005). Spike and slab variable selection: frequentist and bayesian strategies. *Annals of Statistics*.

Karmakar, B., G. Mukherjee, and W. Kar (2024). Using penalized synthetic controls on truncated data: A case study on effect of marijuana legalization on direct payments to physicians by opioid manufacturers.

Kim, S., C. Lee, and S. Gupta (2020). Bayesian synthetic control methods. *Journal of Marketing Research 57*(5), 831–852.

Kouzy, R., J. Abi Jaoude, A. Kraitem, M. B. El Alam, B. Karam, E. Adib, J. Zarka, C. Traboulsi, E. W. Akl, and K. Baddour (2020). Coronavirus goes viral: quantifying the covid-19 misinformation epidemic on twitter. *Cureus 12*(3).

Lewandowsky, S., U. K. H. Ecker, C. M. Seifert, N. Schwarz, and J. Cook (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest 13*(3), 106–131.

Mackie, D. M., L. T. Worth, and A. G. Asuncion (1990). Processing of persuasive in-group messages. *Journal of Personality and Social Psychology 58*(5), 812–822.

Malsiner-Walli, G. and H. Wagner (2018). Comparing spike and slab priors for bayesian variable selection. *arXiv preprint arXiv:1812.07259*.

Mendez, R. (2021). Who says covid misinformation is a major factor driving pandemic around the world. https://www.cnbc.com/2021/08/24/who-says-covid-misinformation-is-a-major-factor-driving-pandemic-around-the-world.html.

Miller, J. W. and M. T. Harrison (2018). Shrink globally, act locally: Sparse bayesian regularization and prediction. *Journal of the American Statistical Association 113*(524), 1691–1703.

Muhammed, S. and S. Mathew (2022). The disaster of misinformation: a review of research in social media. *International journal of data science and analytics 13*(4), 271–285.

Naeem, S. and M. Boulos (2021). Covid-19 misinformation online and health literacy: a brief overview. *International journal of environmental research and public health 18*(15), 8091.

Nyhan, B. (2021). Why the backfire effect does not explain the durability of political misperceptions. *Proceedings of the National Academy of Sciences 118*(15), e1912440117.

Nyhan, B. and J. Reifler (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior 32*(2), 303–330.

Pennycook, G. and D. G. Rand (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences 116*(7), 2521–2526.

Perlis, R. H., K. L. Trujillo, J. Green, A. Safarpour, J. N. Druckman, M. Santillana, K. Ognyanova, and D. Lazer (2023). Misinformation, trust, and use of ivermectin and hydroxychloroquine for covid-19. In *JAMA Health Forum*, Volume 4, pp. e233257–e233257. American Medical Association.

Rao, A. and E. Wang (2017). Demand for 'healthy' products: False claims and ftc regulation. *Journal of Marketing Research 54*(6), 968–989.

Ročková, V. and E. I. George (2018). The spike-and-slab lasso. *Journal of the American Statistical Association 113*(521), 431–444.

Roozenbeek, J., C. R. Schneider, S. Dryhurst, J. Kerr, A. L. J. Freeman, G. Recchia, A. M. van der Bles, and S. van der Linden (2020). Susceptibility to misinformation about covid-19 around the world. *Royal Society Open Science 7*(10), 201199.

Ročková, V. and E. I. George (2018). Spike-and-slab meets lasso: A review of the spike-and-slab lasso. *Statistical Science 33*(4), 485–514.

Rubin, R. (2022). When physicians spread unscientific information about covid-19. *JAMA 327*(10), 904.

Salvatier, J., T. V. Wiecki, and C. Fonnesbeck (2016). Probabilistic programming in python using pymc3. *PeerJ Computer Science 2*, e55.

Scherer, L. D., J. McPhetres, G. Pennycook, A. Kempe, L. A. Allen, C. E. Knoepke, C. E. Tate, and D. D. Matlock (2021). Who is susceptible to online health misinformation? a test of four psychosocial hypotheses. *Health Psychology*.

Song, Q. and F. Liang (2023). Nearly optimal bayesian shrinkage for high-dimensional regression. *Science China Mathematics 66*(2), 409–442.

Steen, R. G. (2011). Misinformation in the medical literature: what role do error and fraud play? *Journal of medical ethics 37*(8), 498–503.

National Institute of Health (2021). 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`. Accessed: 2024-05-11.

New York Times (2016). 2016 presidential election results. `https://www.nytimes.com/elections/2016/results/president`. Accessed: 2024-05-11.

