

Johnsson, Ida; Pesaran, M. Hashem; Yang, Cynthia Fan

**Working Paper**

## Structural Econometric Estimation of the Basic Reproduction Number for Covid-19 across U.S. States and Selected Countries

CESifo Working Paper, No. 10659

**Provided in Cooperation with:**

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Johnsson, Ida; Pesaran, M. Hashem; Yang, Cynthia Fan (2023) : Structural Econometric Estimation of the Basic Reproduction Number for Covid-19 across U.S. States and Selected Countries, CESifo Working Paper, No. 10659, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/282347>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

**Structural Econometric  
Estimation of the Basic  
Reproduction Number for  
Covid-19 across U.S. States and  
Selected Countries**

*Ida Johnsson, M. Hashem Pesaran, Cynthia Fan Yang*

## **Impressum:**

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email [office@cesifo.de](mailto:office@cesifo.de)

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: [www.SSRN.com](http://www.SSRN.com)
- from the RePEc website: [www.RePEc.org](http://www.RePEc.org)
- from the CESifo website: <https://www.cesifo.org/en/wp>

# Structural Econometric Estimation of the Basic Reproduction Number for Covid-19 across U.S. States and Selected Countries

## Abstract

This paper proposes a structural econometric approach to estimating the basic reproduction number ( $R_0$ ) of Covid-19. This approach identifies  $R_0$  in a panel regression model by filtering out the effects of mitigating factors on disease diffusion and is easy to implement. We apply the method to data from 48 contiguous U.S. states and a diverse set of countries. Our results reveal a notable concentration of  $R_0$  estimates with an average value of 4.5. Through a counterfactual analysis, we highlight a significant underestimation of the  $R_0$  when mitigating factors are not appropriately accounted for.

JEL-Codes: C130, C330, I120, I180, J180.

Keywords: basic reproduction number, Covid-19, panel threshold regression model.

*Ida Johnsson*  
*Chewy, Inc. & University of Southern*  
*California, Los Angeles / CA / USA*  
*ida.b.johnsson@gmail.com*

*M. Hashem Pesaran*  
*University of Southern California*  
*Los Angeles / CA / USA*  
*pesaran@usc.edu*

*Cynthia Fan Yang\**  
*Department of Economics*  
*Florida State University*  
*Tallahassee / FL / USA*  
*cynthia.yang@fsu.edu*

\*corresponding author

August 30, 2023

We wish to thank Alessandro Rebucci, Alexander Chudik, and Ron Smith for their valuable comments. We also gratefully acknowledge Alexander Chudik for helpful discussions regarding his codes.

# 1 Introduction

Since the Covid-19 outbreak, a rapidly growing body of literature has been devoted to estimating its reproduction numbers, which are standard epidemiological metrics used to quantify the rate at which the disease spreads at the initial and later stages of the epidemic. Obtaining accurate estimates of the reproduction numbers is critical in understanding the transmissibility of infectious agents and formulating public health responses. Different formal definitions of reproduction numbers have been proposed in the literature. One of the most widely adopted metrics is the basic reproduction number, denoted by  $\mathcal{R}_0$ , which is defined as "the average number of secondary cases produced by one infected individual during the infected individual's entire infectious period assuming a fully susceptible population" (Del Valle et al., 2013). An essential assumption of this definition is that the population must be entirely susceptible without any immunity or interventions. In other words, the value of  $\mathcal{R}_0$  represents the maximum rate at which an epidemic can spread *in the absence of any mitigating actions*, whether voluntary or mandatory, such as improved personal hygiene, mask-wearing, social distancing, isolation, or vaccination. In practice, the day-to-day evolution of the epidemic is governed by public health directives and individual behavior, as well as the internal dynamics of the epidemic, which on its own eventually slows down its rate of spread as the number of susceptible individuals declines (due to immunity, death, or mitigation practices). A measure that quantifies the time-varying transmissibility during the life of the epidemic is the effective reproduction number,  $\mathcal{R}_{e,t}$ .<sup>1</sup> The focus of the present study is on the structural estimation of  $\mathcal{R}_0$  for Covid-19 across U.S. states and countries.

Obtaining a reliable estimate of  $\mathcal{R}_0$ , although seemingly straightforward, can be quite challenging in practice. This is largely due to the difficulty of accurately counting early infections during an epidemic, especially for newly emerging pathogens like SARS-CoV-2. Even if public health officials may strive to establish surveillance systems promptly, the data quality during the initial stages of the outbreak is inevitably poor. Additionally, many existing time-series estimation methods encounter difficulties in identifying an appropriate sample period to estimate  $\mathcal{R}_0$ . Unlike  $\mathcal{R}_{e,t}$ , which varies over time and is affected by mitigation policies, to estimate  $\mathcal{R}_0$  it is important to focus on the initial stages of the epidemic when the infections are spreading exponentially, without impediments. If the selected sample is too long, it will likely include periods of reduced social interactions due to mandatory and/or voluntary behavioral changes, which slow down the spread of the infection and result in a downward bias in the  $\mathcal{R}_0$  estimates. On the other hand, if the sample selected at the initial stages of the epidemic is too short, the under-recording of infected

---

<sup>1</sup>The effective reproduction number is the expected number of secondary cases produced by one infected individual in a population that includes both susceptible and non-susceptible individuals at time  $t$  (assuming that the conditions remained the same after time  $t$ ). The effective reproduction number differs from another time-varying metric, the case reproduction number (Fraser, 2007). The latter is defined as the average number of secondary cases that a primary case actually infects at time  $t$ , and thus its estimation can only be undertaken retrospectively.

cases can be particularly serious, resulting in an underestimation of the disease’s transmission rate and again leading to the underestimation of  $\mathcal{R}_0$ .

The availability of Covid-19 data from diverse populations and locations presents an unprecedented opportunity for researchers to compare the  $\mathcal{R}_0$  estimates using various models and estimation methods. Researchers around the world have reported a wide range of  $\mathcal{R}_0$  values of Covid-19, as summarized in Table 1. Some studies have highlighted the heterogeneity of  $\mathcal{R}_0$  across different countries and regions (Korolev, 2021; Fernández-Villaverde and Jones, 2022), while others find evidence of similarity (Katul et al., 2020; Chudik et al., 2022). Many of the existing estimates were obtained based on complex models that require large amounts of data and involve numerous assumptions made by the modeler. As it is rare to have high-quality data for all components of the model, researchers often have to calibrate multiple model parameters (Delamater et al., 2019). As our knowledge of the disease and the quality of data have improved over time, an increasing number of researchers have argued that earlier estimates of  $\mathcal{R}_0$  were too conservative (Katul et al., 2020; Ke et al., 2021).

In the current paper, we estimate  $\mathcal{R}_0$  of Covid-19 using a structural econometric approach, which allows us to identify  $\mathcal{R}_0$  as the intercept in panel regressions of outcomes on a number of key factors across the contiguous states in the U.S. and a diverse range of countries. This structural method is grounded in the individual-based stochastic network model for epidemic diffusion recently developed by Pesaran and Yang (2022), who established a moment condition for the evolution of the infections. Chudik et al. (2022) use the same moment condition but allow for time variations in the transmission rate, modeled as a function of various factors, including government responses (mandatory mitigation policies and economic support), precautionary behavioral changes such as voluntary social distancing, vaccinations, and virus mutations. These authors focus on nine European countries, which had similar starts at the outset of the epidemic in March 2020 but ended up with differing outcomes. They find robust statistically significant effects of reduced mobility, increased government economic support to comply with containment policies, vaccination, and virus mutations on the transmission rate of Covid-19.

We employ the structural approach proposed by Chudik et al. (2022), CPR hereafter, and estimate  $\mathcal{R}_0$  as the intercept in panel data regressions where suitably transformed case numbers are explained in terms of the mitigating factors identified in the CPR study. Such an estimate can be viewed as a counterfactual outcome that could have resulted if none of the public health policies had been enacted. This method has several advantages. First, by controlling for a number of mitigation measures, we are able to estimate  $\mathcal{R}_0$  using relatively long samples, thus avoiding the difficulties of selecting an early sample. Second, the model easily accommodates the differences in the initial start dates of the Covid-19 outbreaks across countries (or regions) by using an unbalanced panel data model. Third, country-specific  $\mathcal{R}_0$  can be estimated by least squares under very weak assumptions and only requires the regressors in the panel data model

to be weakly exogenous. We also use lagged values of the regressors to avoid simultaneity bias, and we take account of error serial correlation when computing the standard errors to deal with the implications of using overlapping seven-day moving averages, a practice routinely employed in the literature to deal with uneven recording of infected cases over the different days of the week. Fourth, our estimation of  $\mathcal{R}_0$  requires only Covid-19 case data and an assumed value for the recovery rate,  $\gamma$ . This approach serves as a useful complement to other methods that rely on mortality data. Instead of estimating the recovery rate, we set it *a priori* using clinical evidence with a high degree of reliability. Fifth, we allow for the under-reporting of infected cases, which is known to vary across different phases of the epidemic.

We find remarkably similar estimates of  $\mathcal{R}_0$  across 48 contiguous U.S. states, averaging at 4.7 with a range of 4.2 to 4.9 (4.5 to 5.3) under the assumption of a lower (higher) degree of under-reporting. In addition, for a selection of diverse nations,  $\mathcal{R}_0$  remains within a narrow range of 3.4 to 5 across all scenarios considered, with an average value of 4.3. Overall, these estimates align well with the recent findings in the literature. Furthermore, our counterfactual analysis underscores a substantial underestimation of  $\mathcal{R}_0$  that results from neglecting the mitigating factors governing the time profile of the Covid-19 transmission rate.

The rest of the paper is organized as follows. Section 2 reviews the related literature. Section 3 outlines the empirical methodology. Section 4 presents the main findings, and Section 5 concludes. Additional tables and figures are provided in an online supplement.

## 2 Related Literature

The estimation of the reproduction number for infectious diseases, including Covid-19, has been a topic of extensive research in the literature. In the interest of space, we will highlight a selection of recent contributions that are closely related to our study.

There are many different approaches to the estimation of  $\mathcal{R}_0$  that can be broadly categorized into mathematical and statistical approaches (White et al., 2020). The mathematical approaches rely on constructing a theoretical model, which makes explicit hypotheses about the biological mechanisms driving the transmission rate and its dynamics. Such hypotheses range from simple representations of the time needed to complete some part of the disease process (such as the incubation period) to complex agent-based models with explicit modeling of social interactions (Lessler and Cummings, 2016). In contrast, statistical methods derive estimators of  $\mathcal{R}_0$  directly from data using probabilistic models. They do not require specifying a disease-approximating mathematical model and can be implemented using standard statistical software.

The most widely used mathematical models for disease transmission are the compartmental models, most notably the classical SIR model that divides a population into three compartments (susceptible, infected, and recovered) and its variants, such as the SEIR and SIRD models (with

an additional exposed or deceased compartment, respectively). [Tang et al. \(2020\)](#) provide a recent review of the compartmental infectious disease models. Many researchers have made efforts to infer  $\mathcal{R}_0$  of Covid-19 from different mathematical models and adopted various frequentist methods to estimate or calibrate the model parameters (e.g., [Buckman et al., 2020](#); [Katul et al., 2020](#); [Ke et al., 2021](#); [Korolev, 2021](#); [Fernández-Villaverde and Jones, 2022](#)). Another strand of literature estimates similar compartmental models using Bayesian techniques (e.g., [Wu et al., 2020](#); [Arias et al., 2023](#); [Atkeson et al., 2023](#)).<sup>2</sup>

The statistical methods typically require an estimate of the generation time, which is the time lag between infections in primary and secondary cases ([Obadia et al., 2012](#)). A frequently used estimator of  $\mathcal{R}_0$  was proposed by [Wallinga and Lipsitch \(2007\)](#), who show how to estimate  $\mathcal{R}_0$  from the estimated exponential growth rate using data on infected cases during the *early phase* of an outbreak and the moment generating function of the generation time distribution. Since the time lag between all infectee/infecter pairs is not directly observable, the generation time distribution in practice is often substituted with the serial interval distribution that measures the time between symptoms onset. Besides the method of statistical exponential growth, other statistical approaches to estimating  $\mathcal{R}_0$  include sequential Bayesian ([Bettencourt and Ribeiro, 2008](#)) and maximum likelihood estimations ([White and Pagano, 2008](#)). See [White et al. \(2020\)](#) for a recent review of the statistical estimation of the reproduction number. Many researchers have adopted the statistical exponential growth model to estimate the  $\mathcal{R}_0$  of Covid-19 using early case data for China (e.g., [Li et al., 2020](#); [Liu et al., 2020](#); [Sanche et al., 2020](#); [Zhao et al., 2020](#)).

