Working Paper Series

How To Go Viral: A COVID-19 Model with Endogenously Time-Varying Parameters

WP 20-10

Paul Ho
Federal Reserve Bank of Richmond

Thomas A. Lubik
Federal Reserve Bank of Richmond

Christian Matthes
Indiana University

FEDERAL RESERVE BANK OF RICHMOND®
Richmond ▪ Baltimore ▪ Charlotte

How To Go Viral: A COVID-19 Model with Endogenously Time-Varying Parameters*

Paul Ho
Federal Reserve Bank of Richmond†

Thomas A. Lubik
Federal Reserve Bank of Richmond‡

Christian Matthes
Indiana University§

August 21, 2020

Abstract

We estimate a panel model with endogenously time-varying parameters for COVID-19 cases and deaths in U.S. states. The functional form for infections incorporates important features of epidemiological models but is flexibly parameterized to capture different trajectories of the pandemic. Daily deaths are modeled as a spike-and-slab regression on lagged cases. Our Bayesian estimation reveals that social distancing and testing have significant effects on the parameters. For example, a 10 percentage point increase in the positive test rate is associated with a 2 percentage point increase in the death rate among reported cases. The model forecasts perform well, even relative to models from epidemiology and statistics.

JEL Classification: C32, C51

Key Words: Bayesian Estimation, Panel, Time-Varying Parameters

*We thank seminar participants at the Federal Reserve Bank of Richmond for helpful comments. James Geary and James Lee provided exceptional research assistance. This research was supported in part through computational resources provided by the Big-Tex High Performance Computing Group at the Federal Reserve Bank of Dallas. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

†Research Department, P.O. Box 27622, Richmond, VA 23261. Email: paul.ho@rich.frb.org.

‡Research Department, P.O. Box 27622, Richmond, VA 23261. Email: thomas.lubik@rich.frb.org.

§Wylie Hall, 100 South Woodlawn Avenue, Bloomington, IN 47405. Email: matthesc@iu.edu.

1 Introduction

A new form of coronavirus, SARS-CoV-2, which causes the respiratory disease COVID-19, appeared in the U.S. in January 2020.[1] Since then, the U.S. has seen over 5 million cases and 170,000 deaths as of mid-August.[2] Any policy response to the pandemic crucially depends on understanding how the virus spreads, how the disease evolves over time, what its effects on mortality rates are, and how factors such as increased testing and measures such as social distancing affect outcomes. We contribute to this effort from a statistical perspective that pays heed to prior epidemiological research.

To that end, we develop and estimate a time series model for the number of cases and the number of deaths in U.S. states that has three key features: (i) it exploits the panel dimension of the data without forcing dynamics to be the same across states, (ii) it is a statistical model that, while using some insights from epidemiological models, is more flexible than common models in epidemiology, and (iii) it features time variation in parameters tied directly to fluctuations in observable predictors to account for the fact that as the pandemic grew, citizens and governments changed their behavior.[3] The model produces accurate forecasts for COVID-19 cases and deaths in the U.S., outperforming many competing models' forecasts for new and cumulative deaths. Our estimates show that increased social distancing and testing are associated with lower numbers of cases, but this association does not hold in all states. In addition, increased testing is associated with lower death rates among reported cases. We estimate the model using Bayesian methods, which allows us to quantify the uncertainty in our forecasts and estimates explicitly.

Our model for the number of infections is based on the observation that the time path of infections during an epidemic follows a typical pattern. When a pathogen enters a population that is susceptible to infection, the number of infected cases is initially low. However, the growth rate of new infections is high and tends to rise sharply at an exponential rate since each infected person creates a chain of new infections. At some point, however, the pathogen runs out of susceptible hosts because they are already infected, immune, or simply not physically present because of health policies such as social distancing. At this inflection point, the growth rate of infections falls until it eventually declines to zero. We replicate these broad patterns of an epidemic by specifying a flexible functional form that describes the path of infections over time as depending on the current and the lagged levels of the number of infections. In contrast to theoretical epidemiological models, our specification

[1] https://en.wikipedia.org/wiki/COVID-19_pandemic

[2] https://coronavirus.jhu.edu

[3] The usefulness of time-varying parameter models during times of policy changes was first noted by Robert Lucas in his original work on the Lucas Critique (Lucas (1976)).

has more leeway to go where the data tell it to and is not constrained by precise theoretical relationships that may be specified incorrectly.

Since deaths from COVID-19 fundamentally arise from infections, we model the number of deaths as depending on lagged cases. In particular, we use a spike-and-slab regression model (Mitchell and Beauchamp (1988); Ishwaran et al. (2005)), in which the number of deaths on a given day in a particular state depends on the lagged number of daily new cases in that state. Due to the long lag between the time COVID-19 patients test positive and the time they may die, our specification includes 35 lags, introducing a large number of coefficients relative to the length of the sample period. The spike-and-slab structure shrinks the regression coefficients in order to improve forecast performance.

We adapt our empirical specification to account for the fact that over time and across states, there has been heterogeneity in how the pandemic has evolved and how states have responded. First, we introduce endogenous time-varying parameters (TVP). The parameters for the model of infections depend on social distancing and testing, and the death rate depends on testing. We measure social distancing using geolocation data from 16 to 20 million mobile devices and measure the intensity of testing using the ratio of infections to tests conducted. The model thus captures how these predictors alter the predicted path of infections and deaths while providing additional flexibility to match different trajectories in a way that is tightly disciplined by data. Second, we utilize the panel structure of the data. In particular, our estimation allows the data to determine the correlation in parameters across states and imposes that social distancing and testing have the same effect on parameters in all states. Exploiting the panel structure sharpens estimates and forecasts in the presence of a short sample period by leveraging on data from all states to inform the estimates for a given individual state.

The model forecasts from May and June at horizons of up to 4 weeks are generally corroborated by the data. In particular, we check the empirical frequency at which the data realizations fall below various quantiles of our model forecasts. The forecasts match the empirical realizations for daily new cases except during the sharp rise at the end of June and start of July. However, by mid-July, the parameter estimates update and produce forecasts that largely match the data in the second half of July. The density forecasts for daily new deaths also match the empirical realizations well, especially at the upper quantiles. Our forecasts for new and cumulative deaths perform favorably relative to a repository of models from leading teams of epidemiologists and statisticians.

The endogenous TVP allow us to consider how social distancing and testing drive the model-implied paths for cases and deaths. Due to the nonlinearity of our model, we find that the effect of increased social distancing and testing on the predicted number of cases

differs across states even though the parameters determining this dependence are fixed across states. For instance, under the median estimates for Texas, quantitatively plausible increases in social distancing or testing are associated with a reduction in the number of cases by up to 50%. Under the median estimates for New York, the peak number of cases is early and sharp, and neither social distancing nor testing substantially changes the model-implied path for cases. We also find that more testing is associated with a lower death rate since the reported infections are likely to include more asymptomatic or mild cases.

Epidemiologists have long studied the spread of infectious diseases, using both increasingly complex theoretical models and also more purely empirical frameworks. We contribute to the latter by utilizing the toolkit prevalent in the analysis of economic data. In that respect, our work is similar to Harvey and Kattuman (2020), Li and Linton (2020), and Liu et al. (2020), who also use statistical models to forecast the pandemic. Our work is closest to Liu et al. (2020), who similarly use a panel structure. They use a linear time-trend model that allows for an exogenous break, whereas our model is a nonlinear autoregressive model whose parameters are connected to observable predictors. Our functional form shares similarities with the generalized logistic curve used by Harvey and Kattuman (2020) to model the number of cases. Both are flexible models for monotone progress from an initial condition toward a saturation point. In contrast, Li and Linton (2020) use a polynomial time trend for the logarithm of cases that is less flexible. Both Harvey and Kattuman (2020) and Li and Linton (2020) focus on locality-by-locality estimation.

Our paper also connects with recent work that enriches structural models from epidemiology, primarily the so-called Susceptible-Infected-Recovered (SIR) framework. The structural nature of the SIR model allows for the analysis of policy and counterfactual scenarios (e.g., Atkeson (2020); Fernández-Villaverde and Jones (2020); Hornstein (2020)). A hybrid approach is taken by Atkeson et al. (2020), who fit data on daily deaths to a mixture of Weibull functions, then use the model-implied mixture of Weibulls to obtain a time-varying reproduction rate for an SIR model. However, Koroloev (2020) and Kopecky and Zha (2020) highlight identification issues in SIR frameworks, which pose a challenge to accurate forecasting and quantification of uncertainty.

With the growing data on the COVID-19 pandemic, numerous attempts have been made to study the connection between different variables and the spread of the disease. One approach is to incorporate the SIR model into a choice-theoretic framework (e.g., Eichenbaum et al. (2020); Farboodi et al. (2020); Bognanni et al. (2020)) in order to model the feedback between individual or policy decisions and the transmission of the virus. Our reduced form approach seeks to minimize assumptions and gives the data a greater role in informing the researcher. A second approach is to estimate a SIR model with exogenous TVP (e.g., Ar-

royo Marioli et al. (2020); Buckman et al. (2020); Dandekar and Barbastathis (2020)), then check the correlation of the parameters with various observables, such as social distancing or quarantine measures, ex post. In contrast, we estimate the dependence of parameters on these predictors jointly with the rest of the model. Finally, numerous papers have used microeconometric methods that make use of differences across localities (e.g., Almagro and Orane-Hutchinson (2020); Desmet and Wacziarg (2020); Glaeser et al. (2020)). By incorporating the panel structure, we similarly utilize variation across both states and time to determine how social distancing and testing affect the path of the virus.

The paper is structured as follows. In Section 2, we introduce our model specification which we use to capture the evolution of infections and deaths over the course of an epidemic. We describe the data and estimation procedure in Section 3 and present the results in Section 4. Section 5 concludes.

2 A Panel Model for Estimating and Forecasting Pandemics

We now introduce and specify our empirical modeling framework for estimating and forecasting infections and deaths over the course of a pandemic. We formally introduce the model setup before highlighting the distinctive features of our specifications.

2.1 Model Setup

We begin by modeling the number of infections independently since it is the key variable in any theoretical or empirical model that studies the evolution of an epidemic. The number of subsequent deaths is a function of the number of infections, which we consequently model as a function of the lagged number of new cases.

2.1.1 Number of Cases

We specify the following model for the reported number of infections.[4] Given states $i = 1, ..., N$ and time periods $t = 1, ..., T$, denote the cumulative number of reported cases nor-

[4]We model the number of reported cases directly since it is the most common approach in the literature. Testing in the U.S. has increased substantially since the onset of the pandemic, which we capture by allowing the parameters to depend on the positive test rate. Therefore, the estimates and projections should not be interpreted as capturing the unobserved true number of cases in the population.

malized by population by $C_{i,t}$. We assume that $C_{i,t}$ follows:

$$\Delta \log C_{i,t} = \log(1+\gamma_{i,t}) \frac{\phi(C_{i,t-1}; \alpha_{i,t}, \zeta_{i,t}, \eta_{i,t})}{\phi(10^{-5}; \alpha_{i,t}, \zeta_{i,t}, \eta_{i,t})} \exp(u^C_{i,t}) \tag{1}$$

$$\phi(C; \alpha, \zeta, \eta) \equiv \exp[-C^{-\alpha} - (\zeta^\eta - C^\eta)^{-2}] \tag{2}$$

$$u^C_{i,t} = \rho_i u^C_{i,t-1} + \varepsilon^C_{i,t}\,, \tag{3}$$

where $\varepsilon^C_{i,t} \sim \mathcal{N}(0, (1-\rho_i^2)(\sigma^C_{i,t})^2)$. The normalization by $\phi(10^{-5}; \alpha_{i,t}, \zeta_{i,t}, \eta_{i,t})$ ensures that when a fraction 10^{-5} of the population has been infected, the growth rate in the absence of shocks or time-variation in parameters is γ. The AR(1) processes $u^C_{i,t}$ allow for potentially persistent deviations from the deterministic trend. We assume these shocks are stationary.