New York Times (2023). Coronavirus (covid-19) data in the united states. `https://www.nytimes.com/interactive/2021/us/covid-cases.html`. Accessed: 2024-05-11.

Pew Research Center (2024). The partisanship and ideology of american voters. Accessed: 2024-08-12.

Thomson, W., S. Jabbari, A. E. Taylor, W. Arlt, and D. J. Smith (2019). 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.

Vaduganathan, M., J. van Meijgaard, M. R. Mehra, J. Joseph, C. J. O'Donnell, and H. J. Warraich (2020). Prescription fill patterns for commonly used drugs during the covid-19 pandemic in the united states. *JAMA 323*(24), 2524.

Voss, A. (2020). Official statement from international society of antimicrobial chemotherapy (isac). International Society of Antimicrobial Chemotherapy. Accessed: 2024-05-11.

West, J. D. and C. T. Bergstrom (2021). Misinformation in and about science. *Proceedings of the National Academy of Sciences 118*(15), e1912444117.

Supplementary Materials

A Robustness Checks

Our robustness checks broadly fall into two types : data/model related and alternative explanations for our observed effects. We provided these results in the supplements.

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.

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 S1. 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 S1. Comparing with table 3 we see that these results are also substantively the same.

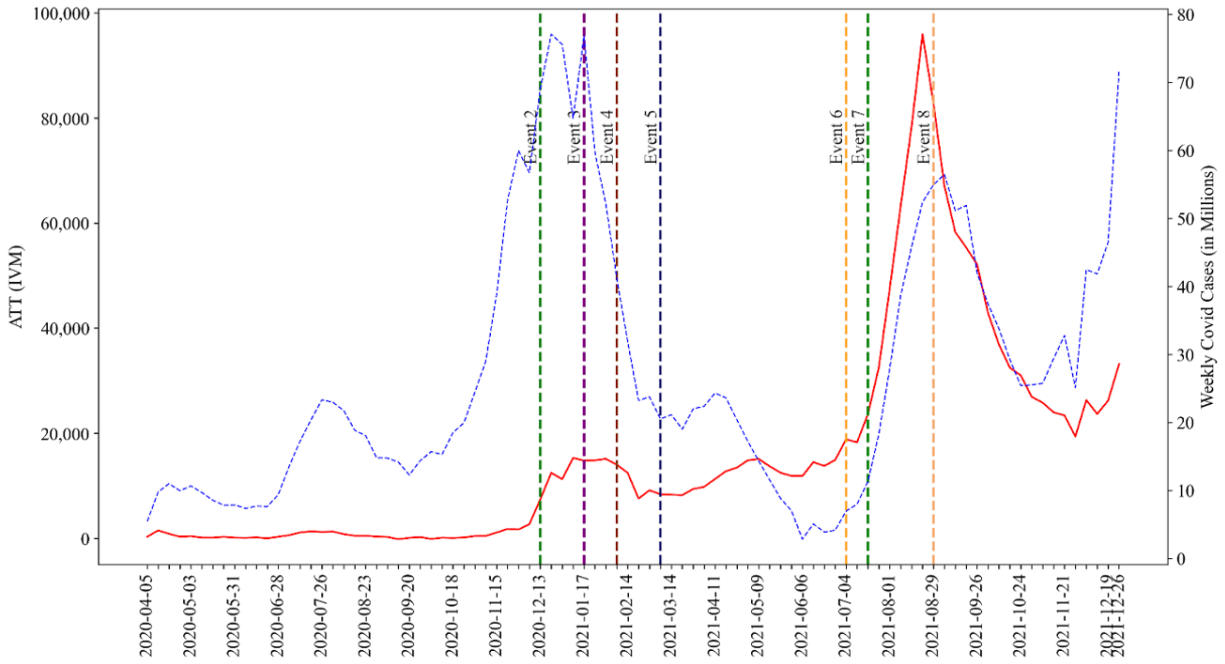

Figure S1: Based on an alternative control set we plot the ATT estimates for IVM prescription claims aggregated across all US states (in red) and of total US weekly Covid case counts (in blue).

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 S2. 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.