To facilitate comparisons, [Table 1](#) provides a summary of reported  $\mathcal{R}_0$  estimates and respective methods from selected studies. This table complements earlier summaries that focus on the  $\mathcal{R}_0$  estimates for China ([Alimohamadi et al., 2020](#); [Liu et al., 2020](#)) by showcasing the more recent estimates for various locations, with special emphasis on the U.S. estimates.<sup>3</sup>

### 3 Empirical Methodology

Our econometric approach follows from a simplified moment condition of an individual-based stochastic network SIR model developed by [Pesaran and Yang \(2022\)](#). Assuming a homogeneous population,<sup>4</sup> they derive the following moment condition in  $c_t$  (per capita number of cumulative

---

<sup>2</sup>Some authors categorize this strand of methods separately as stochastic approaches or stochastic Markov Chain Monte Carlo (MCMC) methods.

<sup>3</sup>[Alimohamadi et al. \(2020\)](#) found that the estimates of  $\mathcal{R}_0$  based on 23 studies using Chinese data ranged from 1.90 to 6.49, with a mean value of 3.38. [Liu et al. \(2020\)](#) surveyed 12 studies reporting a mean  $\mathcal{R}_0$  estimate for China of 3.28, with a range of 1.4 to 6.49.

<sup>4</sup>[Pesaran and Yang \(2022\)](#) also considered a heterogeneous population segmented into multiple groups that can be based on demographic and socioeconomic factors as well as contact locations/schedules. Using simulations, they find that when  $n$  is sufficiently large and the rate of infection is reasonably rapid, there are little differences in the time profiles of infected cases at the aggregate level whether one uses a single-group or multi-group network model, which partly justify our use of single-group moment conditions for estimation of  $\mathcal{R}_0$ .



Table 1: Reported  $\mathcal{R}_0$  estimates for Covid-19 from selected studies

Study	Locations	Methods	$\mathcal{R}_0$ estimates
Arias et al. (2023)	Belgium	SIRD model, MCMC	Slightly above 4
Buckman et al. (2020)	Brazil, China, Italy, US	SEIR model	Range from 4.8 to 11.4 (US=9.7)
Chudik et al. (2022)	9 Euro countries	Structural econometric approach	Range from 5.01 to 5.51
Fernández-Villaverde and Jones (2022)	Various cities, states, countries	SIRD model	Range from 1.5 to 2.4 (US=2.02)
Katul et al. (2020)	57 countries	SIR model	4.5
Ke et al. (2021)	US and 8 Euro countries	SEIR model	Range from 3.6 and 6.4 (US=5.9 [4.7–7.5])
Korolev (2021)	California, Japan, US	SEIRD model	Range from 1.69 to 19.55
Li et al. (2020)	China	Statistical exponential growth model	2.2 [1.4–3.9]
Liu et al. (2020)	China	Statistical exponential growth model	2.90 [2.32–3.63]
		Statistical maximum likelihood estimation	2.92 [2.28–3.67]
Locatelli et al. (2021)	15 Euro countries	Statistical exponential growth model	2.2 [1.9–2.6]
Sanche et al. (2020)	China	Statistical exponential growth model	5.7 [3.8–8.9]
Wu et al. (2020)	China	SEIR model, MCMC	2.68 [2.47–2.86]
Zhao et al. (2020)	China	Statistical exponential growth model	Range from 2.24 [1.96–2.55] to 3.58 [2.89–4.39]

Note: The 95% confidence intervals are in brackets.

cases on day  $t$ ) and  $i_t$  (per capita number of active infected cases on day  $t$ ):

$$E\left(\frac{1 - c_{t+1}}{1 - c_t} \mid i_t\right) = e^{-\beta_t i_t} + O(n^{-1}), \quad (1)$$

where  $\gamma$  is the recovery rate (per day) and is assumed to be time-invariant;  $\beta_t$  denotes the transmission rate, which can vary over time; and  $n$  stands for the population size. Under this setup,  $i_t$  can be computed from current and past values of  $c_t$  for a given choice of  $\gamma$ :

$$i_t = c_t - \gamma \sum_{\ell=0}^{\infty} (1 - \gamma)^\ell c_{t-\ell-1},$$

or recursively, using  $i_t = (1 - \gamma)i_{t-1} + \Delta c_t$ , where  $\Delta c_t = c_t - c_{t-1}$  denotes per-capita daily new cases, with  $c_t = i_t = 0$ , for  $t \leq 0$  (before the start of the epidemic).

Since  $n$  is large and  $\beta_t i_t$  is small, Eq. (1) implies that  $\beta_t$  can be approximated by  $-i_t^{-1} \ln\left(\frac{1 - c_{j,t+1}}{1 - c_{jt}}\right)$ . To account for additional factors that may influence disease transmission, CPR model the time

variations of  $\beta_t$  in terms of a set of variables measuring mandatory or voluntary social distancing, economic support, vaccine uptake, and virus mutations. We adopt CPR’s strategy and estimate the following threshold panel data model:

$$\beta_{jt}/\gamma = \frac{-\ln\left(\frac{1-c_{j,t+1}}{1-c_{jt}}\right)}{\gamma l_{jt}} = \alpha_j + \boldsymbol{\psi}' \mathbf{x}_{j,t-p} + \kappa I(\Delta c_{j,t-p} > \tau) + u_{j,t+1}, \quad (2)$$

in  $n$  states or countries indexed by  $j = 1, 2, \dots, n$ , over the days  $t = 1, 2, \dots, T$ . The dependent variable,  $\beta_{jt}/\gamma$ , measures the transmission rate in country  $j$  scaled by the recovery rate,  $\gamma$ . One can also interpret  $1/\gamma$  as the mean infectious period. Based on the clinical evidence for Covid-19 and in line with quarantine policies, we set  $\gamma = 1/14$ , which is also a commonly adopted value in previous studies.<sup>5</sup>

Turning to the right-hand side variables of Eq. (2),  $\mathbf{x}_{j,t-p}$  is a vector of explanatory variables lagged  $p$  days, and  $u_{j,t+1}$  are the idiosyncratic errors. The precautionary behavior is captured by the term  $I(\Delta c_{j,t-p} > \tau)$ , where  $I(\cdot)$  is the indicator function equal to one if  $\Delta c_{j,t-p} > \tau$  and zero otherwise.  $\Delta c_{j,t-p}$  is the (seven-day moving average of) reported daily new cases per 100,000 people also lagged  $p$  periods. We chose this as the threshold variable since it has been the most closely monitored metric on the spread of Covid-19 in media reports worldwide.  $\tau$  is a threshold parameter and, for simplicity, is assumed to be the same across countries. Intuitively, higher values of  $\Delta c_{j,t-p}$  imply greater perceived risk; individuals are likely to practice voluntary social distancing more vigorously if  $\Delta c_{j,t-p}$  exceeds the threshold value  $\tau$ .

The parameters to be estimated include  $\alpha_j$ ,  $\boldsymbol{\psi}$ ,  $\kappa$ , and  $\tau$ . In the present study, the key structural parameter of interest is the state- or country-specific intercept,  $\alpha_j$ , which we use as the estimator of the mean basic production number,  $\mathcal{R}_0$ , for state (or country)  $j$ .<sup>6</sup> This is valid because in the absence of any mitigating interventions, whether mandatory and/or voluntary, both  $\mathbf{x}_{j,t-p}$  and  $I(\Delta c_{j,t-p} > \tau)$  would equal to zero, and then  $\beta_{jt}/\gamma = \beta_{j0}/\gamma$  for all  $t$ . Hence, the intercept serves as an estimate of  $\beta_{j0}/\gamma$ , which is indeed the same as  $\mathcal{R}_0$  in the standard SIR models. To identify  $\alpha_j$ , however, it is important that we take into account the mitigating effects of behavioral changes in response to the spread of the epidemic, uptake of vaccination, and possible mutations of the virus. It is also worth noting that  $c_{jt}$  is very small at the onset of the outbreak, and hence  $-\ln\left(\frac{1-c_{j,t+1}}{1-c_{jt}}\right) \approx \Delta c_{j,t+1}$ . This approximation, coupled with the fact that  $\mathbf{x}_{j,t-p}$  and  $I(\Delta c_{j,t-p} > \tau)$  are both equal to zero at the initial stages of the epidemic, makes Eq. (2) consistent with the basic idea of estimating  $\mathcal{R}_0$  by the average growth of infections per active

---

<sup>5</sup>If data on the recovered cases are available, one can also estimate the recovery rate using, for example, the recovery moment condition outlined in [Pesaran and Yang \(2022\)](#). However, we have opted to calibrate the recovery rate value based on clinical information due to its greater reliability, and this is the only parameter we need to calibrate.

<sup>6</sup>In contrast to CPR who also considered a common intercept term for the nine European countries in their study, we focus on the fixed-effects (FE) estimation because our study covers a larger set of heterogeneous countries and states in the U.S.

cases over its infectious period.

We estimate model (2) over two sample periods. The first sample covers the period from March 6, 2020, to January 31, 2021, prior to the roll-out of Covid-19 vaccinations. The second sample extends the first one to November 30, 2021. We will refer to these two samples as the pre-vaccination sample and the full sample, respectively. Both samples are unbalanced since the dates of the outbreaks of Covid-19 and the available case data differ across U.S. states and countries.

Apart from the threshold indicator variable, we also include policy stringency and economic support indices as mitigating factors for both sample periods. For the full sample, we include two additional factors, namely the share of the population that is fully vaccinated and the share of the Delta variant of Covid-19. The rationale behind including the latter factor stems from its greater transmissibility as compared to the earlier variants.

For the U.S. state-level regressions, we also include the interaction between the economic support index and a dummy variable for a Republican governor. This allows us to account for the varied effects of economic policies on the spread of Covid-19 across states with different political leanings, since the economic support index does not include support to firms or take into account the total fiscal value of economic support provided (Hale et al., 2022).

We utilized data from multiple sources to conduct our analysis. The policy stringency and economic support indices were retrieved from the Oxford Covid-19 Government Response Tracker (OxCGRT) project.<sup>7</sup> The data on Delta variants were obtained from CoVariants.org (Hodcroft, 2021). Our World in Data was the source of the data on Covid-19 vaccinations as well as the country-level Covid case numbers.<sup>8</sup> The Centers for Disease Control and Prevention’s Covid data tracker was used to obtain state-by-state Covid cases in the U.S.<sup>9</sup>

To reduce the impact of within-week seasonality in the reported daily cases due to delayed reporting or reduced testing on weekends, we follow the standard practice and take seven-day moving averages of the reported data before estimation. Another important data issue that needs to be addressed is the under-reporting of confirmed cases. This occurs mainly due to asymptomatic infections and lack of testing, especially during the early stages of an outbreak. In the epidemiological literature, the magnitude of under-reporting is often measured by the multiplication factor (MF), which is defined as the ratio of true to reported cases (e.g., Gibbons et al., 2014). For Covid-19, with increased testing we would expect MF to fall over time, and simulations carried out by (Pesaran and Yang, 2022) support this. To account for this declining under-reporting, we consider two different specifications of the MF values: one where the MF linearly declines from 5 to 2, and another where it declines from 8 to 2.5.

We compute three types of standard errors in all estimations. The first type is the “usual”

---

<sup>7</sup>Available at <https://github.com/OxCGRT/covid-policy-dataset>.

<sup>8</sup>Available at <https://github.com/owid/covid-19-data/tree/master/public/data>.

<sup>9</sup>Downloaded from <https://data.cdc.gov/api/views/9mfq-cb36/rows.csv?accessType=DOWNLOAD> (last accessed: May 2023).

standard error assuming no cross-sectional or serial correlation in the errors. The second type, labeled as “robust1” in the tables, is the Newey-West type heteroskedasticity and autocorrelation consistent standard error (Newey and West, 1987). The third type, labeled “robust2”, is the Driscoll and Kraay (1998) standard error, which accounts for both cross-sectional and serial correlation. For the latter two robust standard errors, we choose the truncation lag as the integer part of  $T^{1/3}$ , where  $T$  denotes the longest time span of the unbalanced panel.

## 4 The Main Findings

This section presents the estimates of  $\mathcal{R}_0$  for the U.S. states and 19 countries. In assessing the estimates’ statistical significance, we will focus on the most conservative standard errors that are robust to both error cross-sectional and serial correlation. Estimates for the mitigating covariates and their standard errors are provided in the online supplement.

### 4.1 Estimates of $\mathcal{R}_0$ for U.S. states

We use the estimates of  $\alpha_j$  from the panel threshold regressions, (2), after filtering out the effects of the mitigating factors, to obtain an estimate of  $\mathcal{R}_0$  for the  $j^{\text{th}}$  state. These estimates are summarized in Table 2 for all 48 contiguous states in the U.S. We report two sets of estimates for each of the two sample periods that end on January 31, 2021, and November 30, 2021, respectively. We scale up all reported cases by  $MF_t$  that declines linearly from 5 to 2 in one scenario and from 8 to 2.5 in another scenario.<sup>10</sup>

Overall, the results exhibit remarkable similarities across all states and sample periods, with all estimates being statistically highly significant, even when the most robust standard errors are used. The results are also reasonably robust to the choice of MF and the sample period. When using lower MF values, the estimates of  $\mathcal{R}_0$  vary between 4.21 (Oklahoma) to 4.88 (Rhode Island) in the pre-vaccination sample, and between 4.44 (Mississippi) and 4.95 (Rhode Island) in the full sample. With higher MF values, the estimates rise slightly and vary between 4.47 (New Hampshire) to 5.29 (Rhode Island) in the case of the pre-vaccination sample, and 4.47 (Vermont) to 5.26 (Rhode Island) when we use the full sample.