In what follows, we describe the role of each of the parameters in giving flexibility to the model-implied path for number of cases, before describing how the parameters vary across time and states.

Functional Form. A key feature of the model is the functional form for ϕ in equation (2) and the resulting range of trajectories implied by equation (1). We choose ϕ so that the model with fixed parameters follows the general pattern of infections in a pandemic, with an initial sharp increase as the disease spreads, followed by a leveling off and decline due to public health policies or herd immunity. On the other hand, we ensure that the functional form is flexible enough to match a wide range of such paths.

Figure 1 plots the trajectories for a range of time-invariant parameter values, illustrating the role of each parameter. Each path begins with the same initial condition, and each panel shows the effect of changing one parameter while leaving the rest unchanged. The functional form allows for different rates of increase and subsequent decrease in the number of new cases, different peaks, and different asymptotic numbers of cumulative cases. Importantly, there is no fixed relationship between the different stages of the pandemic, imposing neither the tight structure of an SIR model nor the symmetry of functional forms such as quadratic trends.

Identification of the model parameters is based on the growth rate and changes in the growth rate of infections, with the different parameters associated with distinct phases of the epidemic. Initially, the rate of growth is approximately exponential. The effect of increasing α, shown in the top-right panel of Figure 1, is to increase the curvature of the number of new cases, ΔC_t, in the early phase of the epidemic, which captures the appearance of large clusters or the effects of social distancing measures. As the stock of susceptible hosts starts getting smaller, the rise in the growth rate decelerates until it reaches a peak. Afterwards,

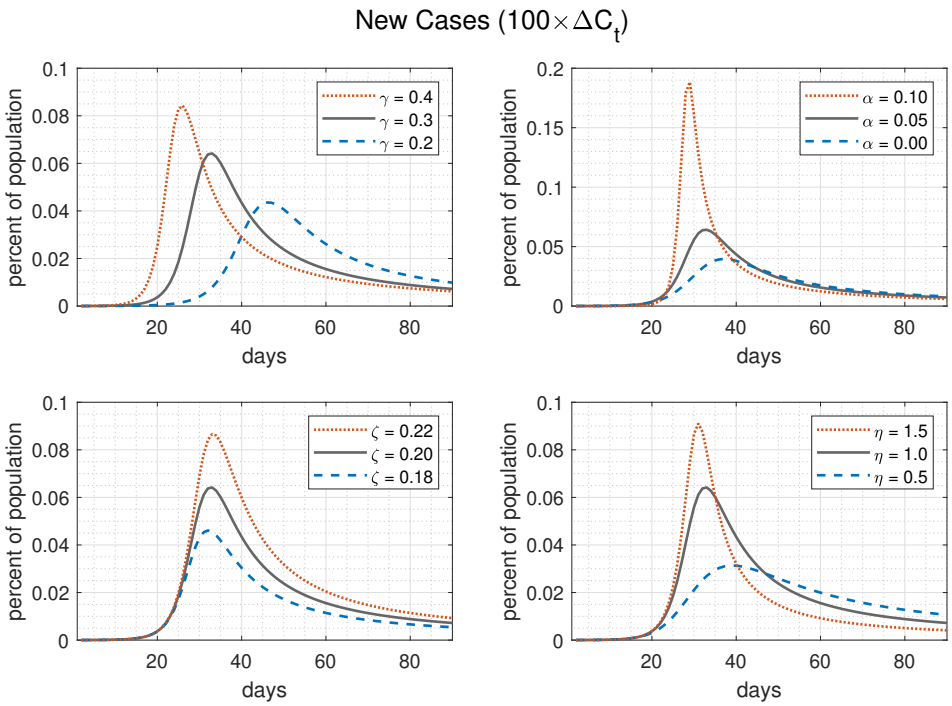

Figure 1: Model-implied daily new cases with time-invariant parameters and no shocks. Gray lines are identical across all panels. Each panel shows change in model-implied number of new cases associated with a change in one parameter.

the growth rate of new infections declines.

The parameters η and ζ determine the long-run number of cumulative cases and the speed at which a population converges to that number, which could depend on factors such as demographics or policies.[5] In particular, the bottom panels of Figure 1 show that increasing ζ or η leaves the initial path of C_t unchanged but increases the number of new cases around the peak. While ζ does not affect the overall shape of the trajectory materially, decreasing η flattens the peak and leads to a slower decline in the number of cases.

Panel Structure and Time-Varying Parameters. Denote a generic parameter by $\theta \in \{\gamma, \alpha, \zeta, \eta, \sigma^C, \rho\}$. We assume that the parameters depend on a vector of observables $X_{i,t}$, which could include demographic variables, social distancing metrics, or the amount of testing:

$$g(\theta_{i,t}) = g(\overline{\theta}_i) + \kappa_\theta' X_{i,t} \tag{4}$$

$$g(\overline{\theta}_i) \sim \mathcal{N}(\mu_\theta, \omega_\theta^2), \tag{5}$$

[5]We also estimated a model that replaces the exponent of 2 on the second term in (2) with a freely estimated parameter. We fix the exponent because it is not well-identified separately from ζ.

where the function g is chosen to map the appropriate support of θ to the real line:

$$g(\theta) = \begin{cases} \theta & \text{supp}(\theta) = (-\infty, \infty) \\ \log(\theta) & \text{supp}(\theta) = (0, \infty) \\ \log(\frac{1}{1-\theta} - 1) & \text{supp}(\theta) = (0, 1) \end{cases} .$$

The model (4) assumes that time variation in the parameters within a state can arise only through time-varying predictors but does not allow for exogenous time variation in parameters. We allow for differences across states through the fixed effect $\overline{\theta}_i$. The joint distribution of parameters across states is determined by the hyperparameters μ_θ and ω_θ, which we estimate.

2.1.2 Number of Deaths

In addition to modeling infections, we also consider the mortality rate. Not all infections are fatal, and an observed death is the outcome of a process that can vary over time. We thus assume that the number of deaths on any given day depends on the lagged number of cases, but allow the data to determine the rate at which infections translate to deaths at different horizons and which lags are most important.

In particular, we consider an extension of the spike-and-slab regression for the number of new deaths $\Delta D_{i,t}$ as a function of lagged new cases $\Delta C_{i,t-\ell}$:

$$\Delta D_{i,t} = \frac{1}{\sum_\ell \iota_{i,\ell}} \sum_{\ell=1}^{L} \iota_{i,\ell} \lambda(\overline{\lambda}_{i,\ell}, \delta; X_{i,t-\ell}) \Delta C_{i,t-\ell} + \varepsilon_{i,t}^D \tag{6}$$

$$\varepsilon_{i,t}^D \sim \mathcal{N}(0, \frac{1}{\sum_\ell \iota_{i,\ell}} \sum_{\ell=1}^{L} \iota_{i,\ell} \Delta C_{i,t-\ell} \times (\sigma_i^D)^2), \tag{7}$$

where $\iota_{i,\ell} \sim \text{Bernoulli}(p_\ell)$ is a variable selection indicator. In the absence of shocks, the setup nests a deterministic SIR model, in which infections lead to deaths at a Poisson rate, and the values of λ will fall geometrically with ℓ at the recovery rate. We provide greater flexibility by allowing the coefficient λ to vary freely across lags. In addition, we include shocks whose variance scales with the number of lagged cases. The scaling captures the trade-off between a lower variance due to a larger number of cases over which to average and a higher variance due to a larger number of expected deaths.

The variable selection parameter $\iota_{i,\ell}$ shrinks small coefficients to zero, which can improve forecast precision since there are a large number of coefficients relative to observations. On one hand, the parameter L is relatively large because COVID-19 patients who do not survive

the illness have a relatively long lag time between testing positive for the virus and dying. On the other hand, COVID-19 is a recent disease for which we have a relatively short panel of data. By making p_ℓ depend on ℓ, the model assumes that lags, which are more important for predicting mortality in one state, are likely important for other states as well.

The death rate λ roughly captures the fraction of infected individuals who die after a given number of days.[6] It depends on a state- and window-specific parameter $\overline{\lambda}_{i,\ell}$ and a coefficient δ that determines the dependence of death rate on the predictors $X_{i,t-\ell}$. For instance, the death rate likely decreases with the extensiveness of testing, as more mild and asymptomatic cases are documented. Here we consider the functional form:

$$\lambda(\overline{\lambda}_{i,\ell}, \delta; X_{i,t-\ell}) = \overline{\lambda}_{i,\ell}(1 + \delta' X_{i,t-\ell}). \tag{8}$$

To allow the death rates to be correlated across states, we specify:

$$(\sigma_i^D)^{-2} \sim \Gamma(a_\sigma, b_\sigma) \tag{9}$$

$$\overline{\lambda}_{i,\ell} \mid \iota_{i,\ell} = 1 \sim \mathcal{N}(\mu_\lambda, (\sigma_i^D)^2/\upsilon), \tag{10}$$

where $(a_\sigma, b_\sigma, \mu_\lambda, \upsilon)$ are hyperparameters to be estimated.

2.2 Discussion of Model Features

Endogenous Time Variation in Parameters. A key feature of our model is the endogenous time variation in parameters. In most models, any time variation in parameters is exogenous. In contrast, we assume that the model-implied path of the pandemic can only change if the observables $X_{i,t}$ fluctuate. While the time variation in parameters offers the flexibility to track a wide range of trajectories for infections and deaths, the endogeneity of the time variation adds discipline to these fluctuations.[7] By restricting the parameters to only vary with observable data, we also rely more on the functional form in (2) to fit the data and produce accurate forecasts.

In addition, we are able to estimate how different observables change the path for infections and deaths. This allows us to compute counterfactual trajectories for the pandemic that condition on different paths for $X_{i,t}$. A more common approach in the literaure to estimate the effect of observables has been to estimate a TVP model with exogenous time variation, then to assess the correlation of the smoothed parameters with observables as a

[6] This is exactly true if the number of new cases is independent across days and $\iota_{i,\ell} = 1$ for all ℓ.

[7] This is similar in spirit to the literature on endogenous Markov regime-switching, for instance, Diebold and Lee (1994); Chang et al. (2017). However, in our model, the observables drive the actual parameter values rather than the probability of moving between regimes.

second step (e.g. Arroyo Marioli et al. (2020); Buckman et al. (2020); Dandekar and Barbastathis (2020)). Our approach estimates the effect of the observables jointly with the rest of the model, allowing for coherent quantification for both point estimates and posterior uncertainty.

Panel Structure. Rather than estimate the model state-by-state, we consider a panel model in which the parameters are correlated across states. This is designed to tighten estimates for states that are in the early stages of transmission, since their state-specific parameter estimates are informed by the data for states that are further along in the pandemic. The panel structure also aids in the estimation of κ_θ and δ. Since these parameters are common across states, our panel estimates leverage the state-level heterogeneity in $X_{i,t}$, yielding tighter estimates of the effect of these predictors.

Statistical Model. Our models for infections and deaths are both statistical, unlike a majority of models that are variants of the SIR model (see, for instance, Table 1 in the Appendix for the list that we compare our forecasts against). Our model's relative flexibility allows us to fit the data well despite the restrictions we place on the time variation in parameters. Nevertheless, the minimal structure that the model imposes on the rise and fall in the number of cases helps generate tighter long-run forecasts.

