Behavior and the Transmission of COVID-19

By Andrew G. Atkeson, Karen Kopecky, and Tao Zha *

Since the outbreak of the COVID-19 pandemic in early 2020, epidemiologists and economists have raced to develop models of the disease usable for forecasting the progression of the pandemic, for evaluating the effectiveness of various interventions aimed at mitigating the spread of the disease, and for understanding the interaction of the pandemic with the economy.The models developed by economists often differ from those developed by epidemiologists in that economists include equations intended to capture the impact of endogenous changes in human behavior undertaken in response to the pandemic on the progression of the pandemic itself.[1] We refer to such models as *behavioral* SIR, or BSIR, models.

We show that a simple BSIR model can match an important feature of the data: the growth rates of daily deaths began at high and highly dispersed levels early on in the pandemic and then fell toward zero fairly rapidly. However, despite this remarkable match between model and data on growth rates of daily deaths, our simple model cannot match many aspects of the evolution of the pandemic in many countries. Large unexplained factors or wedges are required to account for regime shifts in the evolution of daily deaths, disease transmission, and human behavior that occurred in later days of the pandemic. Future research should aim at accounting for these wedges so as to develop new models that are more useful for analysis of this and future pandemics.

I. A Behavioral SIR Model

Our behavioral SIR model, which is commonly used in the literature, is built on the standard SIR epidemiological model that is used to interpret the data on deaths from COVID-19 for a given region.[2] At each moment of time, the population N is divided into four categories (states) that sum to the total population: susceptible S, infected I, resistant R, and dead D. The *transmission rate* $\beta(t)$ is the rate at which infected agents spread the virus to others that they encounter at date t. We use $\mathcal{R}(t)$ to denote the *effective reproduction number* of the disease at date t, the ratio of the rate at which infected agents infect susceptible agents to the recovery rate γ of infected agents. The effective reproduction number can fall either due to changes in the normalized transmission rate, $\beta(t)/\gamma$, or changes in S/N.

By inverting the standard SIR model to interpret data on deaths, one can derive a simple linear relationship between the growth rate of daily deaths and the effec-

* Atkeson: Department of Economics, University of California, Los Angeles, NBER, and Federal Reserve Bank of Minneapolis, andy@atkeson.net. Kopecky: Federal Reserve Bank of Atlanta and Emory University, karen.kopecky@atl.frb.org. Zha: Federal Reserve Bank of Atlanta, Emory University, and NBER, zmail@tzha.net. We are grateful to Hongyi Fu for superlative research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the Federal Reserve Banks of Atlanta and Minneapolis, the Federal Reserve System, or the National Bureau of Economic Research.

[1] See Atkeson, Kopecky and Zha (2020) for the extensive references to the economics literature.

[2] Details of this SIR model, as well as the BSIR model, are presented in Atkeson, Kopecky and Zha (2020) and Droste and Stock (2021).

tive reproduction number

$$\mathcal{R}(t) = 1 + \frac{1}{\gamma}\frac{\frac{d^2D(t)}{dt^2}}{\frac{dD(t)}{dt}}, \tag{1}$$

Our behavioral SIR model assumes that the transmission rate $\beta(t)$ is a function of human activity $Y(t)$

$$\beta(t) = \bar{\beta}Y(t)^{\alpha}\exp(\psi_{\beta}(t)). \tag{2}$$

The parameter $\bar{\beta}$ in equation (2) is a fixed coefficient that captures features of a region's population and environment determined prior to the pandemic that might impact transmission. Factors considered in the literature include population density, modes of transportation, household and demographic structure, cultural norms (bowing, shaking hands or kissing), and temperature and humidity. The parameter α captures the elasticity of transmission with respect to activity.

The parameter $\psi_{\beta}(t)$ represents a potentially time-varying wedge shifting the region-specific relationship between activity and transmission. This wedge may represent the impact of policies such as mask-wearing, ventilation, physical distancing, redesign of workspaces, or other measures implemented after the start of the pandemic that reduce transmission given a fixed level of activity, and/or natural variation in the transmission of the virus over time such as variation driven by seasonality or mutation of the virus.