A closer inspection of results in Table 2 reveals that the estimates of  $\mathcal{R}_0$  tend to be slightly higher in states such as Rhode Island, California, Connecticut, and New Jersey, and lower in states such as Missouri, Oklahoma, and Alabama.<sup>11</sup> Despite these differences, the estimates are very tightly clustered. This can be observed from the histograms in Figure 1, which give the distribution of the estimates for the two sample periods and MF specifications. With lower MF

<sup>10</sup>The parameter estimates for the mitigating factors are provided in Table S.1 of the online supplement.

<sup>11</sup>For a visual representation of the spatial distribution of the  $\mathcal{R}_0$  estimates, refer to the hexagon maps presented in Figures S.1 and S.2 of the online supplement.

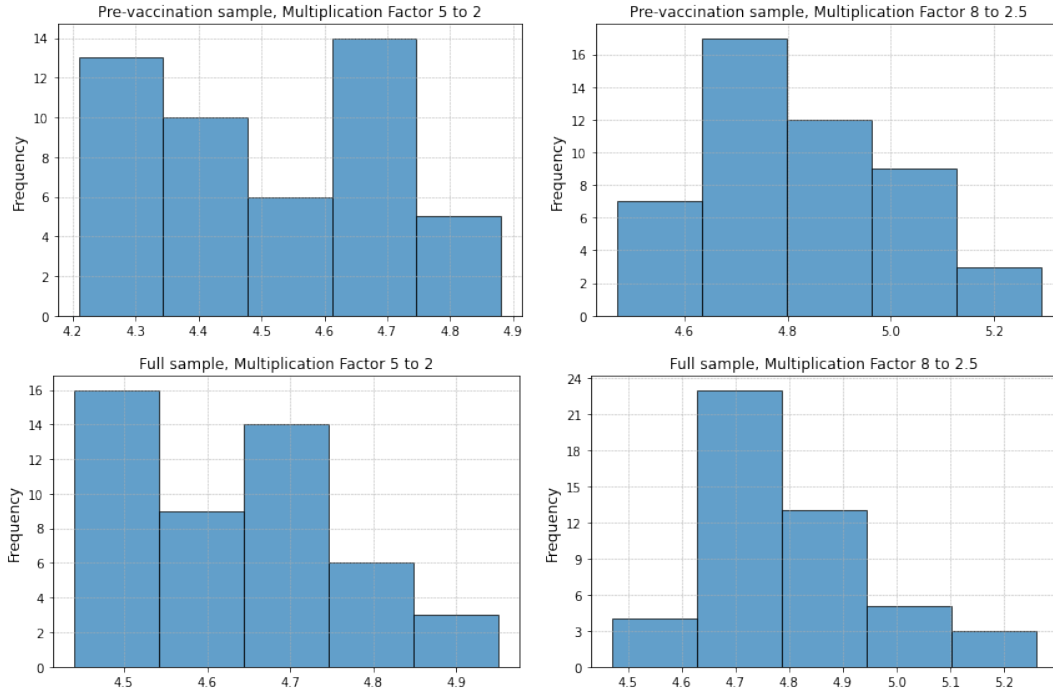
Table 2: Estimates of  $\mathcal{R}_0$  across 48 contiguous U.S. states

State	Pre-vaccination sample		Full sample	
	Ending Jan 31, 2021		Ending Nov 30, 2021	
	MF	MF	MF	MF
	5 to 2	8 to 2.5	5 to 2	8 to 2.5
Alabama	4.29 [0.92]	4.60 [0.86]	4.45 [0.81]	4.68 [0.76]
Arizona	4.62 [0.98]	4.92 [0.90]	4.70 [0.83]	4.94 [0.78]
Arkansas	4.35 [0.97]	4.66 [0.90]	4.50 [0.83]	4.76 [0.79]
California	4.77 [0.97]	5.08 [0.88]	4.94 [0.90]	5.11 [0.87]
Colorado	4.62 [0.96]	4.92 [0.87]	4.67 [0.85]	4.79 [0.80]
Connecticut	4.82 [1.02]	5.13 [0.92]	4.79 [0.87]	4.91 [0.82]
Delaware	4.72 [1.01]	5.04 [0.92]	4.75 [0.87]	4.93 [0.81]
District of Columbia	4.74 [1.09]	5.04 [0.98]	4.68 [0.89]	4.77 [0.84]
Florida	4.41 [0.97]	4.73 [0.89]	4.53 [0.83]	4.74 [0.78]
Georgia	4.41 [0.96]	4.76 [0.88]	4.48 [0.83]	4.70 [0.78]
Idaho	4.36 [1.03]	4.63 [0.95]	4.54 [0.85]	4.77 [0.80]
Illinois	4.68 [0.98]	5.00 [0.89]	4.72 [0.84]	4.90 [0.79]
Indiana	4.34 [0.94]	4.66 [0.87]	4.53 [0.84]	4.76 [0.79]
Iowa	4.29 [0.84]	4.73 [0.79]	4.53 [0.80]	4.78 [0.76]
Kansas	4.65 [1.00]	4.95 [0.92]	4.69 [0.84]	4.88 [0.79]
Kentucky	4.72 [1.06]	5.01 [0.96]	4.73 [0.87]	4.92 [0.82]
Louisiana	4.70 [1.03]	5.01 [0.94]	4.73 [0.89]	4.97 [0.84]
Maine	4.67 [1.15]	4.91 [1.04]	4.71 [0.91]	4.74 [0.86]
Maryland	4.68 [1.01]	4.98 [0.92]	4.66 [0.86]	4.73 [0.81]
Massachusetts	4.60 [1.05]	4.85 [0.96]	4.79 [0.90]	4.95 [0.84]
Michigan	4.56 [1.05]	4.82 [0.96]	4.71 [0.88]	4.86 [0.83]
Minnesota	4.57 [1.03]	4.85 [0.94]	4.77 [0.87]	4.94 [0.82]
Mississippi	4.34 [0.95]	4.67 [0.89]	4.44 [0.82]	4.67 [0.77]
Missouri	4.26 [0.94]	4.53 [0.87]	4.46 [0.82]	4.64 [0.77]
Montana	4.27 [0.99]	4.58 [0.92]	4.55 [0.84]	4.77 [0.79]
Nebraska	4.35 [0.96]	4.71 [0.89]	4.51 [0.82]	4.77 [0.77]
Nevada	4.39 [1.01]	4.69 [0.93]	4.55 [0.86]	4.78 [0.81]
New Hampshire	4.23 [0.96]	4.47 [0.89]	4.49 [0.85]	4.56 [0.80]
New Jersey	4.73 [0.98]	5.07 [0.89]	4.75 [0.85]	4.91 [0.80]
New Mexico	4.79 [1.23]	5.07 [1.11]	4.87 [0.94]	5.06 [0.88]
New York	4.85 [1.01]	5.18 [0.91]	4.66 [0.87]	4.70 [0.82]
North Carolina	4.69 [1.04]	4.96 [0.95]	4.66 [0.87]	4.81 [0.82]
North Dakota	4.24 [0.89]	4.68 [0.84]	4.67 [0.80]	5.17 [0.76]
Ohio	4.41 [1.05]	4.66 [0.96]	4.56 [0.87]	4.72 [0.81]
Oklahoma	4.21 [0.87]	4.50 [0.82]	4.48 [0.82]	4.72 [0.77]
Oregon	4.55 [0.97]	4.82 [0.89]	4.60 [0.86]	4.62 [0.81]
Pennsylvania	4.64 [0.98]	4.92 [0.90]	4.65 [0.84]	4.75 [0.79]
Rhode Island	4.88 [1.01]	5.29 [0.92]	4.95 [0.88]	5.26 [0.83]
South Carolina	4.32 [0.92]	4.64 [0.85]	4.51 [0.82]	4.75 [0.77]
South Dakota	4.25 [0.88]	4.65 [0.83]	4.64 [0.81]	5.05 [0.76]
Tennessee	4.43 [1.03]	4.73 [0.95]	4.59 [0.85]	4.92 [0.80]
Texas	4.45 [1.01]	4.73 [0.93]	4.54 [0.85]	4.73 [0.80]
Utah	4.33 [0.96]	4.65 [0.89]	4.55 [0.82]	4.87 [0.77]
Vermont	4.23 [1.05]	4.59 [0.97]	4.45 [0.88]	4.47 [0.83]
Virginia	4.36 [0.98]	4.64 [0.91]	4.48 [0.85]	4.59 [0.80]
Washington	4.65 [0.99]	4.94 [0.91]	4.64 [0.88]	4.68 [0.83]
Wisconsin	4.57 [0.96]	4.89 [0.88]	4.76 [0.85]	4.99 [0.80]
Wyoming	4.55 [1.02]	4.77 [0.94]	4.58 [0.84]	4.69 [0.79]

Notes: The pre-vaccination sample includes 14,996 observations for  $N = 48$  states, covering a time span of  $T_{\min} = 163$  to  $T_{\max} = 331$  days. The full sample comprises 29,531 observations, spanning  $T_{\min} = 457$  to  $T_{\max} = 634$  days. Both estimation samples are unbalanced at the beginning. Numbers in brackets are standard errors robust to serial correlation and cross-sectional correlation (robust2). The multiplication factor (MF) declines linearly from the high value to the low value. The estimates on the mitigating factors are provided in Table S.1 in the online supplement.

values, the estimates are primarily clustered around 4.2 to 4.8. When the MF is higher, about 40 out of the 48 states have estimates concentrated between 4.6 and 5.1. To summarize, an average estimate (across states, periods, and MF values) of 4.7 seems to provide a reasonable summary number for the U.S.

Figure 1: Histograms of  $\mathcal{R}_0$  estimates across 48 contiguous U.S. states



Notes: The pre-vaccination sample ends on January 31, 2021, and the full sample ends on November 30, 2021, for all states. Both estimation samples are unbalanced at the beginning.

## 4.2 Country-specific Estimates

Table 3 reports the estimates of  $\mathcal{R}_0$  for 19 countries.<sup>12</sup> Similar to our findings for the U.S. states, the estimates fall within a narrow range. Moreover, the results are fairly robust to the choice of MF, with higher MF values only marginally increasing the  $\mathcal{R}_0$  estimates. Argentina, Chile, Peru, and Spain are found to have the highest average  $\mathcal{R}_0$  values, while Nigeria, South Korea, and Indonesia have the lowest average  $\mathcal{R}_0$  values. Specifically, for the pre-vaccination period, the estimates range from 3.85 (South Korea) to 5.01 (Argentina) under the lower MF values, and 3.84 (South Korea) to 5.04 (Argentina) under the higher MF values. The full-sample estimates lie between 3.42 for Nigeria and 4.46 for Chile (3.44 for Nigeria and 4.56 for Spain) in the low

<sup>12</sup>The associated estimation results for the mitigating factors of the country regressions are provided in Table S.4 in the online supplement, where we also present the  $\mathcal{R}_0$  estimates in descending bar charts in Figures S.3 and S.4.