3 Data and Estimation

3.1 Data

We use publicly available data on the daily number of reported COVID-19 cases and deaths in the 50 U.S. states and Washington, D.C. from The New York Times[8] from January 21, 2020 through August 11, 2020. For each state, we start the sample when the state has a cumulative number of cases of at least 20. The data set collects the cumulative number of infections at the end of each day reported by local government and health authorities.

We also use two predictors for the variation in the parameters: the Mobility and Engagement Index (MEI), constructed by the Federal Reserve Bank of Dallas from January 3, 2020 through August 8, 2020, and positive test rates from The Atlantic's Covid Tracking Project from March 1, 2020 through August 8, 2020. We allow the parameters of the model of infections to depend on both the MEI and testing, and allow the parameters of the model

[8]For full details, refer to the associated GitHub repository.

for deaths to depend on positive test rates only. See Figure 9 in the Appendix for the time paths of these predictors.

The MEI summarizes the deviation from normal mobility behavior since the start of the COVID-19 outbreak. The index is formed using principal components on seven variables measured using geolocation data from 16 to 20 million mobile devices. Each variable is a measure of how much individuals travel away from home, and the index is normalized so that a higher value corresponds to greater mobility (i.e., less social distancing).[9] We take a seven-day lagged moving average to smooth out seasonal fluctuations. While all states show a common pattern of declining mobility in March followed by an increase in mobility from the second half of April, there is heterogeneity across states in how much and how quickly mobility changed at different points of the pandemic cycle.

We define the positive test rate as the total number of reported cases over the past seven days divided by the total number of tests conducted over the past seven days. A lower positive test rate is an indication of more extensive testing. As reporting errors occasionally lead to a positive test rate that is negative or greater than one, we truncate the positive test rate to be within the $[0, 1]$ interval. While the positive test rate for the U.S. declined in aggregate as states increased their testing capacities in March and April, the path for the positive test rates has differed greatly across states.

In what follows, we estimate the model using data through August 8, 2020, when our samples for the MEI and testing data end. We also estimate the model using data through every other Sunday from May 3, 2020 through June 14, 2020, and check the performance of our forecasts at a horizon of 1 to 28 days. This covers the period during which states were reopening and until the point when many states experienced a second wave of sharp increases in case numbers. Finally, we also take forecasts using data through July 15, 2020, in order to show how the estimates update around the peak of the second wave.

3.2 Estimation

We draw from the posterior of the model for the number of infections using the following Gibbs sampler:

1. Condition on $\overline{\theta}_{1:N}$.

   (a) Draw $(\mu_\theta, \omega_\theta^2)$ from a normal-inverse-gamma distribution.

   (b) Draw κ_θ using Metropolis-Hastings.

[9] See Atkinson et al. (2020) for details on the construction and comparison with other measures of social distancing.

2. Draw $\overline{\theta}_{1:N} \mid \kappa_\theta$ using Metropolis-Hastings.

Step 1(a) is standard and uses the property that the normal-inverse-gamma distribution is a conjugate prior (see, for instance, Zellner (1971)). Step 1(b) requires computing the likelihood contribution from equation (1) for the entire panel. Step 2 can be done state-by-state, similar to the estimation of the baseline model without time-varying parameters. Hence, Step 2 could also be parallelized if a researcher wanted to use our model on a larger set of locations.

To draw from the posterior of the model for mortality, we make use of the spike-and-slab structure. In particular, we take the following steps:

1. Conditional on $\mu_\lambda, \upsilon, a_\sigma, b_\sigma, p_\ell$,
   (a) Conditional on δ,
      i. Draw $\iota_{i,\ell}$ state-by-state using Metropolis-Hastings.
      ii. Draw $\overline{\lambda}_{i,\ell}, \sigma_i^D \mid \iota_{i,\ell}$ from a normal-inverse-gamma distribution.
   (b) Draw $\delta \mid \iota_{i,\ell}, \overline{\lambda}_{i,\ell}, \sigma_i^D$ using Metropolis-Hastings.
2. Draw $\mu_\lambda, \upsilon \mid \overline{\lambda}_{i,\ell}, \sigma_i^D$ from a normal-inverse-gamma distribution.
3. Draw $a_\sigma, b_\sigma \mid \sigma_i^D$ using Metropolis-Hastings.
4. Draw $p_\ell \mid \iota_{i,\ell}$ using the conjugate form of the beta prior.

Step 1(a)(i) uses the fact that a normal-inverse-gamma distribution is a conjugate prior for a linear regression. In particular, for a given i, we can compute the marginal likelihood for equation (6) given a candidate draw $\{\iota_{i,\ell}\}_{\ell=1}^L$, integrating out $\overline{\lambda}_{i,\ell}$ and σ_i^D. Given $\iota_{i,\ell}$, we have a standard regression for Step 1(a)(ii). Conditional on $(\iota_{i,\ell}, \overline{\lambda}_{i,\ell}, \sigma_i^D)$, we can draw δ using Metropolis-Hastings, since the likelihood is straightforward to compute. In Step 2, it is straightforward to draw (μ_λ, υ) from a normal-inverse-gamma distribution, since $\overline{\lambda}_{i,\ell}$ is distributed according to a generalized least squares regression on a constant, in which the standard deviations of the shocks are known to be σ_i^D. In Step 4, we utilize the fact that the beta distribution is the conjugate prior for a binomial distribution. We pick $L = 35$, allowing the number of deaths to depend on the number of new cases over a month ago. The spike-and-slab structure allows the data to determine which lags are most important.

3.3 Prior

For the model for number of cases, we consider a relatively uninformative normal-inverse-gamma conjugate prior for the hyperparameters $(\mu_\theta, \omega_\theta^2)$ for $\theta \in \{\gamma, \alpha, \zeta, \eta, \sigma^C, \rho\}$:

$$\begin{aligned} \omega_\theta^{-2} &\sim \Gamma(1, 0.25) \\ \mu_\theta \mid \omega_\theta &\sim \mathcal{N}(0, \omega_\theta^2). \end{aligned}$$

In addition, we impose a Gaussian prior for κ_θ for $\theta \in \{\gamma, \alpha, \zeta, \eta, \sigma^C, \rho\}$:

$$\kappa_\theta \sim \mathcal{N}(0, 0.5^2 V^{-1}),$$

where V is a diagonal matrix with the sample variances of each corresponding predictor $X_{i,t}$. The prior thus represents the belief that each predictor contributes equally to the variance of the transformed parameters $g(\theta_{i,t})$.

For the model of mortality, we similarly impose a Gaussian prior for δ:

$$\delta \sim \mathcal{N}(0, 0.5^2 V^{-1})$$

to match the prior on κ_θ. For the variance σ_i^D of the shocks, we use the prior:

$$\begin{aligned} a_\sigma &\sim \Gamma(2, 1) \\ b_\sigma &\sim \Gamma(2, 3 \times 10^{-7}), \end{aligned}$$

which is calibrated to the scale of the number of deaths. In particular, scale parameter for b_σ of 3×10^{-7} is chosen so that the mode of the prior is approximately the average state-specific variance of the number of new deaths divided by the square root of the number of new cases, $\frac{1}{N} \sum_i \widehat{V}_i \left[\Delta D_{i,t} / \sqrt{\Delta C_{i,t}} \right]$. The shape parameters of 2 for a_σ and b_σ are chosen to make the prior relatively uninformative. For the distribution of $\overline{\lambda}_{i,\ell}$, we use the prior

$$\begin{aligned} \upsilon &\sim \Gamma(1, 10^{-3}) \\ \mu_\lambda \mid \upsilon &\sim \mathcal{N}(0, (0.05^2 \times 10^{-3})/\upsilon). \end{aligned}$$

The prior for ν is chosen to be relatively flat and is scaled such that the standard deviation of a state-lag-specific coefficient $\overline{\lambda}_{i,\ell}$ is of order 10^{-2}. The conditional variance of μ_λ is scaled

by 10^{-3} to account for the scale of ν. Finally, we consider the prior

$$p_\ell \sim \text{Uniform}(0, 1)$$

for the probability of including a lag.

3.4 Forecasting

To forecast the number of cases and deaths in each state, we need to condition on a path for the time-varying predictors. To that end, we estimate an AR(1) model independently for each predictor in each state using the last 14 days of data. If the absolute value of the predictor is declining, we assume a long-run mean of zero. If the absolute value of the predictor is increasing, we assume the long-run mean is the maximum value of the absolute value of that predictor in the full sample. We extrapolate from the last data point using this AR(1) model without shocks. The path we condition on captures the general trend of the predictor in the most recent data.

4 Results

We present two sets of empirical results. We first discuss the parameter estimates of the panel, whereby we give an overview of the results from the 50 U.S. states and D.C. We then show how these parameter estimates depend on measures of social distancing and testing. In the next step, we check the forecast performance of the model, whereby we focus on three large states that have exhibited different patterns for the evolution of the pandemic.

4.1 Parameter Estimates

We first discuss the parameter estimates based on data through August 8, 2020.

4.1.1 State-Specific Parameters

Figure 2 shows the marginal posterior distributions from the infections model for both the state-specific components of γ, α, ζ, η, σ^C, and ρ, as defined in equation (4), and the aggregate distribution from equation (5). A large amount of heterogeneity across states is required to match the wide range of trajectories across states even after accounting for the MEI and positive test rates. Nevertheless, the posteriors for the aggregate distributions of γ, α, η, and σ^C are substantially tighter than their priors, and the data are informative as indicated by the shifts of the posteriors.

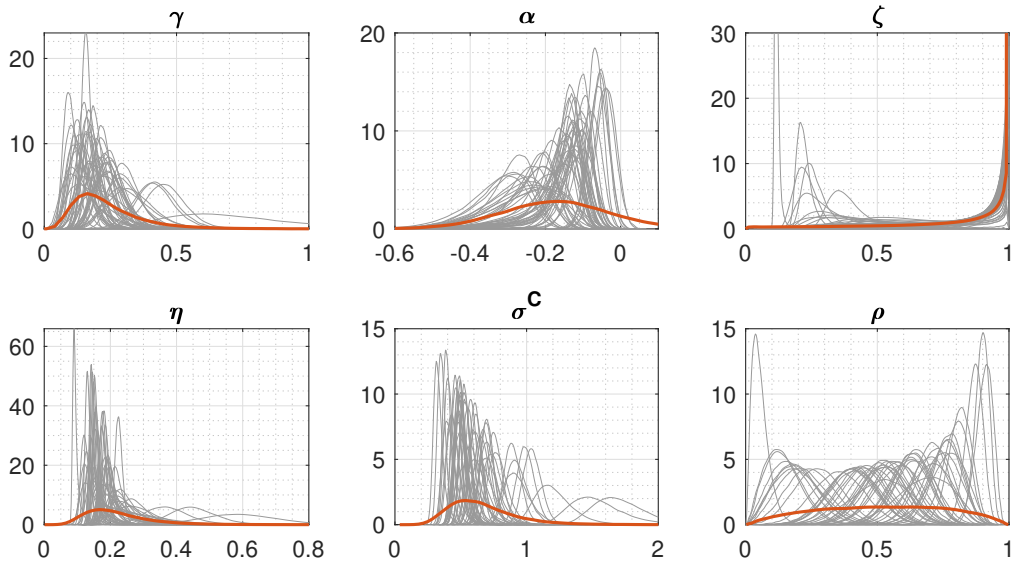

Figure 2: Marginal posteriors for γ, α, ζ, η, σ^C, and ρ. **Thin gray lines:** posteriors for each state; **Thick red line:** aggregate distribution across states.