Table S1: Moderating Effect of Susceptibility based on an alternative control set. Results from regression analysis based on (3) 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^2	0.126	0.110	0.289
Adjusted R^2	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)

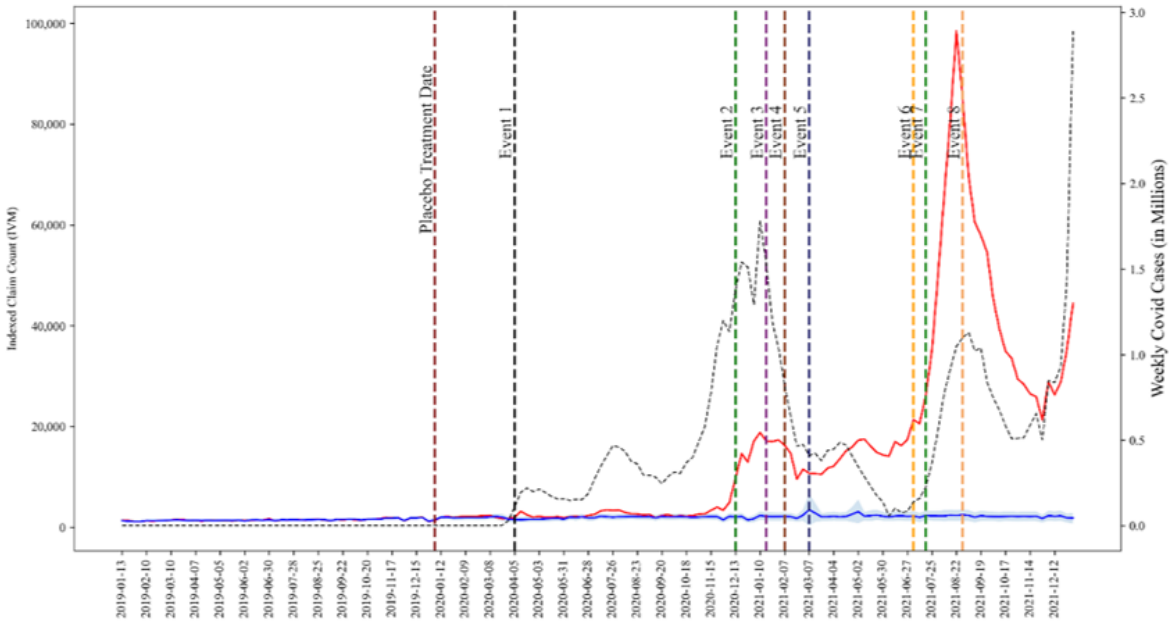

Figure S2: In red and blue we have the actual and predicted IVM claims when the treatment date was preponed by 12 weeks for a placebo test. In dotted black line we have Covid case counts.

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 Theorem 1

Consider the setup for a particular state s. We remove subscripts pertaining to state s for ease of notation. In this case for the pre-treatment IVM claims $\{Y_t : t = 1, \cdots, n\}$, assume the model for a single state in accordance with (1)

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

For fixed σ, τ and a fixed sequence of λ_c, the negative log likelihood of posterior distribution of β_c without the prior structure on λ_c is

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

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

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

Let σ^* be the true standard deviation of ϵ_t in (1), then by redefining $\tilde{Y}_t = \sqrt{\tau}\sigma^{-1}Y_t$ and $\tilde{X}_{ct} = \sqrt{\tau}\sigma^{-1}X_{ct}$, our optimization boils down to

$$\min_{\beta}\sum_{t=1}^{n}\left(\tilde{Y}_t - \sum_{c=1}^{C}\beta_c\tilde{X}_{ct}\right)^2 \qquad \text{s.t.} \quad \sum_{c=1}^{C}\frac{\beta_c^2}{\lambda_c^2} \leq K\,. \tag{4}$$

Note that while the variance of Y_t was σ^{*2}, by using the transformation, the variance of $\tilde{Y}_t$ is $\tau\sigma^{*2}/\sigma^2$. We next provide the construction to choose prior hyperparamaters for λ_c.