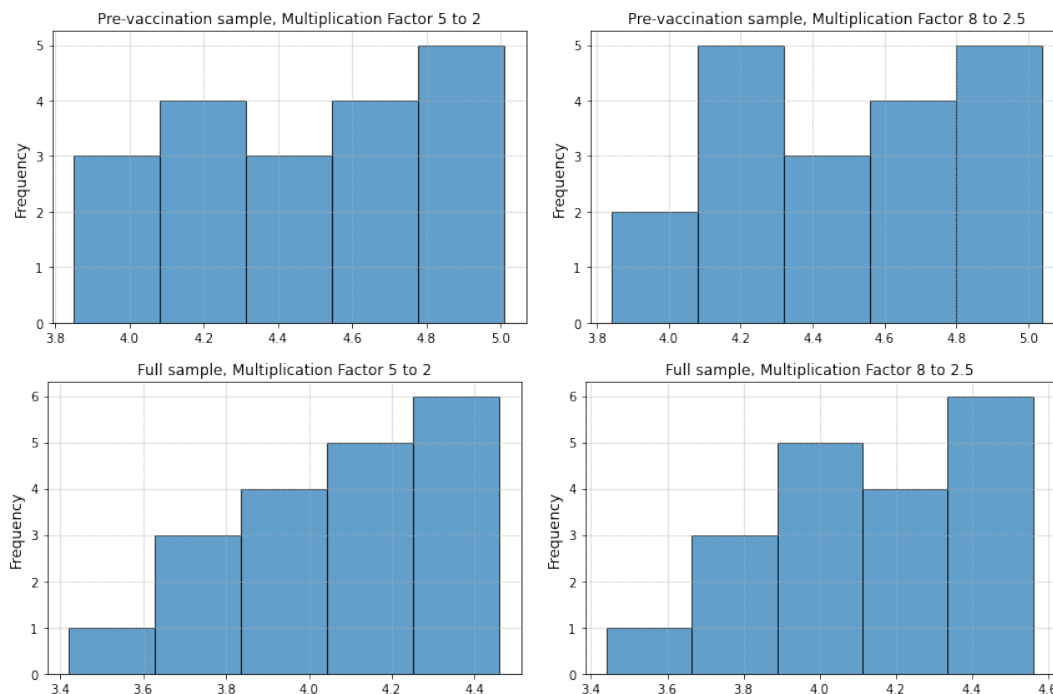
Table 3: Estimates of  $\mathcal{R}_0$  across 19 countries

Country	Pre-vaccination sample		Full sample	
	Ending Jan 31, 2021		Ending Nov 30, 2021	
	MF 5 to 2	MF 8 to 2.5	MF 5 to 2	MF 8 to 2.5
Argentina	5.01 [0.51]	5.04 [0.50]	4.41 [0.61]	4.51 [0.59]
Australia	4.27 [0.49]	4.26 [0.48]	3.92 [0.59]	3.92 [0.57]
Brazil	4.47 [0.45]	4.52 [0.44]	4.08 [0.56]	4.17 [0.54]
Chile	4.85 [0.48]	4.90 [0.47]	4.46 [0.61]	4.53 [0.59]
Colombia	4.83 [0.46]	4.86 [0.45]	4.32 [0.57]	4.41 [0.55]
Ecuador	4.69 [0.48]	4.70 [0.48]	4.22 [0.59]	4.24 [0.57]
Egypt	4.54 [0.41]	4.53 [0.41]	3.95 [0.53]	3.97 [0.51]
France	4.58 [0.45]	4.61 [0.44]	4.35 [0.57]	4.47 [0.55]
Germany	4.18 [0.44]	4.19 [0.43]	4.08 [0.58]	4.11 [0.56]
Indonesia	4.08 [0.40]	4.08 [0.39]	3.80 [0.54]	3.82 [0.52]
Japan	4.16 [0.36]	4.14 [0.36]	3.97 [0.52]	3.97 [0.50]
Mexico	4.21 [0.40]	4.22 [0.40]	3.83 [0.52]	3.85 [0.50]
Nigeria	3.88 [0.40]	3.88 [0.40]	3.42 [0.49]	3.44 [0.48]
Peru	4.89 [0.45]	4.93 [0.45]	4.31 [0.59]	4.38 [0.56]
Philippines	4.37 [0.43]	4.38 [0.42]	3.88 [0.56]	3.90 [0.54]
South Korea	3.85 [0.37]	3.84 [0.37]	3.78 [0.51]	3.77 [0.49]
Spain	4.83 [0.49]	4.87 [0.48]	4.44 [0.61]	4.56 [0.58]
Thailand	4.64 [0.54]	4.63 [0.53]	4.12 [0.57]	4.14 [0.55]
Turkey	4.69 [0.52]	4.70 [0.51]	4.22 [0.60]	4.28 [0.58]

Notes: The pre-vaccination sample includes 5,928 observations for 19 countries, covering a time span of  $T_{\min} = 140$  to  $T_{\max} = 329$  days. The full sample comprises 11,661 observations for the same 19 countries, spanning  $T_{\min} = 443$  to  $T_{\max} = 632$  days. Both estimation samples are unbalanced at the beginning. Numbers in brackets are standard errors robust to serial correlation and cross-sectional correlation (robust2). The multiplication factor declines linearly from the high value to the low value. The estimates on the mitigating factors are provided in Table S.4 in the online supplement.

(high) MF scenarios. The slight differences in estimated  $\mathcal{R}_0$  values across countries might be associated with varying degrees of under-reporting. To further examine the distribution of the estimates, Figure 2 displays the histograms for each sample period and MF specification. We see that the estimates from the pre-vaccination sample are relatively evenly distributed within the range, whereas about 15 out of the 19 countries have full sample estimates falling in the range of 3.8 to 4.5. In sum, the average estimate of  $\mathcal{R}_0$  for 19 countries across both sample periods and MF values is around 4.3, which is quite close to the average U.S. estimate of 4.7. Overall, both our U.S. states and international estimates align with the recent literature and suggest that earlier studies have underestimated  $\mathcal{R}_0$ , as documented in Table 1.

Figure 2: Histograms of  $\mathcal{R}_0$  estimates across 19 countries



Notes: The pre-vaccination sample ends on January 31, 2021, and the full sample ends on November 30, 2021, for all states. Both estimation samples are unbalanced at the beginning.

### 4.3 Estimates of $\mathcal{R}_0$ without Mitigating Factors

To demonstrate the importance of filtering out the effects of the mitigating factors in the estimation of  $\mathcal{R}_0$ , we estimated Eq. (2) without any of the mitigating factors (or by setting both  $\mathbf{x}_{j,t-p}$  and  $I(\Delta c_{j,t-p} > \tau)$  to zero). The state- and country-specific estimates are presented in Tables S.7 and S.8 in the online supplement. Evidently, the estimated  $\mathcal{R}_0$ 's are significantly biased downward when the mitigating factors are not accounted for. The average estimate across sample periods and MF specifications is only 1.5 (1.3) for U.S. states (19 countries). These results underscore the problem of omitted variable bias resulting from neglecting mitigating factors.

## 5 Conclusions

In this paper, we propose a novel approach to the estimation of the basic reproduction number,  $\mathcal{R}_0$ , of Covid-19, and provide estimates for U.S. states and a selected number of countries. Our approach falls under the category of counterfactual causal analysis, where the focus is to filter out the effects of mitigating factors on the diffusion of the virus, and thus identify  $\mathcal{R}_0$  when the values of the mitigating factors are set to zero in the counterfactual exercise. Our estimates of  $\mathcal{R}_0$  turn out to be centered around 4.5, clustered closely across U.S. states and a selected number



of countries. Not allowing for the mitigating factors results in estimates of  $\mathcal{R}_0$  that suffer from substantial downward bias.

While our estimation approach is relatively simple to implement and yields satisfactory estimates, it is subject to an important limitation. In order to identify and provide accurate estimates of  $\mathcal{R}_0$ , it requires the availability of reliable data on mitigating factors, ideally taking into account all such factors in the analysis. In practice, this means that the method might not be applicable at the very early stages of an epidemic. Nevertheless, we believe our approach offers a useful alternative to the existing methods for  $\mathcal{R}_0$  estimation, which also face the challenge of obtaining reliable early samples that are not subject to any mitigating interventions.

## References

- Alimohamadi, Y., M. Taghdir, and M. Sepandi (2020). Estimate of the basic reproduction number for COVID-19: A systematic review and meta-analysis. *Journal of Preventive Medicine and Public Health* 53(3), 151–157.
- Arias, J. E., J. Fernández-Villaverde, J. F. Rubio-Ramírez, and M. Shin (2023). The causal effects of lockdown policies on health and macroeconomic outcomes. *American Economic Journal: Macroeconomics* 15(3), 287–319.
- Atkeson, A., K. A. Kopecky, and T. A. Zha (2023). Four stylized facts about COVID-19. *International Economic Review* (forthcoming).
- Bettencourt, L. M. A. and R. M. Ribeiro (2008). Real time Bayesian estimation of the epidemic potential of emerging infectious diseases. *PLoS ONE* 3(5), e2185.
- Buckman, S. R., R. Glick, K. J. Lansing, N. Petrosky-Nadeau, and L. M. Seitzman (2020). Replicating and projecting the path of COVID-19 with a model-implied reproduction number. *Infectious Disease Modelling* 5, 635–651.
- Chudik, A., M. H. Pesaran, and A. Rebucci (2022). Social distancing, vaccination and evolution of COVID-19 transmission rates in Europe. *IMF Economic Review*, 1–35.
- Del Valle, S. Y., J. M. Hyman, and N. Chitnis (2013). Mathematical models of contact patterns between age groups for predicting the spread of infectious diseases. *Mathematical Biosciences and Engineering* 10, 1475.
- Delamater, P. L., E. J. Street, T. F. Leslie, Y. T. Yang, and K. H. Jacobsen (2019). Complexity of the basic reproduction number ( $R_0$ ). *Emerging Infectious Diseases* 25(1), 1–4.
- Driscoll, J. C. and A. C. Kraay (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics* 80(4), 549–560.
- Fernández-Villaverde, J. and C. I. Jones (2022). Estimating and simulating a SIRD model of COVID-19 for many countries, states, and cities. *Journal of Economic Dynamics and Control*, 104318.
- Fraser, C. (2007). Estimating individual and household reproduction numbers in an emerging epidemic. *PLoS ONE* 2(8), e758.
- Gibbons, C. L., M.-J. J. Mangen, D. Plass, A. H. Havelaar, R. J. Brooke, P. Kramarz, et al. (2014). Measuring underreporting and under-ascertainment in infectious disease datasets: A comparison of methods. *BMC public health* 14(1), 1–17.
- Hale, T., N. Angrist, B. Kira, A. Petherick, T. Phillips, S. Webster, et al. (2022). Variation in government responses to COVID-19. BSG Working Paper 2020/032.
- Hodcroft, E. B. (2021). CoVariants: SARS-CoV-2 mutations and variants of interest. <https://covariants.org/>.

- Katul, G. G., A. Mrad, S. Bonetti, G. Manoli, and A. J. Parolari (2020). Global convergence of COVID-19 basic reproduction number and estimation from early-time SIR dynamics. *PLOS ONE* 15(9), e0239800.
- Ke, R., E. Romero-Severson, S. Sanche, and N. Hengartner (2021). Estimating the reproductive number  $R_0$  of SARS-CoV-2 in the United States and eight European countries and implications for vaccination. *Journal of Theoretical Biology* 517, 110621.
- Korolev, I. (2021). Identification and estimation of the SEIRD epidemic model for COVID-19. *Journal of Econometrics* 220(1), 63–85.
- Lessler, J. and D. A. T. Cummings (2016). Mechanistic models of infectious disease and their impact on public health. *American Journal of Epidemiology* 183(5), 415–422.
- Li, Q., X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, et al. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *New England Journal of Medicine* 382(13), 1199–1207.
- Liu, T., J. Hu, J. Xiao, G. He, M. Kang, Z. Rong, et al. (2020). Time-varying transmission dynamics of Novel Coronavirus Pneumonia in China. *bioRxiv* 2020.
- Liu, Y., A. A. Gayle, A. Wilder-Smith, and J. Rocklöv (2020). The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine* 27(2), 1–4.
- Locatelli, I., B. Trächsel, and V. Rousson (2021). Estimating the basic reproduction number for COVID-19 in Western Europe. *PloS ONE* 16(3), e0248731.
- Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55(3), 703.
- Obadia, T., R. Haneef, and P.-Y. Boëlle (2012). The  $R_0$  package: A toolbox to estimate reproduction numbers for epidemic outbreaks. *BMC Medical Informatics and Decision Making* 12(1), 1–9.
- Pesaran, M. H. and C. F. Yang (2022). Matching theory and evidence on Covid-19 using a stochastic network SIR model. *Journal of Applied Econometrics* 37(6), 1204–1229.
- Sanche, S., Y. T. Lin, C. Xu, E. Romero-Severson, N. Hengartner, and R. Ke (2020). High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. *Emerging Infectious Diseases* 26(7), 1470–1477.
- Tang, L., Y. Zhou, L. Wang, S. Purkayastha, L. Zhang, J. He, F. Wang, and P. X.-K. Song (2020). A review of multi-compartment infectious disease models. *International Statistical Review* 88(2), 462–513.
- Wallinga, J. and M. Lipsitch (2007). How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences* 274(1609), 599–604.
- White, L. F., C. B. Moser, R. N. Thompson, and M. Pagano (2020, oct). Statistical estimation of the reproductive number from case notification data. *American Journal of Epidemiology* 190(4), 611–620.
- White, L. F. and M. Pagano (2008). A likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic. *Statistics in Medicine* 27(16), 2999–3016.
- Wu, J. T., K. Leung, and G. M. Leung (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *The Lancet* 395(10225), 689–697.
- Zhao, S., Q. Lin, J. Ran, S. S. Musa, G. Yang, W. Wang, et al. (2020). Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in china, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *International Journal of Infectious Diseases* 92, 214–217.

## Statements and Declarations

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript. The authors have no relevant financial or non-financial interests to disclose.

# Online Supplement to “Structural Econometric Estimation of the Basic Reproduction Number for Covid-19 Across U.S. States and Selected Countries”

Ida Johnsson

Chewy, Inc., and University of Southern California

M. Hashem Pesaran

University of Southern California, USA, and Trinity College, Cambridge, UK

Cynthia Fan Yang

Florida State University

August 30, 2023

This online supplement provides the coefficient estimates for the mitigating factors in the panel threshold regressions and their summary statistics for the U.S. states and 19 nations considered in the main paper. Furthermore, it presents additional graphs illustrating the  $\mathcal{R}_0$  estimates to facilitate the understanding of their clustering and spatial distribution patterns. It also reports the estimated  $\mathcal{R}_0$  without any mitigating factors to showcase the biases in the resulting estimates.