For the mortality model, we define $\lambda_{i,\ell} \equiv \lambda(\overline{\lambda}_{i,\ell}, \delta; \frac{1}{T}\sum_{t=1}^{T} X_{i,t-\ell})$ and plot the posterior means of $\iota_{i,\ell}$ and $\lambda_{i,\ell} \mid \iota_{i,\ell} = 1$ in Figure 3. The former is the probability of including a lag, while the latter captures the average death rate in a state for a given lag. We also plot the mean of these parameters across states on the same axes. Both parameters show a clear weekly seasonal component, potentially reflecting measurement error due to different rates of processing test results or documenting deaths over the week. Nevertheless, the degree of seasonality differs greatly across states.

The posterior estimates for the mortality model also indicate that the number of deaths on a given day depends on the number of new cases up to five weeks prior. While there is a decreasing trend in the estimates for $\iota_{i,\ell}$ as ℓ increases, the data favors including cases at long lags to predict future deaths. For instance, the mean estimate for $\ell = 35$ is 0.07, which is roughly half the mean estimate for $\ell = 1$. The estimates for $\lambda_{i,\ell} \mid \iota_{i,\ell} = 1$ show a small upward trend. In terms of magnitude, the average posterior estimates of $\lambda_{i,\ell} \mid \iota_{i,\ell} = 1$ across states lie between 0.01 and 0.04 across lags, corresponding to the typical range of death rates reported for the U.S.. These estimates reflect the relatively long lag time between infection and death.

Notably, the lag time between infection and death stands in contrast with the assumption of Poisson death and recovery rates in standard SIR models. This assumption is generally made for modeling convenience. However, our coefficient estimates do not appear to be generated from a Poisson structure. Specifically, the fatality and recovery rates used in the

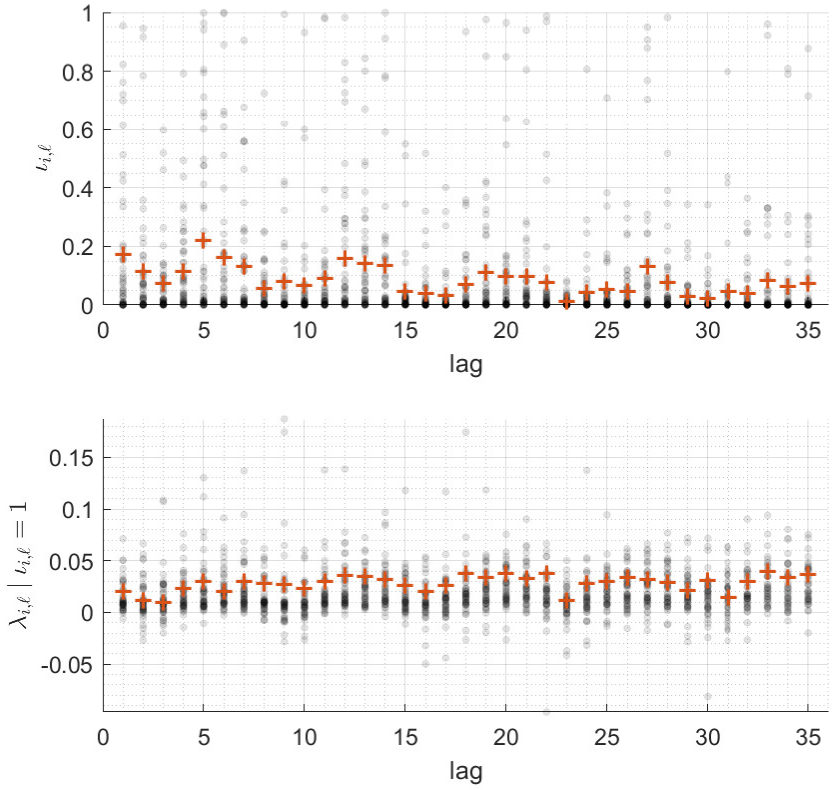

Figure 3: Posterior mean estimates for $\iota_{i,\ell}$ and $\lambda_{i,\ell} \mid \iota_{i,\ell} = 1$. Gray dots correspond to posterior means for individual states. Red crosses indicate average across states.

recent COVID literature range between 0.2% - 1.4% and 1/4 - 1/14, respectively.[10] An OLS regression, in which the numbers of cases are independent across lags, would likely show the regression coefficients decaying rapidly. Intuitively, the spike-and-slab regression likely inherits a similar structure both for $\iota_{i,\ell}$ and $\lambda_{i,\ell} \mid \iota_{i,\ell} = 1$. The long lag between infection and death is further evidenced from the recent second wave of cases in the U.S., as the rise in cases was not followed by a corresponding increase in the number of deaths until several weeks later.

4.1.2 Dependence of Parameters on Social Distancing and Testing

Our estimates show that differences in the MEI and positive test rates are associated with significant variation in the model parameters. In particular, Figure 4 shows that the parameters in both the models for cases and deaths are significantly connected to the MEI and the positive test rate through (4) and (8). These correlations are statistical and do not

[10] Atkeson et al. (2020) provides an overview and use baseline fatality and recovery rates of 0.5% and 1/5, respectively.

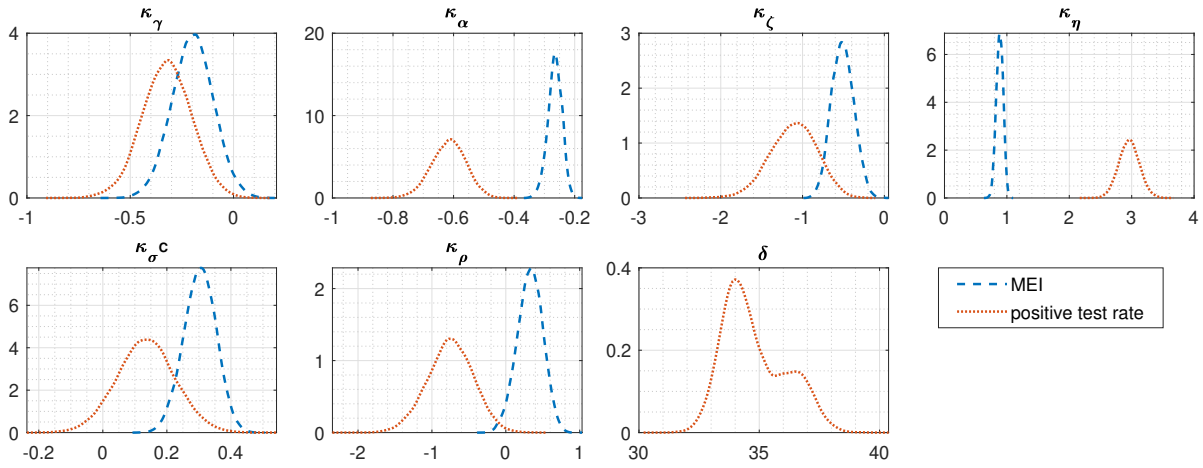

Figure 4: Marginal posteriors for dependence κ_θ and δ of parameters on predictors. **Blue dashed line:** social distancing; **Red dotted line:** amount of testing.

identify causality. In general, one would expect a greater level of social distancing when cases increase (e.g. Glaeser et al. (2020)) due to an endogenous response from both households and governments. On the other hand, a higher number of cases mechanically increases the positive test rate if the number of tests remains constant.

In order to give a sense of how the parameter estimates in Figure 4 for κ_γ, κ_α, κ_ζ, and κ_η map into the behavior of the nonlinear model, we compare the model-implied path of new cases under baseline paths for the MEI and positive test rate against alternative paths with more social distancing or testing in Figure 5. The respective paths of the MEI follow the typical path in the data: it decreases in the first 60 days, then increases and levels off below the initial level of zero. For testing, we consider constant positive test rates 0.1 and 0.2. To show how social distancing and testing can affect the model-implied paths differently across states, we consider the model-implied paths under the median parameter estimates using New York and Texas as examples. In particular, we initialize the number of cases at $10^{-4}\%$ of the population, then simulate the model forward without shocks.

Social distancing and testing can be associated with a lower number of infections, but this relationship depends on the underlying trajectory of cases. Under the Texas parameter estimates, both a lower MEI and a lower positive test rate are associated with flatter curves. In contrast, under the New York parameter estimates, the model-implied trajectories remain relative unchanged for different MEI and positive test rates. These differences arise because of the different trajectories that New York and Texas faced: New York had a relatively rapid rise and fall in the number of cases, whereas in Texas infections increased only gradually at first. Under the baseline paths for the MEI and positive test rate, the number of new cases

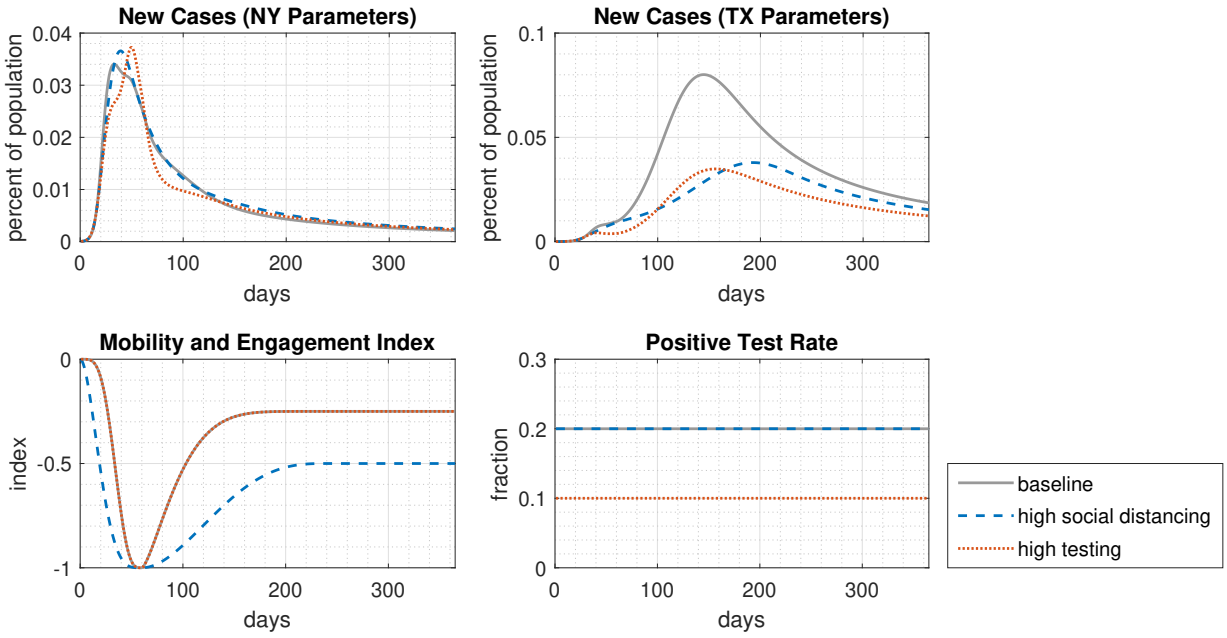

Figure 5: Model-implied paths of cases for different levels of social distancing and testing, using posterior median for $(\gamma, \alpha, \zeta, \eta)$ for New York and Texas and posterior median for κ_θ. **Gray solid line:** baseline; **Blue dashed line:** increased social distancing (lower MEI); **Red dotted line:** increased testing (higher positive test rate).

under the New York parameters falls to around 70% of its peak level by the 60-day mark, while the number of new cases under the Texas parameters continues to rise. Our result that the underlying trajectory of cases matters is consistent with Atkeson et al. (2020), who use an estimated SIR model to show that the effects of distancing measures depend on the precise scenario considered.

The MEI and the positive test rate are also associated with variation in the variance and persistence of the shock $u_{i,t}^C$ in (1). Lower levels of social distancing are associated with shocks of higher variance and higher persistence, while lower levels of testing coincide with shocks of higher variance and lower persistence. Reduced social distancing likely leads to more clusters developing and could cause any temporary spikes in cases to last longer. On the other hand, lower levels of testing can also result in more measurement error.