Lemma 1. *Let $X \sim$ `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 $t \to \infty$, for some $u > 0$, there exists constants k_1, k_2 such that

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

Using the above lemma, we set the prior mean $-Cf(C)$ and variance C^2 for λ_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 λ_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 λ_c's as C grows. Without loss of generality assume that $\lambda_1, \lambda_2, \cdots, \lambda_s$ are greater than $1/\sqrt{C}$ and $\lambda_{s+1}, \lambda_{s+2}, \cdots, \lambda_C \leq 1/\sqrt{C}$. Consider $2s\sqrt{\sum_{c=1}^{C}\beta_c^2/\lambda_c^2}$,

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

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

$$\min_{\beta} \sum_{t=1}^{n} \left(\tilde{Y}_t - \sum_{c=1}^{C} \beta_c \tilde{X}_{ct}\right)^2 \qquad \text{s.t.} \quad \|\boldsymbol{\beta}\|_1 \leq 2s\sqrt{K}\,. \tag{5}$$

While the objective function of the optimization problems in both (4) and (5) are the same, since $\sum_{c=1}^{C} \beta_c^2/\lambda_c^2 \leq K$ implies that $\|\boldsymbol{\beta}\|_1 \leq 2s\sqrt{K}$, thus, the constraint in (5) is a smaller set compared to the constraint set in (4). This means that the minimum value of the objective function achieved in (5) will be larger than the one achieved in the other optimization problem in (4). We thus shift to solving for the stricter minimization problem in instead. The optimization problem in (5) 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 the errors have variance* σ^2*. Let MSPE stand for the mean squared prediction error, defined as* $E\left(\tilde{Y} - \hat{Y}\right)^2$ *where* $\hat{Y}$ *is the prediction based on* $\hat{\beta}$ *and* $\tilde{Y}$ *is the best possible prediction based on the true parameter* $\boldsymbol{\beta}^*$*, then*

$$E\left(\tilde{Y} - \hat{Y}\right)^2 = c_1 M\sigma\sqrt{\frac{\log(p)}{n}} + c_2 M^2\sqrt{\frac{\log(p)}{n}}\,,$$

for some constants c_1 *and* c_2*, where* $\max|X_{ct}| \leq M$*.*

In our case, the variance is $\tau\sigma^{*2}/\sigma^2$. Also M is the maximum of n of X_{ct}'s which is of the order $\sqrt{\log n}$. With this we get for out case, given fixed τ and σ,

$$E\left[\left(\tilde{Y}_t^{(0)} - \hat{Y}_t^{(0)}\right)^2 |\tau, \sigma\right] = c_1\frac{\sigma^*}{\sigma}\sqrt{\frac{\tau \log n \log C}{n}} + c_2 \log n\sqrt{\frac{\log C}{n}}$$

If $\sigma \sim$ `Inverse-Gamma`(a_0, b_0) and $\tau \sim$ `Half-Cauchy`(c_0, γ), then $E[1/\sigma] = a_0/b_0$ and

$E[\sqrt{\tau}] = c_0 + \sqrt{2/\gamma}$. Since σ and τ are independent of each other, we have

$$\begin{aligned} E\left(\tilde{Y}_t^{(0)} - \hat{Y}_t^{(0)}\right)^2 &= E\left[E\left[\left(\tilde{Y}_t^{(0)} - \hat{Y}_t^{(0)}\right)^2 |\tau, \sigma\right]\right] \\ &= c_1 \frac{a_0}{b_0}(c_0 + \sqrt{2/\gamma})\sigma^* \sqrt{\frac{\log n \log C}{n}} + c_2 \log n \sqrt{\frac{\log C}{n}} \\ &= O\left(\sigma^* \sqrt{\frac{\log n \log C}{n}} + \log n \sqrt{\frac{\log C}{n}}\right) \\ &= O\left(\log n \sqrt{\frac{\log C}{n}}\right) = O\left(\sqrt{\frac{(\log C)^3}{n}}\right), \end{aligned} \tag{6}$$

where the last inequalities follows since $C \geq n$ and $\log n$ dominates $\sqrt{\log n}$. Next, we consider $E(Y_t^{(0)} - \tilde{Y}_t^{(0)})^2$ which constitutes the random error as well as the training error based on the approximation of the factor model based on the controls.

Note that based on our factor model, $Y_t^{(0)} = \sum_{k=1}^K \phi_k \mu_{kt} + \epsilon_t$ and similarly $E[X_{ct}] = \sum_{k=1}^K \psi_{ck}\mu_{kt}$. Based on these definitions, $Y_t^{(0)} - \tilde{Y}_t^{(0)} = \sum_{k=1}^K \mu_k R_k(\boldsymbol{\beta}^*) + \epsilon_t$. Since $\|\mu_k\| \leq m$, thus, $E(Y_t^{(0)} - \tilde{Y}_t^{(0)})^2 \leq m^2 \|R(\beta^*)\|_2^2 + \sigma^{*2} \leq m^2\delta^2 + \sigma^{*2}$. Combined with (6), this proves our final bound since

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

B.2 Proof of Theorem 2

We first describe some properties pertaining to our choice of prior in our model coefficients β_c. The coefficients are mean zero Gaussian distributed with the variances being controlled by τ and λ_c^2. While τ is distributed with a Half Cauchy prior, λ_c's are Logit Normals with hyperpaparmeter as defined Lemma 1.