First, Table [S.1](#) reports the estimates on the mitigating factors in the panel regressions for the 48 contiguous U.S. states. It complements the  $\mathcal{R}_0$  (or intercept) estimates shown in Table [2](#) of the main paper. The values of the multiplication factor (MF) are assumed to decline linearly from 5 to 2 and from 8 to 2.5. The summary statistics of the mitigating factors are given in Tables [S.2](#) and [S.3](#). The results reveal that the stringency index is negatively associated with the transmission rate, and the association is highly significant for the full sample period. In states with a Democrat governor, economic support is found to contribute to the slowdown of transmission in most cases, although its effect is statistically and practically insignificant. In comparison, having a Republican governor tends to adversely influence the effectiveness of economic support in reducing disease transmission. In addition, a higher proportion of the fully vaccinated population is found to significantly lower the transmission rate, while a larger share of the Delta variant leads to a significant increase in disease spread, as expected. The threshold value reflects when individuals begin to take precautionary measures due to the fear of getting infected with the increasing number of cases. The low threshold values (0.7 for the high-MF pre-vaccination sample and 0.4 for all other cases) suggest that voluntary social distancing took place soon after the outbreak, and it significantly lowers the transmission rate, as evidenced by the large negative estimates on the threshold variable. Changing the MF specifications have minimal impact on our findings.

Turning next to the country-level regressions, we provide the coefficient estimates for the mitigating factors in Table [S.4](#), which complements the intercept estimates given in Table [3](#) of

the main paper. The summary statistics of the mitigating factors can be found in Tables S.5 and S.6. Both the stringency index and economic support index are found to significantly reduce Covid-19 transmission, with the stringency policies having a greater effect. The threshold effect is also negative and highly significant, reflecting a considerable impact of precautionary behaviors on disease propagation. In all cases, the threshold parameter is estimated to be 0.01, indicating that the voluntary behavioral changes occurred shortly after the outbreak, which is consistent with our finding for the U.S. In addition, the full-sample results suggest that the share of the population fully vaccinated has a negative and strongly significant effect in decreasing the spread of the disease. The estimates on the share of the Delta variant are positive although insignificant. Their lack of accuracy is likely due to the poor data quality on the Delta variant in less developed countries in our sample. Overall, altering the MF assumptions produces negligible changes in the coefficients on all factors.

In addition, we provide supplementary figures to illustrate the distributions of the  $\mathcal{R}_0$  estimates presented in Tables 2 and 3 of the main paper. Hexagon maps, displayed in Figures S.1 and S.2, have been included to illustrate the spatial distribution of the  $\mathcal{R}_0$  estimates across U.S. states. To visualize the country-specific  $\mathcal{R}_0$  estimates, Figures S.3 and S.4 display the estimates in descending order for different sample periods and MF specifications.

Lastly, as discussed in Section 4.3 of the main paper, we provide the  $\mathcal{R}_0$  estimates without mitigating factors in Tables S.7 and S.8 for the U.S. states and countries, respectively. These tables serve to demonstrate the substantial underestimation of  $\mathcal{R}_0$  resulting from disregarding mitigating factors.

Table S.1: Estimation results for the mitigating factors across 48 contiguous U.S. states

State	Pre-vaccination sample		Full sample	
	Ending Jan 31, 2021		Ending Nov 30, 2021	
	MF	MF	MF	MF
	5 to 2	8 to 2.5	5 to 2	8 to 2.5
<b>Stringency Index</b>	-1.28	-1.27	-1.07	-1.51
standard s.e. (t-ratio)	0.08 (-15.4)	0.09 (-14.5)	0.11 (-9.6)	0.13 (-11.5)
robust1 s.e. (t-ratio)	0.15 (-8.7)	0.14 (-9.1)	0.14 (-7.7)	0.15 (-9.8)
robust2 s.e. (t-ratio)	1.16 (-1.1)	0.97 (-1.3)	0.64 (-1.7)	0.60 (-2.5)
<b>Economic Support</b>	-0.39	-0.63	-0.03	0.02
standard s.e. (t-ratio)	0.06 (-6.3)	0.06 (-9.9)	0.07 (-0.4)	0.09 (0.3)
robust1 s.e. (t-ratio)	0.12 (-3.2)	0.12 (-5.2)	0.09 (-0.3)	0.10 (0.2)
robust2 s.e. (t-ratio)	0.32 (-1.2)	0.33 (-1.9)	0.20 (-0.1)	0.23 (0.1)
<b>Vaccinated Share</b>			-1.65	-1.51
standard s.e. (t-ratio)	-	-	0.10 (-17.2)	0.11 (-13.3)
robust1 s.e. (t-ratio)	-	-	0.08 (-19.8)	0.10 (-15.2)
robust2 s.e. (t-ratio)	-	-	0.40 (-4.2)	0.44 (-3.4)
<b>Delta Variant Share</b>			1.00	1.00
standard s.e. (t-ratio)	-	-	0.04 (22.8)	0.05 (19.4)
robust1 s.e. (t-ratio)	-	-	0.04 (24.7)	0.05 (18.9)
robust2 s.e. (t-ratio)	-	-	0.18 (5.5)	0.24 (4.1)
<b>Rep × Economic Support</b>	0.42	0.39	0.26	0.18
standard s.e. (t-ratio)	0.07 (6.1)	0.07 (5.4)	0.09 (3.0)	0.10 (1.8)
robust1 s.e. (t-ratio)	0.15 (2.8)	0.14 (2.7)	0.10 (2.7)	0.11 (1.7)
robust2 s.e. (t-ratio)	0.14 (3.0)	0.14 (2.7)	0.09 (2.8)	0.10 (1.8)
<b>Threshold Indicator</b>	-2.31	-2.37	-2.71	-2.49
standard s.e. (t-ratio)	0.04 (-62.2)	0.04 (-54.6)	0.07 (-39.1)	0.08 (-30.6)
robust1 s.e. (t-ratio)	0.11 (-20.3)	0.12 (-19.1)	0.13 (-20.7)	0.13 (-19.3)
robust2 s.e. (t-ratio)	0.50 (-4.6)	0.53 (-4.5)	0.59 (-4.6)	0.55 (-4.5)
Threshold Value	0.70	0.40	0.40	0.40
R-squared	0.31	0.29	0.09	0.08

Notes: The pre-vaccination sample includes 14,996 observations for 48 contiguous states, covering a time span of  $T_{\min} = 163$  to  $T_{\max} = 331$  days. The full sample comprises 29,531 observations, spanning  $T_{\min} = 457$  to  $T_{\max} = 634$  days. Both estimation samples are unbalanced at the beginning. Rep is a dummy variable equal to one for a Republican governor. “Robust1” standard errors are robust to serial correlation only (Newey-West type correction), whereas “robust2” standard errors are robust to serial correlation as well as any cross-sectional correlation. Figures in parentheses are t-ratios. Oxford stringency and economic support indices are divided by 100 so that they take values between zero and one. Lag order is set to  $p = 10$  days in all regressions. The multiplication factor (MF) declines linearly from the high value to the low value.

Table S.2: Summary statistics by state for the pre-vaccination sample (ending January 31, 2021)

	stringency		support		thr	
	mean	std	mean	std	mean	std
Alabama	0.52	(0.12)	0.43	(0.18)	0.97	(0.18)
Arizona	0.58	(0.07)	0.47	(0.14)	0.97	(0.18)
Arkansas	0.57	(0.09)	0.44	(0.18)	0.97	(0.16)
California	0.65	(0.12)	0.70	(0.16)	0.93	(0.25)
Colorado	0.62	(0.12)	0.70	(0.18)	0.97	(0.18)
Connecticut	0.69	(0.08)	0.76	(0.16)	0.97	(0.17)
Delaware	0.67	(0.12)	0.71	(0.16)	0.98	(0.13)
District of Columbia	0.73	(0.08)	0.68	(0.18)	0.99	(0.10)
Florida	0.59	(0.17)	0.57	(0.27)	0.95	(0.21)
Georgia	0.61	(0.11)	0.64	(0.25)	0.96	(0.20)
Idaho	0.61	(0.12)	0.40	(0.21)	0.97	(0.17)
Illinois	0.62	(0.14)	0.61	(0.24)	0.96	(0.19)
Indiana	0.57	(0.09)	0.59	(0.25)	0.96	(0.19)
Iowa	0.53	(0.10)	0.81	(0.15)	0.97	(0.17)
Kansas	0.61	(0.08)	0.51	(0.23)	0.97	(0.17)
Kentucky	0.69	(0.11)	0.61	(0.25)	0.97	(0.18)
Louisiana	0.64	(0.11)	0.50	(0.25)	0.97	(0.17)
Maine	0.76	(0.12)	0.60	(0.24)	0.99	(0.10)
Maryland	0.67	(0.15)	0.68	(0.19)	0.96	(0.19)
Massachusetts	0.64	(0.11)	0.42	(0.19)	0.96	(0.20)
Michigan	0.64	(0.15)	0.44	(0.17)	0.96	(0.19)
Minnesota	0.63	(0.09)	0.46	(0.14)	0.97	(0.18)
Mississippi	0.56	(0.13)	0.48	(0.24)	0.97	(0.16)
Missouri	0.53	(0.12)	0.42	(0.18)	0.97	(0.18)
Montana	0.57	(0.07)	0.41	(0.20)	0.97	(0.17)
Nebraska	0.59	(0.10)	0.58	(0.15)	0.97	(0.16)
Nevada	0.61	(0.09)	0.50	(0.23)	0.97	(0.17)
New Hampshire	0.57	(0.16)	0.50	(0.29)	0.98	(0.13)
New Jersey	0.62	(0.13)	0.70	(0.18)	0.97	(0.18)
New Mexico	0.81	(0.04)	0.48	(0.12)	0.98	(0.15)
New York	0.72	(0.11)	0.95	(0.22)	0.95	(0.23)
North Carolina	0.67	(0.08)	0.59	(0.10)	0.96	(0.21)
North Dakota	0.47	(0.05)	0.41	(0.20)	1.00	(0.00)
Ohio	0.64	(0.09)	0.46	(0.15)	0.96	(0.21)
Oklahoma	0.47	(0.14)	0.39	(0.21)	0.97	(0.18)
Oregon	0.64	(0.09)	0.71	(0.16)	0.97	(0.18)
Pennsylvania	0.63	(0.11)	0.63	(0.19)	0.96	(0.20)
Rhode Island	0.70	(0.07)	0.86	(0.07)	0.99	(0.10)
South Carolina	0.54	(0.11)	0.51	(0.19)	0.97	(0.18)
South Dakota	0.47	(0.08)	0.43	(0.17)	0.99	(0.08)
Tennessee	0.61	(0.09)	0.42	(0.19)	0.97	(0.18)
Texas	0.59	(0.10)	0.43	(0.18)	0.95	(0.21)
Utah	0.54	(0.07)	0.39	(0.21)	0.97	(0.17)
Vermont	0.69	(0.09)	0.67	(0.17)	0.96	(0.19)
Virginia	0.61	(0.10)	0.57	(0.21)	0.96	(0.20)
Washington	0.66	(0.12)	0.75	(0.19)	0.97	(0.18)
West Virginia	0.62	(0.09)	0.49	(0.14)	0.98	(0.14)
Wisconsin	0.58	(0.11)	0.52	(0.21)	0.97	(0.18)
All	0.62	(0.13)	0.56	(0.24)	0.97	(0.18)