The estimate of δ, the dependence of the death rate on the positive test rate, shows that a higher positive test rate is associated with a substantially higher death rate. At the posterior mean for the average death rate μ_λ of 0.006, an increase in the positive test rate of 10 percentage points corresponds to an increase in death rate of 2 percentage points. Intuitively, a higher positive test rate occurs when individuals who are tested tend to have

a higher ex-ante probability of being infected. These individuals tend to have more severe symptoms, leading to a higher reported death rate.

4.2 Forecasts

We now assess the forecast performance of our model. This is a critical aspect of our analysis since the global pandemic is still ongoing with no end in sight. Moreover, the course of the pandemic in the U.S. has proven to be very volatile and heterogeneous across the states as reported above. The most recent aggregate U.S. data even show the appearance of a second peak in infections. The ability to forecast well in this changing environment is a key aspect for an empirical epidemiological model that our panel framework with endogenous TVP is designed to accomplish.

To check the forecast performance of our model, we estimate the model using an initial subsample of the data and compare the model forecasts to the actual realizations. We do this every two weeks from May 3, 2020 to June 14, 2020, covering the period during which the aggregate number of cases in the U.S. was declining after the initial peak until the sharp spike in cases leading to the second peak. During this time, states reopened at different rates, and the number of cases and deaths across states followed a wide range of paths. The heterogeneity across states provides a test for whether our model is sufficiently flexible to match the numerous possible paths for the pandemic.

4.2.1 Coverage

Figure 6 shows Q-Q plots to compare the empirical realizations to the quantiles for our forecasts of new cases and new deaths at the 1- to 28-day horizon. In particular, for each horizon, we check the fraction of states whose realized number of new cases or deaths fall below the qth quantile of our forecast for that state. We also average over each week to remove any weekly seasonality, by counting the fraction of state-horizon observations that fall below the respective qth quantile of our forecasts.

Overall, our forecasts match the realized data well. For new cases, the posterior quantiles of our forecasts match the empirical frequency closely, except for the forecasts from June 14, 2020. This coincides with the sharp spike in cases in numerous states. In many cases, our model predicts a rise in cases, but one that is smaller than the eventual spike. One likely reason for the underprediction is that we tie the time-variation in parameters solely to the time-variation in the MEI and positive test rates. The forecasts may be improved by the inclusion of more variables, including more detailed mobility measures or disaggregated data. Nevertheless, later forecasts from July 15, 2020 indicate that the parameter estimates

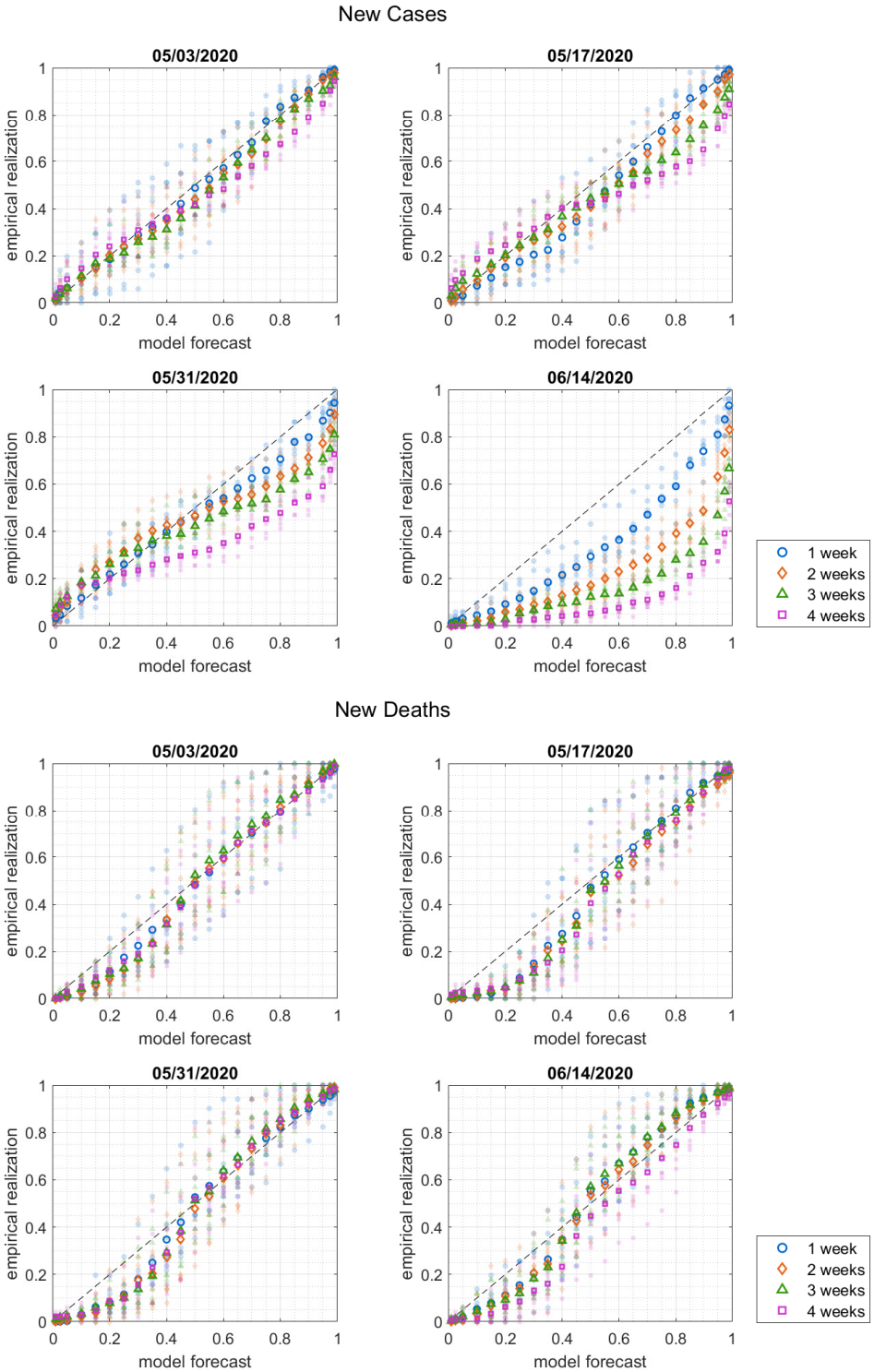

Figure 6: Q-Q plots for forecasts of increase in number of cases and deaths one to 28 days ahead, comparison with other models. **Translucent markers:** fraction of states whose realized number of new cases or deaths falls below given quantile of forecast for a specific horizon; **Opaque outline markers:** average over each week. Marker colors and shapes indicate week of forecast.

were updated in response to the new inflow of data. Figures 10 shows that the 95% error bands of the corresponding forecast largely contain the realized data in late July and early August.

The forecasts for new deaths match the empirical frequencies well at the upper quantiles, but tend to undershoot slightly at the lower quantiles. The undershooting arises largely among states that have many days without deaths. This occurs more regularly in the early part of the sample and in states with a low number of cases. Indeed, by June 14, 2020, the forecasts undershoot less as the zeros no longer bias the forecasts downward as much.

Figure 7 compares the forecast performance of our model to the forecasts compiled by the COVID-19 Forecast Hub, which collects forecasts from leading teams of epidemiologists and statisticians that are curated to ensure overall accuracy. As the repository primarily provides forecasts for the number of deaths, we focus our assessment on cumulative and new deaths. Since many models only update their forecasts once a week on Sunday or Monday, we compare our forecasts taken on Sunday to any forecasts in the COVID-19 Forecast Hub from the same day or one day later. We plot these competing forecasts on the same Q-Q plot as our forecasts for the one- to four-week horizon.

Our model performs relatively well compared with alternative models at the horizons and dates considered.[11] This is notable given that the competing models include both statistical and richly specified theoretical SIR models, many estimated using more data than we have used. Table 1 in the Appendix provides further details on these models.

4.2.2 New York, California, and Texas

For further insight into how the model forecasts adapt to the data, we plot expanding window forecasts for New York, California, and Texas in Figure 8. We focus on these three states not only because they have among the largest populations in the U.S., but also because the epidemic progressed differently in each, thereby providing a template for assessing the forecasts in the other states. We plot forecasts from May 17, June 14, and July 15, 2020. These correspond roughly to the decline in cases after the initial wave, the increase in cases moving into the second wave, and the peak of the second wave. While all three states have been severely affected by the COVID-19 pandemic, they have displayed different paths for the number of cases, number of deaths, social distancing, and testing. The model forecasts reflect these differences.

[11] Since we check the forecasts against the New York Times data, the relative performance may be attributed partly to differences in the data used by different models. However, to fully explain the wide dispersion in performance across models, one would require implausibly large and systematic differences across data sets.

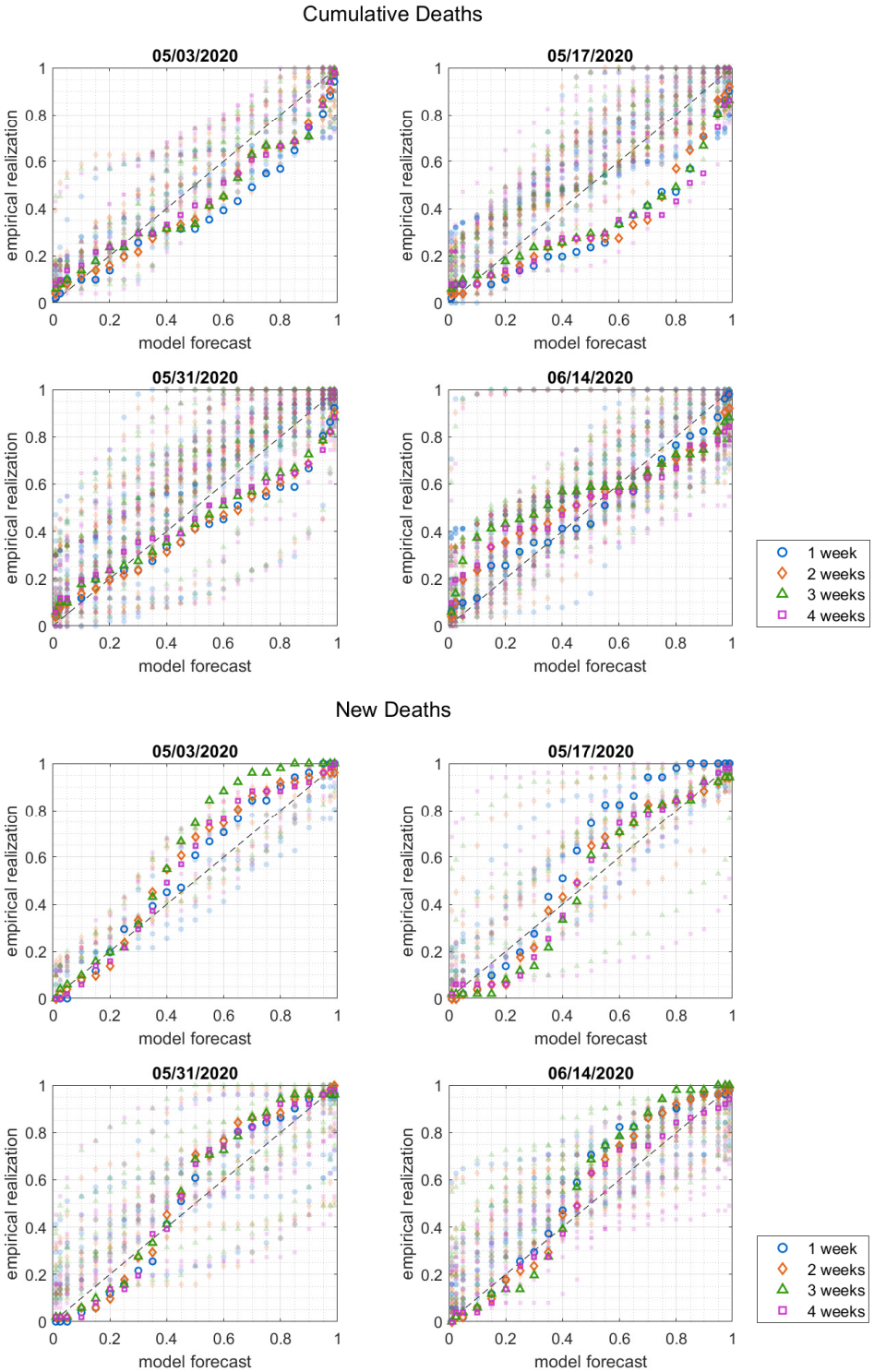

Figure 7: Q-Q plots for forecasts of cumulative and new deaths one to four weeks ahead, comparison with other models. **Opaque outline markers:** fraction of states whose realized number of new or cumulative deaths falls below given quantile of forecast from this paper at given horizons; **Translucent markers:** forecasts from other models on same date or one day later for same horizon. Marker colors and shapes indicate horizon.