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

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

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

1. The scaled covariates $X_{ct}/\sqrt{\log nC}$ 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 $\sqrt{\log nC}$ and the other two statements follow from Assumption 1. Lemma 2 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 $A_n = \{$ At least $\tilde{C}$ entries of $|\boldsymbol{\beta}/\sigma|$ is larger than $a_n\} \cup \{\big(\big\|\boldsymbol{x}'_{st}\boldsymbol{\beta}_s - \boldsymbol{x}'_{st}\boldsymbol{b}_s^*\big\| \geq c_1\sigma^*\epsilon_n\} \cup \{\sigma^2/\sigma^{*2} > (1+\epsilon_n)/(1-\epsilon_n)$ or $\sigma^2/\sigma^{*2} < (1-\epsilon_n)/(1+\epsilon_n)\}$, $B_n = \{$ At least $\tilde{C}$ entries of $|\boldsymbol{\beta}/\sigma|$ is larger than $a_n\}$ and $C_n = A_n \backslash B_n$, 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,

$$\phi'_n = \max_{\{\xi \supseteq \xi^*, |\xi| \leq \tilde{C}+l_n\}} 1\{\big|^T(I - H_\xi)/(n-|\xi|)\sigma^{*2} - 1\big| \geq \epsilon_n\}, \text{ and}$$

$$\tilde{\phi}_n = \max_{\{\xi \supseteq \xi^*, |\xi| \leq \tilde{C}+l_n\}} 1\{\|X_\xi(X_\xi^T X_\xi)^{-1}X_\xi^T y - X_\xi\boldsymbol{\beta}_\xi^*)\| \geq c_1\sigma^*\epsilon_n\}\,,$$

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

With these testing functions, we finally get

$$P\left(\{\pi\big(\big\|\boldsymbol{x}'_{st}\boldsymbol{\beta}_s - \boldsymbol{x}'_{st}\boldsymbol{b}_s^*\big\| \geq c_1\epsilon_n\big|\mathcal{D}_n(s)\big)\} \geq e^{-c_2\, n\,\epsilon_n^2}\right) \leq \exp(-c_3\, n\,\epsilon_n^2)\,. \tag{7}$$

Next, let $\tilde{\mathcal{D}}_n(s) \subseteq \mathcal{D}_n(s)$, such that $\pi\big(\big\|\boldsymbol{x}'_{st}\boldsymbol{\beta}_s - \boldsymbol{x}'_{st}\boldsymbol{b}^*_s\big\| \geq c_1\epsilon_n\big)\} \leq e^{-c_2\, n\, \epsilon_n^2}$. This implies that $\pi\big(\boldsymbol{x}'_{st}\boldsymbol{\beta}_s \geq c_1\epsilon_n + \boldsymbol{x}'_{st}\boldsymbol{b}^*_s\big)\} \leq e^{-c_2\, n\, \epsilon_n^2}$ and $\pi\big(\boldsymbol{x}'_{st}\boldsymbol{\beta}_s \leq -c_1\epsilon_n + \boldsymbol{x}'_{st}\boldsymbol{b}^*_s\big)\} \leq e^{-c_2\, n\, \epsilon_n^2}$. Thus, if we consider the trimmed mean of the posterior predictions converges.

B.3 Proof of Lemma 1

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

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

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 θ 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, $1 - \Phi\left(f(t)\right) = O(\exp(-f(t)^2/2)/f(t)) = O(t^{-(1+u)})$ or $1 - \Phi\left(f(t)\right) = k_1 t^{-(1+u)}$ for some constant k_1 which proves the required.

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

B.4 Proof of Lemma 2

For any set of hyperparameters τ, λ_c, the β_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 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 τ and λ_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

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

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, $P(\mathcal{B}) = P(\tau \leq c_0 + C^{1+u})P(\lambda_c \leq C^{-(4+u)/2})$. Consider $P(\lambda_c \leq C^{-(4+u)/2})$,

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

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, $P(\tau \leq c_0 + C^{1+u}) \approx 1 - 2C^{-(1+u)}/\pi$. Thus,

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

Also, in this set $\mathcal{B}$, the marginal variance of β_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.