The behavioral component of the model assumes that individuals' decisions to engage in activity in a given region at time t, $Y(t)$, are a declining function of the time derivative of cumulative deaths, $\dot{D}(t)$, which is measured by the current level of *daily deaths*. We specify this function describing behavior as

$$Y(t) = \exp\left(-\kappa\dot{D}(t) + \psi_y(t)\right), \tag{3}$$

where $\kappa > 0$ represents the semi-elasticity of activity $Y(t)$ with respect to daily deaths. The variable $\psi_y(t)$ represents a time-varying shifter to the region-specific relationship between deaths and activity that may be induced by lockdowns or changes in behavior in response to the disease.

By substituting equation (3) into equation (2), we obtain a reduced-form relationship between the current level of daily deaths and the transmission rate, given by

$$\beta(t) = \bar{\beta}\exp(-\alpha\kappa\dot{D}(t) + \psi(t)), \tag{4}$$

where $\psi(t) \equiv \alpha\psi_y(t) + \psi_{\beta}(t)$ is the composite wedge.

II. Estimation

In Atkeson, Kopecky and Zha (2020), we discuss how to recover consistent estimates of the growth rate of daily deaths, the level of daily deaths, and the level of cumulative deaths from noisy reported data on daily deaths.[3] With the estimated death growth and level, one can study the epidemiological dynamics implied by our BSIR model. For this exercise we also need to set values of several model parameters. Specifically, we set $\nu = 0.005$, a level used by the CDC for those aged 50-69.[4] The level of the fatality rate, if held constant, does not impact our estimates of the evolution of the effective reproduction number in equation (1).

The rest of the parameters are set or estimated as follows. We set $\gamma = 0.2$. We set

[3] Atkeson, Kopecky and Zha (2020) study 103 regions (34 states and 69 countries) with the estimation period starting at the location-specific date when cumulative deaths reached 25 and ending on November 12, 2020.

[4] See Table 1 at `https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html`.

$\alpha = 2$ for all locations to capture the idea that the number of interactions between individuals goes up with the square of the activity level of all individuals. We allow the semi-elasticity κ to vary by location, depending both on individuals' opportunities to reduce activity by working from home and on their beliefs about personal tradeoffs involved in exposing themselves to virus transmission. We normalize $\psi_\beta(0) = 0$ and the level of activity at the beginning of the pandemic to $Y(0) = 1$. Given these normalizations, the region-specific parameter $\bar{\beta}$ determines the transmission rate of the virus in a given region at the beginning of the pandemic, with the basic reproduction number of the virus in this region given by $\mathcal{R}_0 = \bar{\beta}/\gamma$.

To solve the BSIR model, we set the initial fractions of susceptible, infected, recovered, and dead at date $t_0 > 0$ to their values based on estimated daily deaths in the data and the SIR model as described in Section I. For each value of κ, $\bar{\beta}$ is set such that the transmission rate at date t_0 in the BSIR model corresponds to the rate implied by the SIR model and the death data. This is achieved by setting $\bar{\beta} = \beta^{data}(t_0)\exp(\alpha\kappa\dot{D}^{data}(t))$. After setting the initial phase, for each location, we choose κ to minimize the distance between daily deaths in the data as given by our Bayesian estimation procedure and daily deaths generated by the BSIR model with no wedges such that $\sum_{t=t_0}^{T}\left[\dot{D}^{data}(t) - \dot{D}(t)\right]^2$ is minimized, where T indicates the terminal date.

III. Key Findings and Challenges

Our estimation yields important findings about the COVID-19 pandemic that pose challenges to future empirical work.