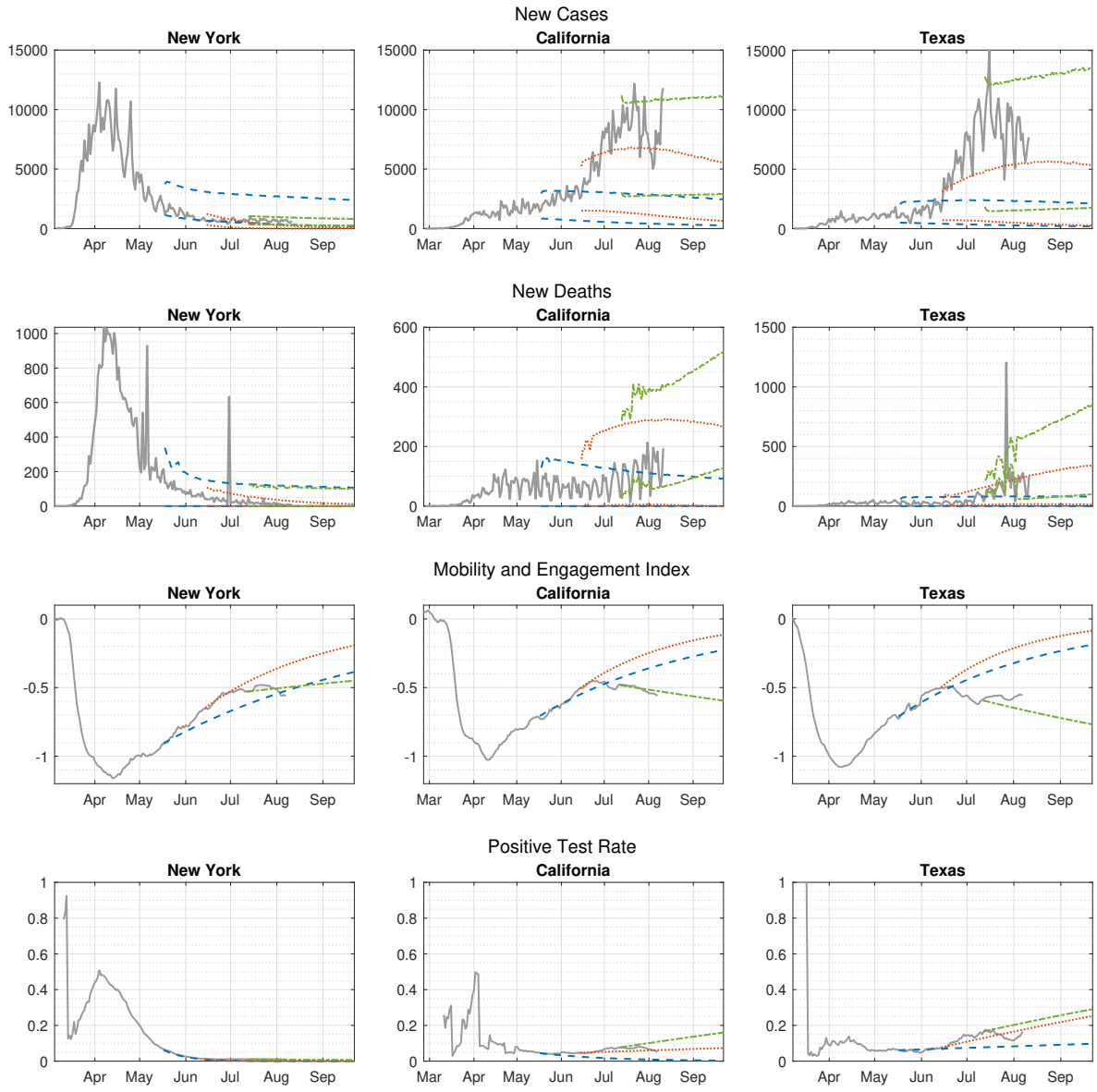

Figure 8: Forecasts for daily new cases, new deaths, MEI, and positive test rates in New York, California, and Texas, using data through May 17, June 14, and July 15, 2020. **Top two panels:** 95% error bands; **Bottom two panels:** extrapolation based on AR(1) model for last 14 days.

New York. While New York was one of the hardest hit states during the early part of the pandemic, the number of cases has steadily decreased since the first half of April. By May, the number of cases was significantly lower than its peak. This is the typical path of cases predicted by standard SIR models, and our model is able to fit this path well. In particular, the model forecasts match the realized gradual decline in number of cases and deaths, with the realized data falling within or close to the 95% error bands, which are relatively tight.

California. The number of cases in California plateaued in April, began to increase in May, and accelerated in June before beginning to stabilize at the end of July. The model forecasts qualitatively match these patterns. In May, we forecast a relatively stable number of cases up till September at least. The data fall mostly within the 95% error bands until mid-June when the increase in number of cases begins to accelerate. When we forecast the number of cases from mid-June, the model predicts a possible rise in cases for several weeks, albeit a smaller one than what actually occurs in the second half of June. Finally, even though the number of cases continued to rise through the first half of July, the model forecasts a plateau in the number of cases between mid-July and early September. The data in the second half of July corroborate this forecast, as the number of cases has fluctuated around the upper half of the 95% error bands.

The forecasts for the number of deaths mirrors those for the number of cases, except for the mid-July forecast. In particular, the model predicts an increase in the number of daily new deaths. This prediction is borne out by the data, with the number of deaths rising in late July even as the number of new cases began to decline. This reflects the result that the number of new deaths depends on the number of new cases up to five weeks prior.

Texas. The data and forecasts for Texas are qualitatively similar to California. In mid-July, the model forecasts slightly faster growth in the number of cases in Texas than in California. However, there is substantial uncertainty about the rate of this increase, as the error bands are about twice as wide than those of the June forecast. The realized data in the second half of July show a slight decrease that is comfortably within the error bands. While the patterns of cases in California and Texas were relatively similar, the positive test rate during the first half of July increased more rapidly in Texas than in California. The forecasts for the two states thus condition on different projected paths for testing, leading to the contrasting predicted trends in new cases.

5 Conclusion

We develop and estimate a statistical model of the COVID-19 pandemic that has three key features. First, parameters are allowed to vary over time, but only in line with observable variables. Second, the model has a panel structure that sharpens estimates and forecasts. Third, the underlying functional forms for the model are flexible and able to track the typical paths of cases and deaths in a pandemic. The model's forecasts perform favorably relative to alternative epidemiological and statistical models.

By allowing parameters to depend on social distancing and testing, our estimates highlight the interaction between these predictors and underlying state-specific parameters in generating model predictions. Specifically, while both increased social distancing and more intensive testing can be associated with lower case numbers, this does not occur when the peak in a locality is relatively early and followed immediately by a sharp decline. In addition, we estimate a decline in death rates associated with a lower positive test rate, accounting for the different composition of cases reported as testing becomes more widely available.

Our functional form captures the trajectory of cases as well as connection between infections and deaths that motivate SIR models. However, our statistical approach minimizes modeling assumptions relative to the structural SIR literature, providing estimates that can help inform the calibration or specification of these models. The autoregressive structure of our setup is akin to time series econometric models and is conducive to forecasting. At the same time, we introduce a panel structure to leverage the variation across states and time that microeconometric methods often rely on. The Bayesian estimation transparently quantifies parameter and forecast uncertainty. Our framework thus bridges a range of approaches to provide insights into the evolution of this global pandemic.

References

Almagro, Milena and Angelo Orane-Hutchinson (2020), "The Determinants of the Differential Exposure to COVID-19 in New York City and Their Evolution over Time." *Covid Economics: Vetted and Real-Time Papers*.

Arroyo Marioli, Francisco, Francisco Bullano, Simas Kučinskas, and Carlos Rondón-Moreno (2020), "Tracking R of COVID-19: A New Real-Time Estimation Using the Kalman Filter." Working paper.

Atkeson, Andrew (2020), "What Will Be the Economic Impact of COVID-19 in the U.S.? Rough Estimates of Disease Scenarios." NBER Working Paper 26867, National Bureau of Economic Research.

Atkeson, Andrew, Karen A. Kopecky, and Tao Zha (2020), "Estimating and Forecasting Disease Scenarios for COVID-19 with an SIR Model." NBER Working Paper 27335, National Bureau of Economic Research.

Atkinson, Tyler, Jim Dolmas, Christoffer Koch, Evan Koenig, Karel Mertens, Anthony Murphy, and Kei-Mu Yi (2020), "Mobility and Engagement Following the SARS-Cov-2 Outbreak." Federal Reserve Bank of Dallas Working Paper 2014.

Bognanni, Mark, Doug Hanley, Daniel Kolliner, and Kurt Mitman (2020), "Economic Activity and COVID-19 Transmission: Evidence from an Estimated Economic-Epidemiological Model." Working paper.

Buckman, Shelby R., Reuven Glick, Kevin J Lansing, Nicolas Petrosky-Nadeau, and Lily M Seitelman (2020), "Replicating and Projecting the Path of COVID-19 with a Model-Implied Reproduction Number." Federal Reserve Bank of San Francisco Working Paper 2020-24.

Chang, Yoosoon, Yongok Choi, and Joon Y Park (2017), "A New Approach to Model Regime Switching." *Journal of Econometrics*, 196, 127–143.

Dandekar, Raj and George Barbastathis (2020), "Quantifying the Effect of Quarantine Control in COVID-19 Infectious Spread Using Machine Learning." Working paper.

Desmet, Klaus and Romain Wacziarg (2020), "Understanding Spatial Variation in COVID-19 across the United States." NBER Working Paper 27329, National Bureau of Economic Research.

Diebold, Francis X. and Joon-Haeng Lee (1994), "Regime Switching with Time-Varying Transition Probabilities." In *Non-Stationary Time Series Analysis and Cointegration* (Colin Hargreaves, ed.), 283–302, Oxford University Press.

Eichenbaum, Martin S., Sergio Rebelo, and Mathias Trabandt (2020), "The Macroeconomics of Epidemics." NBER Working Paper 26882, National Bureau of Economic Research.

Farboodi, Maryam, Gregor Jarosch, and Robert Shimer (2020), "Internal and External Effects of Social Distancing in a Pandemic." NBER Working Paper 27059, National Bureau of Economic Research.

Fernández-Villaverde, Jesús and Charles I. Jones (2020), "Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities." NBER Working Paper 27128, National Bureau of Economic Research.