Table S.3: Summary statistics by state for the full sample (ending November 30, 2021)

	delta		stringency		support		thr		vaxx	
	mean	std	mean	std	mean	std	mean	std	mean	std
Alabama	0.26	(0.41)	0.43	(0.13)	0.39	(0.14)	0.98	(0.13)	0.13	(0.17)
Arizona	0.25	(0.41)	0.46	(0.16)	0.41	(0.12)	0.98	(0.13)	0.17	(0.21)
Arkansas	0.28	(0.43)	0.48	(0.13)	0.41	(0.15)	0.99	(0.11)	0.14	(0.18)
California	0.26	(0.41)	0.55	(0.17)	0.67	(0.12)	0.96	(0.19)	0.19	(0.24)
Colorado	0.27	(0.42)	0.52	(0.16)	0.55	(0.24)	0.98	(0.13)	0.20	(0.25)
Connecticut	0.25	(0.41)	0.55	(0.19)	0.67	(0.21)	0.98	(0.13)	0.23	(0.29)
Delaware	0.25	(0.41)	0.55	(0.16)	0.62	(0.17)	0.99	(0.09)	0.19	(0.24)
District of Columbia	0.24	(0.41)	0.59	(0.17)	0.65	(0.15)	1.00	(0.07)	0.20	(0.25)
Florida	0.24	(0.41)	0.46	(0.20)	0.44	(0.27)	0.98	(0.15)	0.18	(0.23)
Georgia	0.25	(0.41)	0.50	(0.16)	0.49	(0.29)	0.98	(0.15)	0.14	(0.18)
Idaho	0.24	(0.41)	0.49	(0.16)	0.39	(0.16)	0.99	(0.12)	0.14	(0.18)
Illinois	0.25	(0.42)	0.51	(0.18)	0.59	(0.18)	0.98	(0.14)	0.18	(0.23)
Indiana	0.26	(0.42)	0.48	(0.14)	0.47	(0.22)	0.98	(0.14)	0.16	(0.20)
Iowa	0.27	(0.42)	0.43	(0.14)	0.58	(0.28)	0.99	(0.12)	0.19	(0.23)
Kansas	0.28	(0.43)	0.49	(0.15)	0.53	(0.18)	0.99	(0.12)	0.17	(0.21)
Kentucky	0.24	(0.41)	0.56	(0.18)	0.48	(0.22)	0.98	(0.13)	0.17	(0.21)
Louisiana	0.24	(0.41)	0.58	(0.10)	0.42	(0.20)	0.98	(0.13)	0.15	(0.18)
Maine	0.23	(0.41)	0.58	(0.21)	0.49	(0.21)	1.00	(0.07)	0.23	(0.29)
Maryland	0.24	(0.41)	0.54	(0.19)	0.64	(0.21)	0.98	(0.14)	0.21	(0.27)
Massachusetts	0.25	(0.41)	0.56	(0.16)	0.37	(0.18)	0.98	(0.15)	0.22	(0.29)
Michigan	0.24	(0.41)	0.53	(0.19)	0.40	(0.14)	0.98	(0.14)	0.17	(0.22)
Minnesota	0.25	(0.42)	0.51	(0.17)	0.37	(0.18)	0.98	(0.13)	0.20	(0.25)
Mississippi	0.27	(0.42)	0.46	(0.16)	0.43	(0.19)	0.99	(0.12)	0.13	(0.17)
Missouri	0.29	(0.44)	0.45	(0.14)	0.40	(0.15)	0.98	(0.14)	0.16	(0.20)
Montana	0.27	(0.42)	0.45	(0.14)	0.39	(0.15)	0.99	(0.12)	0.19	(0.21)
Nebraska	0.27	(0.43)	0.46	(0.15)	0.64	(0.22)	0.98	(0.14)	0.19	(0.23)
Nevada	0.27	(0.42)	0.51	(0.16)	0.39	(0.23)	0.99	(0.12)	0.17	(0.21)
New Hampshire	0.25	(0.41)	0.48	(0.15)	0.44	(0.22)	0.99	(0.09)	0.21	(0.26)
New Jersey	0.26	(0.42)	0.51	(0.17)	0.66	(0.15)	0.98	(0.13)	0.21	(0.27)
New Mexico	0.25	(0.41)	0.63	(0.21)	0.43	(0.12)	0.99	(0.11)	0.21	(0.26)
New York	0.25	(0.41)	0.58	(0.20)	0.84	(0.29)	0.97	(0.17)	0.20	(0.26)
North Carolina	0.25	(0.41)	0.56	(0.16)	0.50	(0.22)	0.98	(0.15)	0.16	(0.20)
North Dakota	0.25	(0.42)	0.40	(0.09)	0.46	(0.18)	1.00	(0.00)	0.16	(0.19)
Ohio	0.25	(0.41)	0.52	(0.16)	0.40	(0.13)	0.98	(0.15)	0.17	(0.21)
Oklahoma	0.27	(0.42)	0.44	(0.11)	0.38	(0.17)	0.98	(0.13)	0.16	(0.19)
Oregon	0.23	(0.40)	0.55	(0.16)	0.65	(0.13)	0.98	(0.13)	0.20	(0.25)
Pennsylvania	0.25	(0.42)	0.49	(0.18)	0.56	(0.16)	0.98	(0.14)	0.19	(0.24)
Rhode Island	0.25	(0.41)	0.57	(0.17)	0.70	(0.19)	1.00	(0.07)	0.23	(0.28)
South Carolina	0.25	(0.41)	0.46	(0.13)	0.40	(0.22)	0.98	(0.13)	0.15	(0.19)
South Dakota	0.27	(0.43)	0.40	(0.11)	0.39	(0.13)	1.00	(0.06)	0.19	(0.22)
Tennessee	0.25	(0.41)	0.50	(0.16)	0.43	(0.14)	0.98	(0.13)	0.15	(0.18)
Texas	0.25	(0.41)	0.49	(0.13)	0.41	(0.13)	0.98	(0.15)	0.16	(0.21)
Utah	0.28	(0.43)	0.44	(0.12)	0.38	(0.16)	0.99	(0.12)	0.16	(0.20)
Vermont	0.33	(0.45)	0.51	(0.21)	0.69	(0.21)	0.99	(0.11)	0.30	(0.30)
Virginia	0.25	(0.41)	0.50	(0.16)	0.55	(0.20)	0.98	(0.14)	0.20	(0.25)
Washington	0.25	(0.41)	0.57	(0.17)	0.69	(0.20)	0.98	(0.13)	0.20	(0.26)
West Virginia	0.24	(0.41)	0.49	(0.17)	0.44	(0.13)	0.99	(0.10)	0.15	(0.17)
Wisconsin	0.25	(0.41)	0.49	(0.14)	0.41	(0.21)	0.98	(0.13)	0.19	(0.24)
All	0.26	(0.42)	0.51	(0.17)	0.50	(0.22)	0.98	(0.13)	0.18	(0.23)

Table S.4: Estimation results for the mitigating factors across 19 countries

Country	Pre-vaccination sample		Full sample	
	Ending Jan 31, 2021		Ending Nov 30, 2021	
	MF 5 to 2	MF 8 to 2.5	MF 5 to 2	MF 8 to 2.5
<b>Stringency Index</b>	-1.68	-1.71	-1.04	-1.07
standard s.e. (t-ratio)	0.14 (-11.8)	0.14 (-11.9)	0.09 (-11.1)	0.10 (-10.8)
robust1 s.e. (t-ratio)	0.26 (-6.5)	0.26 (-6.7)	0.19 (-5.6)	0.19 (-5.7)
robust2 s.e. (t-ratio)	0.35 (-4.7)	0.35 (-4.9)	0.32 (-3.2)	0.31 (-3.4)
<b>Economic Support</b>	-1.03	-1.01	-0.61	-0.63
standard s.e. (t-ratio)	0.09 (-11.9)	0.09 (-11.6)	0.05 (-11.1)	0.06 (-10.8)
robust1 s.e. (t-ratio)	0.13 (-7.8)	0.13 (-7.7)	0.09 (-6.4)	0.09 (-6.6)
robust2 s.e. (t-ratio)	0.15 (-6.8)	0.15 (-6.8)	0.14 (-4.4)	0.13 (-4.7)
<b>Vaccinated Share</b>			-0.78	-0.62
standard s.e. (t-ratio)	-	-	0.09 (-8.6)	0.09 (-6.6)
robust1 s.e. (t-ratio)	-	-	0.12 (-6.6)	0.12 (-5.0)
robust2 s.e. (t-ratio)	-	-	0.18 (-4.3)	0.19 (-3.4)
<b>Delta Variant Share</b>			0.12	0.07
standard s.e. (t-ratio)	-	-	0.06 (2.2)	0.06 (1.2)
robust1 s.e. (t-ratio)	-	-	0.07 (1.7)	0.08 (1.0)
robust2 s.e. (t-ratio)	-	-	0.12 (1.0)	0.12 (0.6)
<b>Threshold Indicator</b>	-1.44	-1.42	-1.81	-1.79
standard s.e. (t-ratio)	0.12 (-12.2)	0.12 (-11.9)	0.09 (-19.1)	0.10 (-18.0)
robust1 s.e. (t-ratio)	0.31 (-4.7)	0.30 (-4.7)	0.32 (-5.6)	0.32 (-5.6)
robust2 s.e. (t-ratio)	0.24 (-6.0)	0.24 (-5.9)	0.35 (-5.2)	0.34 (-5.2)
Threshold Value	0.01	0.01	0.01	0.01
R-squared	0.13	0.13	0.09	0.08

Notes: The pre-vaccination sample includes 5,928 observations for 19 countries, covering a time span of  $T_{\min} = 140$  to  $T_{\max} = 329$  days. The full sample comprises 11,661 observations for the same 19 countries, spanning  $T_{\min} = 443$  to  $T_{\max} = 632$  days. Both estimation samples are unbalanced at the beginning. “Robust1” standard errors are robust to serial correlation only (Newey-West type correction), whereas “robust2” standard errors are robust to serial correlation as well as any cross-sectional correlation. Figures in parentheses are t-ratios. Oxford stringency and economic support indices are divided by 100 so that they take values between zero and one. Lag order is set to  $p = 10$  days in all regressions. The multiplication factor (MF) declines linearly from the high value to the low value.



Table S.5: Summary statistics by country for the pre-vaccination sample (ending January 31, 2021)

	Stringency		Support		Threshold Ind.	
	mean	std	mean	std	mean	std
<b>Argentina</b>	0.86	(0.11)	0.72	(0.15)	0.99	(0.08)
<b>Australia</b>	0.65	(0.12)	0.69	(0.17)	1.00	(0.00)
<b>Brazil</b>	0.68	(0.13)	0.42	(0.17)	0.97	(0.16)
<b>Chile</b>	0.79	(0.15)	0.71	(0.31)	0.99	(0.11)
<b>Colombia</b>	0.76	(0.13)	0.72	(0.12)	0.98	(0.12)
<b>Ecuador</b>	0.71	(0.17)	0.72	(0.12)	0.96	(0.21)
<b>Egypt</b>	0.72	(0.15)	0.68	(0.17)	0.98	(0.14)
<b>France</b>	0.64	(0.17)	0.68	(0.24)	0.97	(0.17)
<b>Germany</b>	0.62	(0.13)	0.42	(0.16)	0.98	(0.13)
<b>Indonesia</b>	0.63	(0.09)	0.28	(0.10)	0.96	(0.19)
<b>Japan</b>	0.37	(0.07)	0.85	(0.26)	0.98	(0.13)
<b>Mexico</b>	0.70	(0.15)	0.25	(0.35)	0.98	(0.15)
<b>Nigeria</b>	0.68	(0.14)	0.15	(0.19)	0.93	(0.26)
<b>Peru</b>	0.81	(0.13)	0.68	(0.16)	0.98	(0.14)
<b>Philippines</b>	0.76	(0.15)	0.42	(0.28)	0.97	(0.17)
<b>South Korea</b>	0.56	(0.09)	0.45	(0.13)	0.98	(0.13)
<b>Spain</b>	0.66	(0.15)	0.81	(0.21)	0.98	(0.13)
<b>Thailand</b>	0.60	(0.16)	0.81	(0.39)	0.84	(0.37)
<b>Turkey</b>	0.64	(0.12)	0.75	(0.26)	0.97	(0.16)
<b>All</b>	0.68	(0.17)	0.58	(0.30)	0.97	(0.16)

Table S.6: Summary statistics by country for the full sample (ending November 30, 2021)

	Delta		Stringency		Support		Threshold Ind.		Vaccinated	
	mean	std	mean	std	mean	std	mean	std	mean	std
<b>Argentina</b>	0.10	(0.23)	0.78	(0.15)	0.60	(0.18)	1.00	(0.06)	0.18	(0.26)
<b>Australia</b>	0.31	(0.43)	0.63	(0.12)	0.41	(0.35)	1.00	(0.00)	0.13	(0.22)
<b>Brazil</b>	0.17	(0.34)	0.64	(0.12)	0.27	(0.20)	0.99	(0.12)	0.17	(0.26)
<b>Chile</b>	0.14	(0.31)	0.74	(0.16)	0.67	(0.39)	0.99	(0.08)	0.26	(0.32)
<b>Colombia</b>	0.14	(0.29)	0.69	(0.16)	0.73	(0.09)	0.99	(0.09)	0.12	(0.20)
<b>Ecuador</b>	0.15	(0.31)	0.67	(0.13)	0.74	(0.08)	0.98	(0.15)	0.14	(0.24)
<b>Egypt</b>	0.21	(0.39)	0.58	(0.18)	0.72	(0.13)	0.99	(0.10)	0.02	(0.05)
<b>France</b>	0.24	(0.41)	0.59	(0.16)	0.57	(0.23)	0.97	(0.16)	0.21	(0.29)
<b>Germany</b>	0.25	(0.41)	0.62	(0.16)	0.40	(0.12)	0.99	(0.10)	0.21	(0.28)
<b>Indonesia</b>	0.29	(0.43)	0.66	(0.08)	0.33	(0.09)	0.98	(0.14)	0.07	(0.13)
<b>Japan</b>	0.21	(0.39)	0.43	(0.08)	0.92	(0.20)	0.99	(0.10)	0.15	(0.27)
<b>Mexico</b>	0.24	(0.40)	0.60	(0.17)	0.26	(0.27)	0.99	(0.11)	0.13	(0.19)
<b>Nigeria</b>	0.26	(0.41)	0.57	(0.15)	0.08	(0.15)	0.96	(0.19)	0.01	(0.01)
<b>Peru</b>	0.13	(0.29)	0.77	(0.12)	0.60	(0.14)	0.99	(0.10)	0.10	(0.18)
<b>Philippines</b>	0.20	(0.37)	0.75	(0.11)	0.24	(0.29)	0.99	(0.12)	0.05	(0.10)
<b>South Korea</b>	0.23	(0.39)	0.54	(0.09)	0.44	(0.12)	0.99	(0.10)	0.15	(0.26)
<b>Spain</b>	0.24	(0.40)	0.61	(0.14)	0.84	(0.15)	0.98	(0.14)	0.22	(0.31)
<b>Thailand</b>	0.34	(0.42)	0.58	(0.11)	0.72	(0.25)	0.95	(0.22)	0.13	(0.21)
<b>Turkey</b>	0.25	(0.39)	0.62	(0.15)	0.54	(0.33)	0.99	(0.11)	0.17	(0.24)
<b>All</b>	0.21	(0.38)	0.64	(0.16)	0.53	(0.31)	0.98	(0.12)	0.14	(0.24)