Glaeser, Edward L., Caitlin S. Gorback, and Stephen J. Redding (2020), "How Much Does COVID-19 Increase with Mobility? Evidence from New York and Four Other U.S. Cities." NBER Working Paper 27519, National Bureau of Economic Research.

Harvey, Andrew and Paul Kattuman (2020), "Time Series Models Based on Growth Curves with Applications to Forecasting Coronavirus." *Covid Economics, Vetted and Real-Time Papers*.

Hornstein, Andreas (2020), "Social Distancing, Quarantine, Contact Tracing, and Testing: Implications of an Augmented SEIR Model." Federal Reserve Bank of Richmond Working Paper 20-04.

Ishwaran, Hemant, J. Sunil Rao, et al. (2005), "Spike and Slab Variable Selection: Frequentist and Bayesian Strategies." *Annals of Statistics*, 33, 730–773.

Kopecky, Karen A. and Tao Zha (2020), "Impacts of COVID-19: Mitigation Efforts Versus Herd Immunity." Policy Hub 03-2020, Federal Reserve Bank of Atlanta.

Koroloev, Ivan (2020), "Identification and Estimation of the SEIRD Epidemic Model for COVID-19." Working paper, Binghamton.

Li, Shaoran and Oliver Linton (2020), "When Will the COVID-19 Pandemic Peak?" Cambridge Working Papers in Economics 2025.

Liu, Laura, Hyungsik Roger Moon, and Frank Schorfheide (2020), "Panel Forecasts of Country-Level COVID-19 Infections." NBER Working Paper 27248, National Bureau of Economic Research.

Lucas, Robert Jr (1976), "Econometric Policy Evaluation: A Critique." *Carnegie-Rochester Conference Series on Public Policy*, 1, 19–46.

Mitchell, Toby J. and John J. Beauchamp (1988), "Bayesian Variable Selection in Linear Regression." *Journal of the American Statistical Association*, 83, 1023–1032.

Zellner, Arnold (1971), *An Introduction to Bayesian Inference in Econometrics*. Wiley.

A Additional Figures and Tables

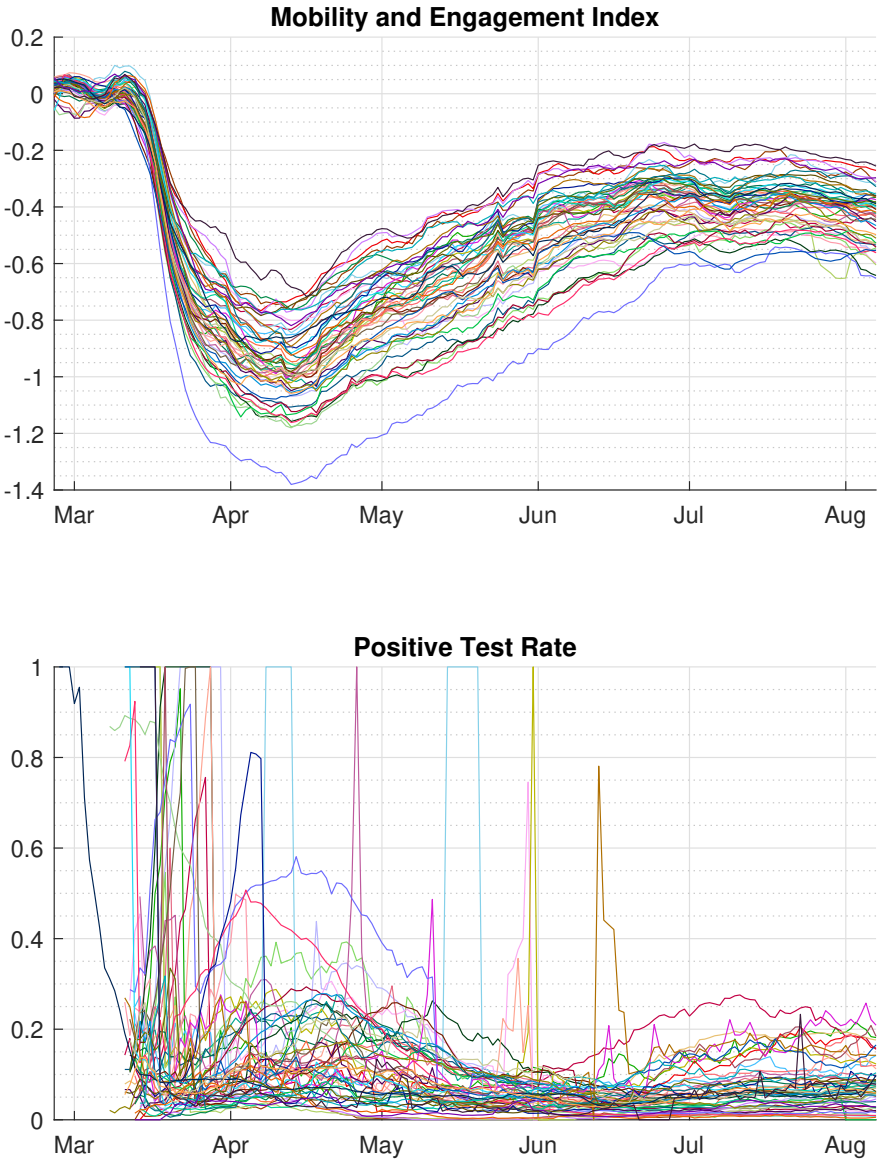

Figure 9: MEI and positive test rate for all U.S. states.

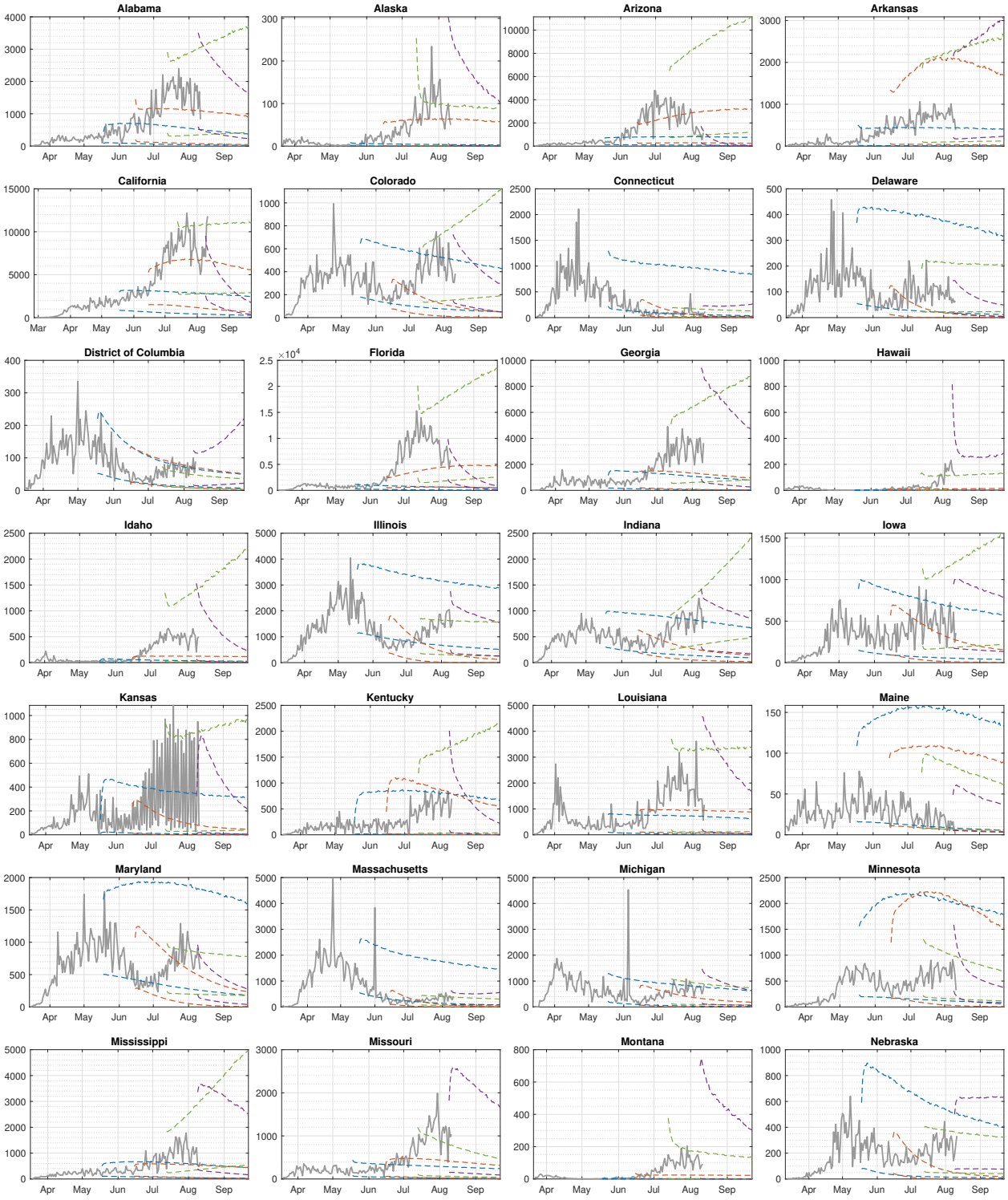

Figure 10a: Forecasts (95% error bands) for daily new cases in U.S. states, using data through May 17, June 14, July 15, and August 8, 2020.

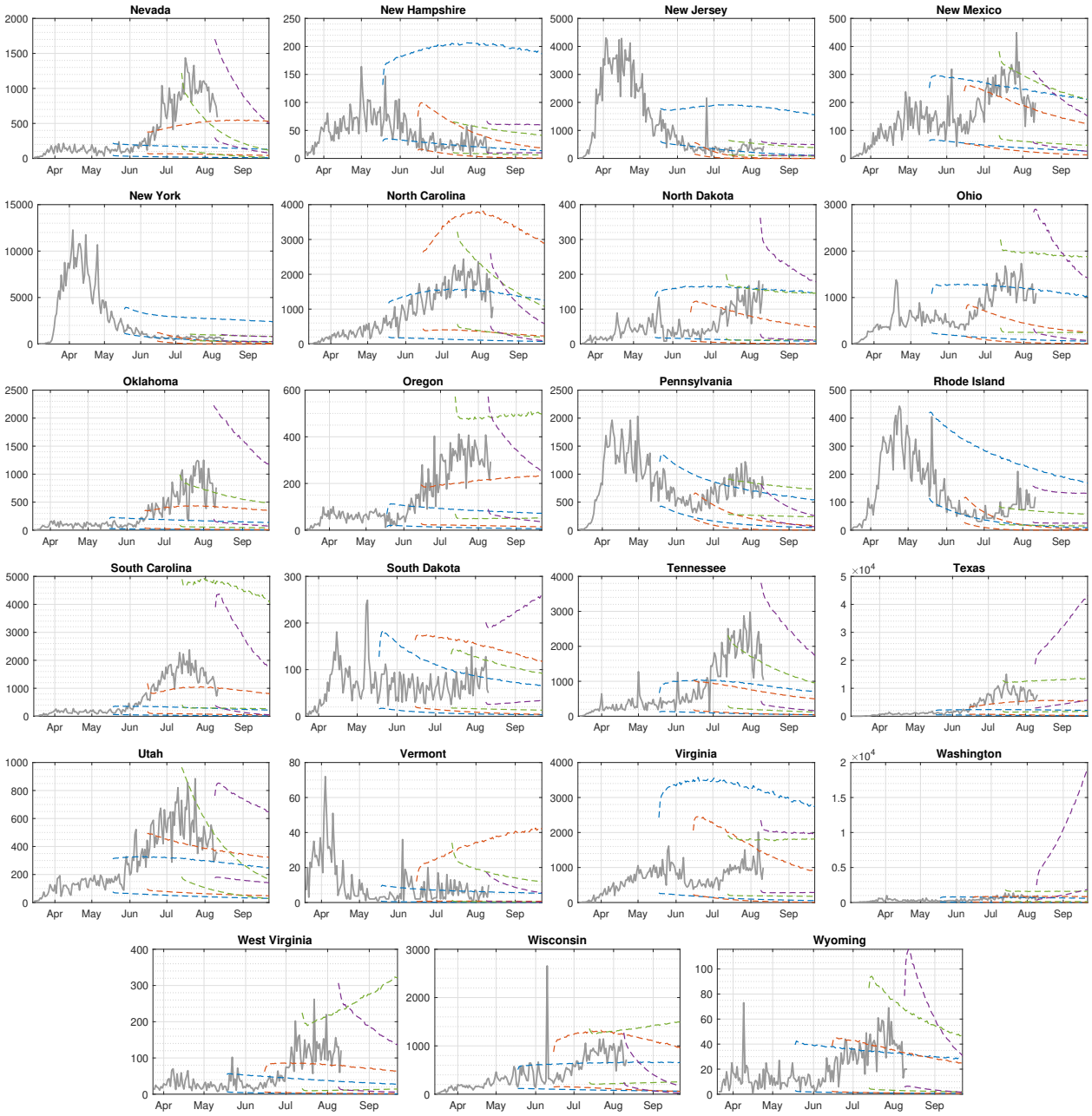

Figure 10b: Forecasts (95% error bands) for daily new cases in U.S. states, using data through May 17, June 14, July 15, and August 8, 2020.

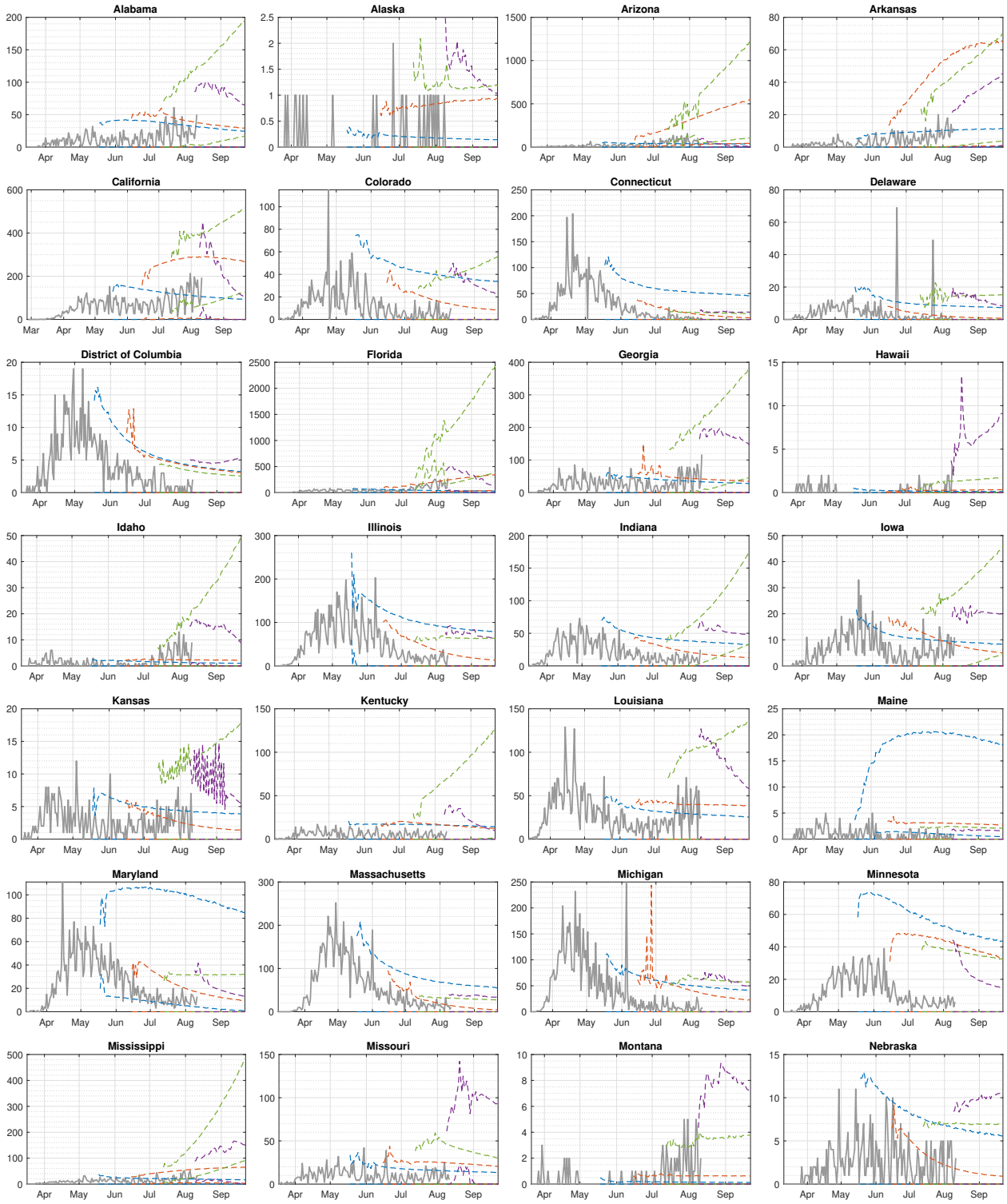

Figure 11a: Forecasts (95% error bands) for daily new deaths in U.S. states, using data through May 17, June 14, July 15, and August 8, 2020.

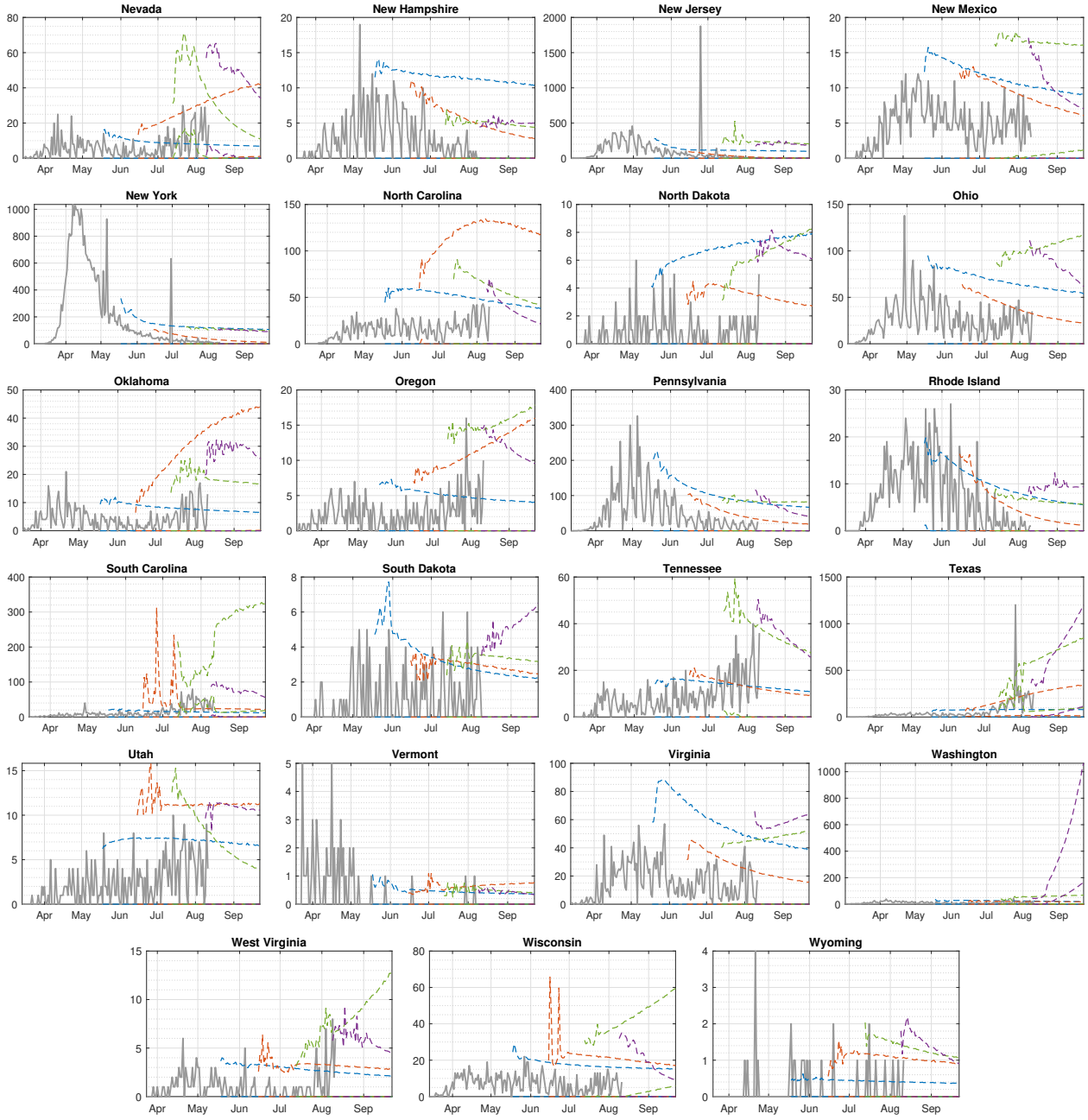

Figure 11b: Forecasts (95% error bands) for daily new deaths in U.S. states, using data through May 17, June 14, July 15, and August 8, 2020.

Team/Model	Description	Cumulative Deaths				New Deaths			
		05/03	**05/17**	**05/31**	**06/14**	**05/03**	**05/17**	**05/31**	**06/14**
Auquan	SEIR with non-linear mixed effects curve-fitting		M	S	M				
Covid Act Now	TVP-SEIR			S	S			S	S
COVID-19 Forecast Hub									
• Baseline	Extrapolation from most recent observations	M	M	M	M	M	M	M	M
• Ensemble	Combined forecast from selected models	M	M	M	M				M
Columbia University	County-level metapopulation SEIR								
	• Constant contact rate				S				
	• High transmission rate	S	S	S	S				
	• Low transmission rate		S	S	S				
	• Moderate transmission rate		S	S	S				
	• Most plausible	S	S	S	S				
IHME	TVP-SEIR with parameters linked to drivers		M		M		M		M
Johns Hopkins ID Dynamics	County-level metapopulation SEIR		S	S	S		S	S	S
Los Alamos National Laboratory	Statistical dynamical growth model	S	M	M	S				S
MGH/HMS COVID-19 Simulator	TVP-SEIR			S	S			S	S
MIT Covid Analytics DELPHI	TVP-SEIR with nonlinear infection rate	M	M	M	M				
Northeastern MOBS Lab GLEAM	Metapopulation, age structured SLIR model	M	M	M	M	M	M	M	M
Predictive Science Inc DRAFT	SEIRX with age distribution, disease severity		M	M	M				
University of Arizona EpiGro	SIR with data assimilation			S	S				
UCLA Stat. Machine Learning Lab	Machine learning SEIR with unreported cases		S	S	S				S
UMass-Amherst MechBayes	Bayesian TVP-SEIR	S	S	S	S		S	S	S
U.S. Army ERDC	Bayesian TVP-SEIR			M	M			M	M
UT-Austin COVID-19 Consortium	Bayesian multilevel negative binomial regression	M	M	M	M	M	M	M	M
Youyang Gu ParamSearch	Machine learning SEIR with reopening	S	S	S	S	S	S	S	S

Table 1: Summary of models from Figure 7. 'S' and 'M' indicate that forecast is from Sunday and Monday, respectively.