Table S.7: Estimates of  $\mathcal{R}_0$  across 48 contiguous U.S. states without mitigating factors

State	Pre-vaccination sample		Full sample	
	Ending Jan 31, 2021		Ending Nov 30, 2021	
	MF 5 to 2	MF 8 to 2.5	MF 5 to 2	MF 8 to 2.5
Arizona	1.51 [0.07]	1.81 [0.08]	1.48 [0.11]	1.61 [0.11]
Arkansas	1.45 [0.06]	1.73 [0.07]	1.40 [0.07]	1.52 [0.07]
California	1.66 [0.32]	1.86 [0.37]	1.53 [0.14]	1.60 [0.14]
Colorado	1.38 [0.07]	1.54 [0.07]	1.34 [0.11]	1.41 [0.11]
Connecticut	1.38 [0.09]	1.54 [0.09]	1.41 [0.13]	1.49 [0.13]
Delaware	1.39 [0.07]	1.60 [0.08]	1.34 [0.09]	1.42 [0.09]
District of Columbia	1.25 [0.06]	1.35 [0.07]	1.27 [0.09]	1.34 [0.09]
Florida	1.44 [0.09]	1.67 [0.09]	1.48 [0.13]	1.58 [0.13]
Georgia	1.41 [0.08]	1.63 [0.09]	1.46 [0.11]	1.56 [0.11]
Idaho	1.43 [0.07]	1.67 [0.08]	1.36 [0.10]	1.47 [0.10]
Illinois	1.46 [0.09]	1.68 [0.09]	1.45 [0.13]	1.56 [0.13]
Indiana	1.46 [0.08]	1.70 [0.09]	1.42 [0.12]	1.52 [0.12]
Iowa	1.49 [0.07]	1.79 [0.08]	1.42 [0.09]	1.56 [0.09]
Kansas	1.47 [0.06]	1.71 [0.07]	1.44 [0.08]	1.55 [0.09]
Kentucky	1.42 [0.07]	1.63 [0.08]	1.39 [0.08]	1.47 [0.08]
Louisiana	1.43 [0.09]	1.67 [0.10]	1.47 [0.14]	1.59 [0.14]
Maine	1.22 [0.06]	1.27 [0.06]	1.22 [0.07]	1.24 [0.07]
Maryland	1.30 [0.08]	1.41 [0.08]	1.37 [0.11]	1.43 [0.11]
Massachusetts	1.41 [0.10]	1.59 [0.10]	1.42 [0.14]	1.51 [0.14]
Michigan	1.43 [0.09]	1.60 [0.10]	1.37 [0.13]	1.45 [0.13]
Minnesota	1.47 [0.08]	1.68 [0.09]	1.39 [0.09]	1.48 [0.09]
Mississippi	1.42 [0.07]	1.67 [0.08]	1.41 [0.09]	1.54 [0.09]
Missouri	1.45 [0.07]	1.66 [0.08]	1.38 [0.10]	1.48 [0.10]
Montana	1.44 [0.06]	1.69 [0.07]	1.36 [0.09]	1.47 [0.09]
Nebraska	1.47 [0.07]	1.75 [0.08]	1.38 [0.09]	1.52 [0.09]
Nevada	1.42 [0.07]	1.65 [0.07]	1.39 [0.10]	1.49 [0.10]
New Hampshire	1.29 [0.07]	1.38 [0.07]	1.27 [0.08]	1.30 [0.08]
New Jersey	1.44 [0.10]	1.66 [0.11]	1.46 [0.17]	1.56 [0.17]
New Mexico	1.40 [0.06]	1.58 [0.07]	1.34 [0.08]	1.42 [0.09]
New York	1.30 [0.11]	1.37 [0.11]	1.40 [0.18]	1.44 [0.17]
North Carolina	1.37 [0.07]	1.54 [0.07]	1.41 [0.09]	1.48 [0.09]
North Dakota	1.62 [0.08]	2.17 [0.11]	1.36 [0.07]	1.62 [0.08]
Ohio	1.40 [0.08]	1.57 [0.08]	1.41 [0.11]	1.48 [0.11]
Oklahoma	1.44 [0.07]	1.72 [0.08]	1.40 [0.08]	1.52 [0.08]
Oregon	1.23 [0.06]	1.29 [0.06]	1.26 [0.08]	1.29 [0.08]
Pennsylvania	1.39 [0.09]	1.54 [0.09]	1.40 [0.13]	1.46 [0.13]
Rhode Island	1.49 [0.08]	1.84 [0.10]	1.38 [0.10]	1.52 [0.11]
South Carolina	1.44 [0.08]	1.70 [0.09]	1.44 [0.09]	1.55 [0.09]
South Dakota	1.57 [0.08]	2.04 [0.11]	1.38 [0.09]	1.59 [0.10]
Tennessee	1.49 [0.08]	1.83 [0.10]	1.43 [0.09]	1.56 [0.09]
Texas	1.44 [0.10]	1.65 [0.10]	1.53 [0.15]	1.62 [0.15]
Utah	1.52 [0.06]	1.86 [0.07]	1.43 [0.09]	1.57 [0.09]
Vermont	1.21 [0.08]	1.25 [0.08]	1.26 [0.15]	1.28 [0.15]
Virginia	1.33 [0.07]	1.45 [0.08]	1.40 [0.11]	1.46 [0.11]
Washington	1.26 [0.07]	1.33 [0.07]	1.30 [0.11]	1.34 [0.11]
West Virginia	1.35 [0.06]	1.50 [0.07]	1.31 [0.07]	1.35 [0.07]
Wisconsin	1.49 [0.07]	1.77 [0.08]	1.41 [0.08]	1.54 [0.09]

Notes: Numbers in brackets are standard errors robust to serial correlation and cross-sectional correlation (robust2). The multiplication factor (MF) declines linearly from the high value to the low value.

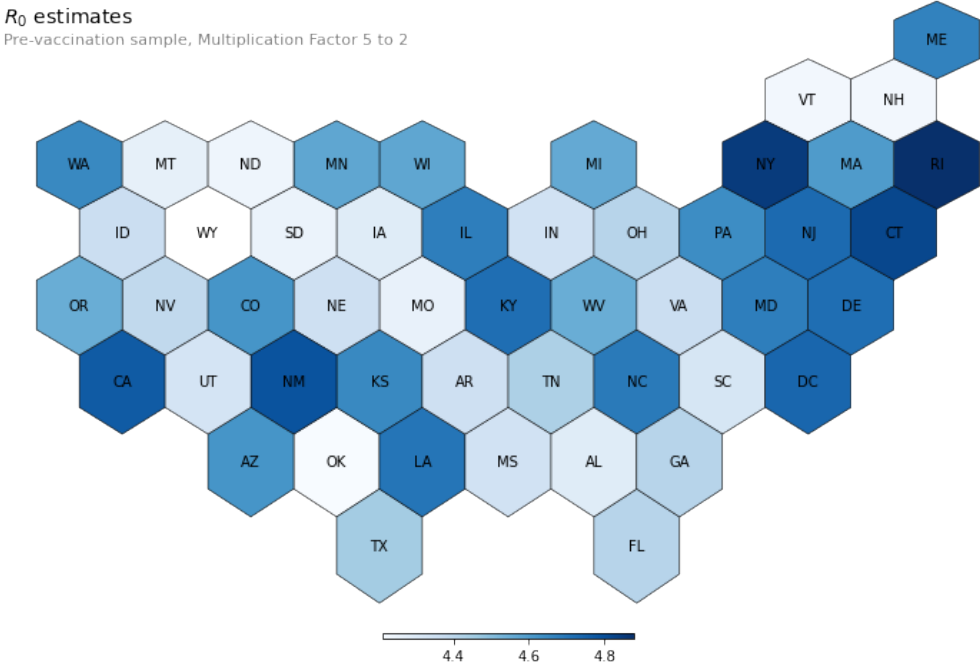
Table S.8: Estimates of  $\mathcal{R}_0$  across 19 countries without mitigating factors

Country	Pre-vaccination sample		Full sample	
	Ending Jan 31, 2021		Ending Nov 30, 2021	
	MF 5 to 2	MF 8 to 2.5	MF 5 to 2	MF 8 to 2.5
Australia	1.14 [0.09]	1.14 [0.09]	1.04 [0.13]	1.04 [0.13]
Brazil	1.45 [0.07]	1.45 [0.07]	1.54 [0.12]	1.54 [0.12]
Chile	1.38 [0.06]	1.38 [0.06]	1.42 [0.09]	1.42 [0.09]
Colombia	1.37 [0.05]	1.37 [0.05]	1.44 [0.06]	1.44 [0.06]
Ecuador	1.23 [0.08]	1.23 [0.08]	1.39 [0.15]	1.39 [0.15]
Egypt	1.12 [0.05]	1.12 [0.05]	1.22 [0.08]	1.22 [0.08]
France	1.69 [0.17]	1.69 [0.17]	1.55 [0.18]	1.55 [0.18]
Germany	1.36 [0.10]	1.36 [0.10]	1.41 [0.14]	1.41 [0.14]
Indonesia	1.12 [0.06]	1.12 [0.06]	1.34 [0.07]	1.34 [0.07]
Japan	1.09 [0.07]	1.09 [0.07]	1.26 [0.08]	1.26 [0.08]
Mexico	1.20 [0.06]	1.20 [0.06]	1.38 [0.08]	1.38 [0.08]
Nigeria	1.07 [0.06]	1.07 [0.06]	1.25 [0.08]	1.25 [0.08]
Peru	1.35 [0.07]	1.35 [0.07]	1.47 [0.11]	1.47 [0.11]
Philippines	1.17 [0.06]	1.17 [0.06]	1.27 [0.09]	1.27 [0.09]
South Korea	1.10 [0.05]	1.10 [0.05]	1.09 [0.09]	1.09 [0.09]
Spain	1.53 [0.12]	1.53 [0.12]	1.59 [0.18]	1.59 [0.18]
Thailand	1.31 [0.14]	1.31 [0.14]	1.59 [0.37]	1.59 [0.37]
Turkey	1.43 [0.12]	1.43 [0.12]	1.46 [0.22]	1.46 [0.22]

Notes: Numbers in brackets are standard errors robust to serial correlation and cross-sectional correlation (robust2). The multiplication factor (MF) declines linearly from the high value to the low value.

Figure S.1: Estimates of  $\mathcal{R}_0$  across 48 contiguous U.S. states, pre-vaccination sample (ending January 31, 2021)

(a) Multiplication Factor 5 to 2



(b) Multiplication Factor 8 to 2.5

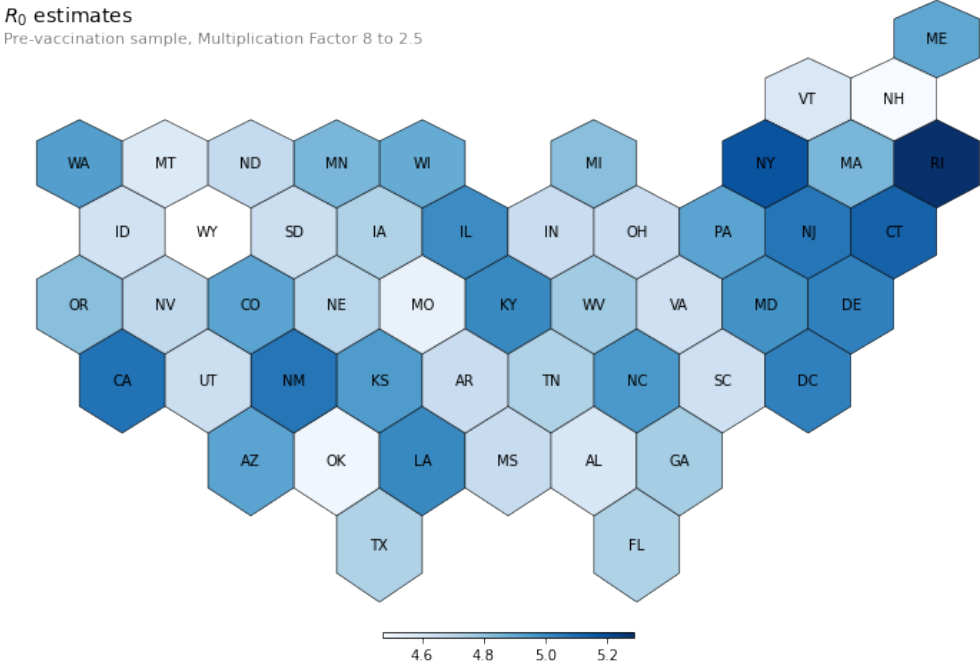
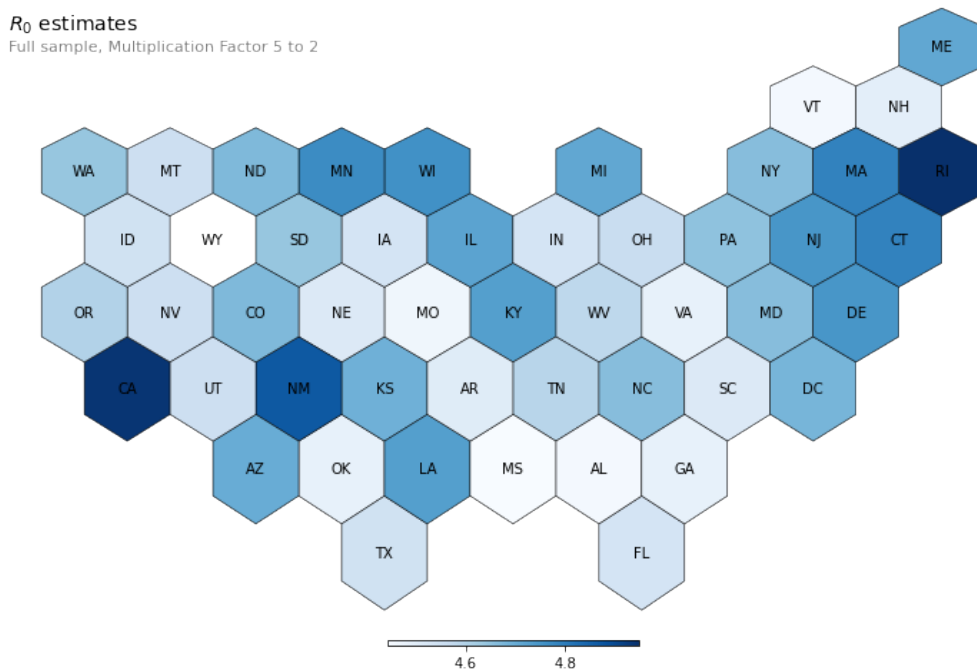


Figure S.2: Estimates of  $\mathcal{R}_0$  across 48 contiguous U.S. states, full sample (ending November 30, 2021)

(a) Multiplication Factor 5 to 2



(b) Multiplication Factor 8 to 2.5

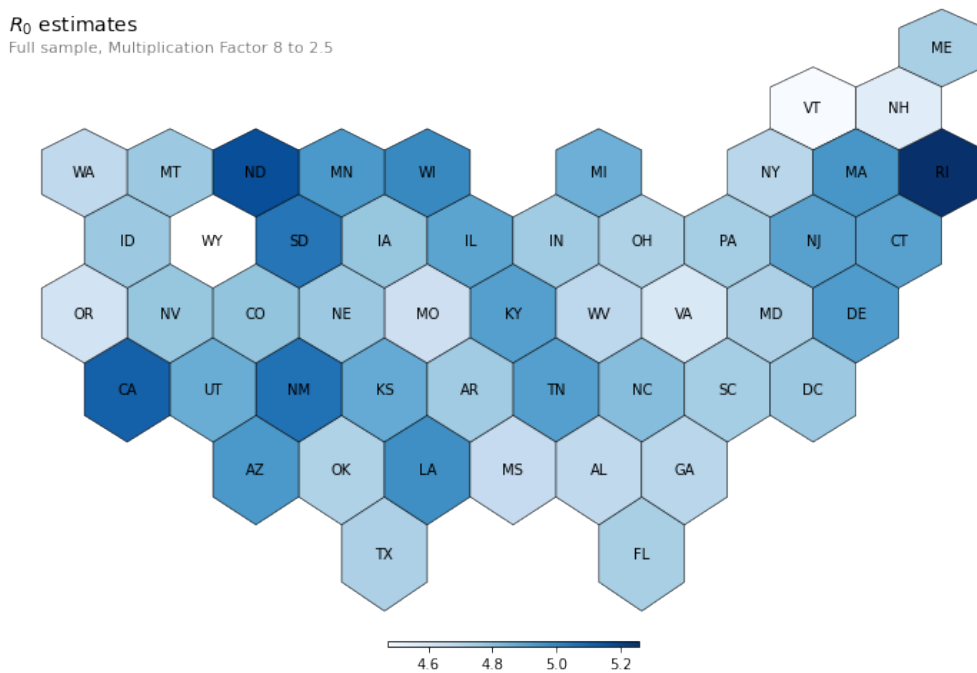
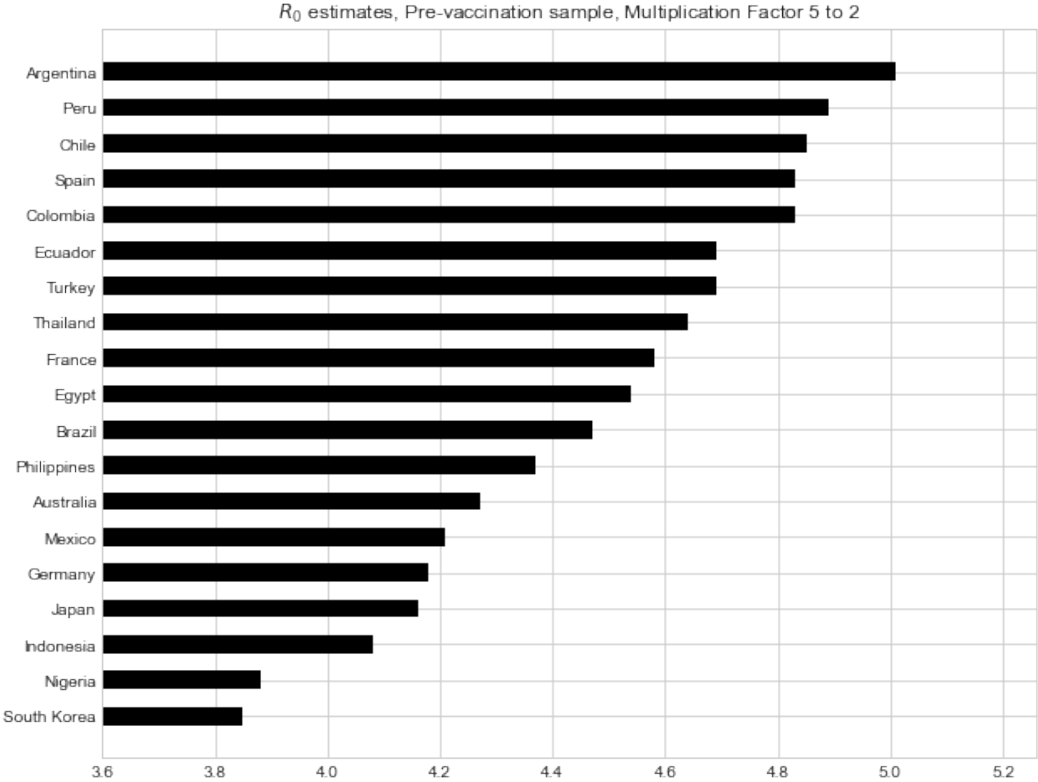


Figure S.3: Estimates of  $\mathcal{R}_0$  across 19 countries, pre-vaccination sample (ending January 31, 2021)

(a) Multiplication Factor 5 to 2



(b) Multiplication Factor 8 to 2.5

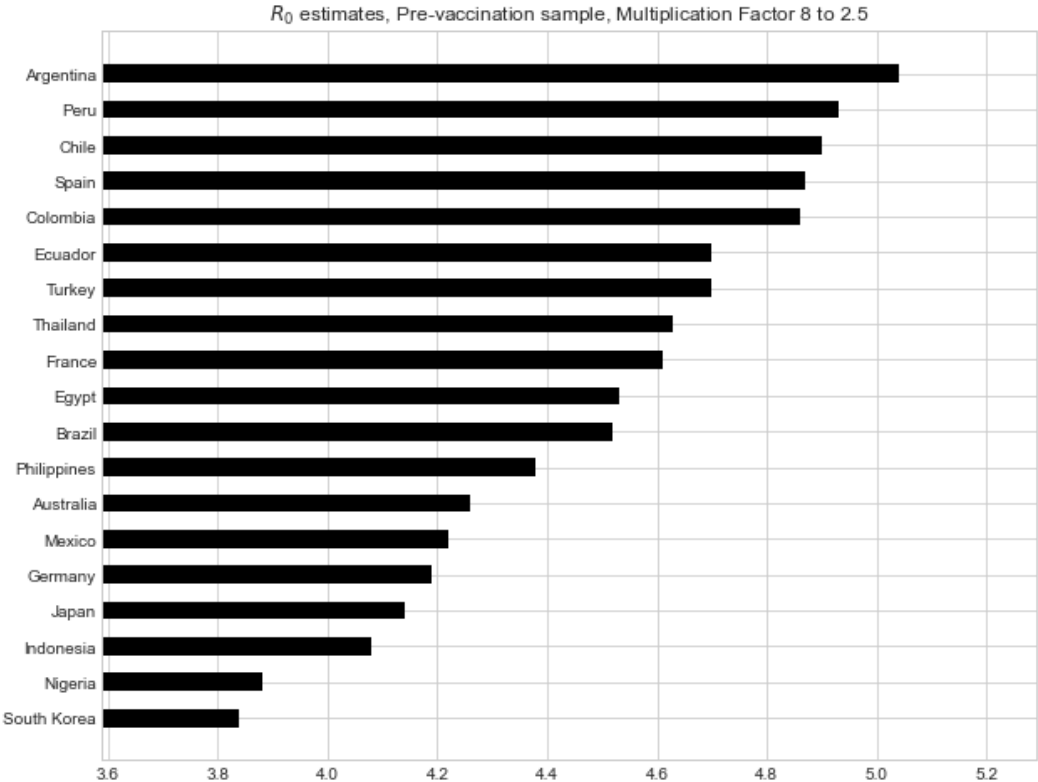
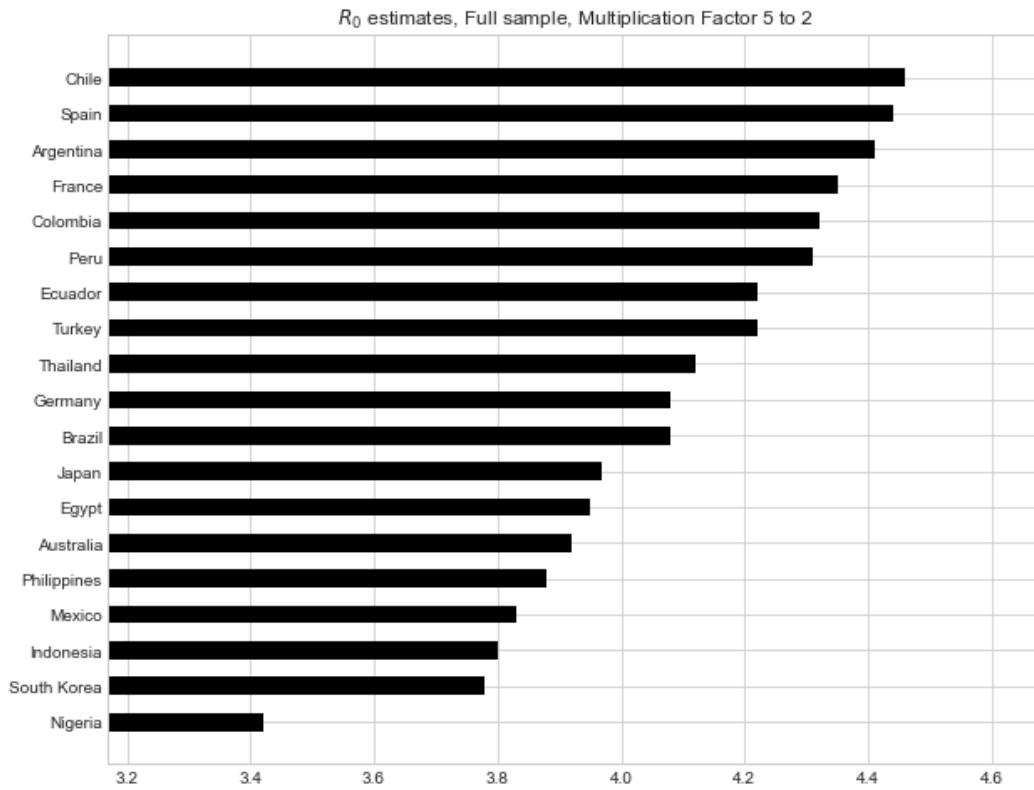


Figure S.4: Estimates of  $\mathcal{R}_0$  across 19 countries, full sample (ending November 30, 2021)

(a) Multiplication Factor 5 to 2



(b) Multiplication Factor 8 to 2.5