Finding I One key finding, documented in Atkeson, Kopecky and Zha (2020), is that the growth rate of daily deaths from COVID-19 fell rapidly almost everywhere within the first 30 days after each region reached 25 cumulative deaths. The solid black line in Figure 1 shows that the median estimated growth rate of daily deaths fell rapidly from an initial level of about 12 percent to zero within the first 30 days of the estimation period. During this initial phase, there was a wide dispersion of growth rates of daily deaths observed in most of 103 locations; then there was a regime shift and the dispersion remained in a relatively narrow range around zero

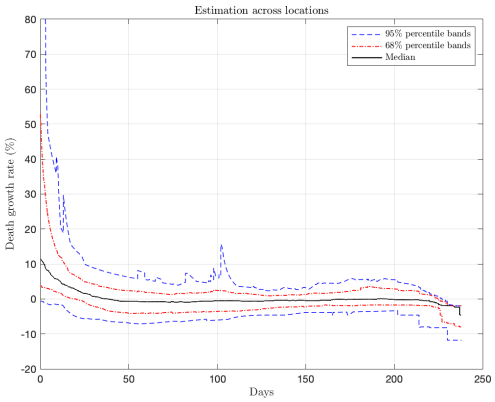

Figure 1. : Location and sampling uncertainty.

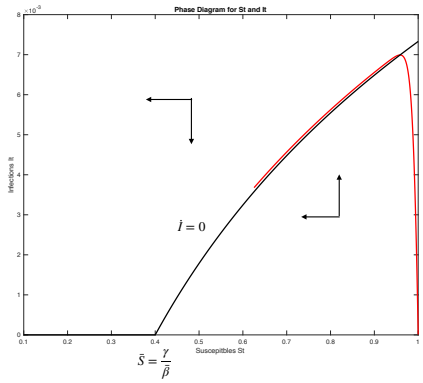

Figure 2. : Phase diagram for the dynamics implied by the standard BSIR model.

into mid November of 2020.[5] This fact was true even with substantial second and third waves of daily deaths observed in many locations in the fall and winter of 2020.

According to equation (1), the effective reproduction numbers, as well as our model-implied disease transmission rates, have a similar pattern to Figure 1. That is, they fell sharply in the early phase of the pandemic from widely dispersed initial levels and have since remained close to 1.[6] A challenge for empirical work is how to account for this stark switch in regime from the early period to the remaining period of the pandemic.

Finding II The key dynamics of the BSIR model can be summarized by the phase diagram presented in Figure 2 with $S(t)$ on the x-axis and $I(t)$ on the y-axis. The black curve is the locus of points (S, I) such that $\dot{I} = 0$. When (S, I) lies above the black curve, $\dot{I} < 0$ and $\dot{S} < 0$. When (S, I) lies below the black curve, $\dot{I} > 0$ and $\dot{S} < 0$. The red curve in Figure 2 shows the model-implied path of $(S(t), I(t))$. The pair $(S(t), I(t))$ starts in the lower left corner of the figure with $S(0)$ very close to one and $I(0)$ positive but very close to zero. $I(t)$ rises rapidly initially and crosses over the black locus of points such that $\dot{I} = 0$ and then falls slowly, remaining above that locus until $S(t)$ falls below $\bar{S}$ where the locus of points such that $\dot{I} = 0$ intersects the x-axis and $I(t)$ asymptotes to zero.

These dynamics imply two results: (a) the path of $I(t)$ and hence the path of daily deaths $\dot{D}(t)$ are both single-peaked; (b) after the peak of infections and daily deaths, these outcomes cannot fall rapidly to a low level without a substantial decline of $S(t)$. Thus, with no wedges, the model cannot account for multiple waves of infections and daily deaths or the patterns of daily deaths seen in many of the regions that were initially hard hit followed by a steep decline in daily deaths to very low levels (Finding I).

Finding III The distribution of the growth rates of daily deaths for all the regions we study, generated by our estimated BSIR model without wedges and reported in the top panel of Figure 3, is qualitatively similar to the data (Figure 1) in that the growth rates of daily deaths fall toward zero (the effective reproduction numbers fell toward one) over the course of 30 days or less—a regime shift. But the dispersion in the growth rates of daily deaths predicted by the BSIR model after the initial period is substantially smaller than the dispersion in the growth rates of these deaths observed in the data. Unlike the data, in the BSIR model without wedges, there are no realizations of growth rates substantially above or below zero. As a result, both the magnitude and the dispersion of the wedges grow larger over time (bottom panel of Figure 3).

The main reason for growing wedges in both magnitude and dispersion is that the relationship between the level of daily deaths and the logarithm of the transmission rate in the standard BSIR model (equation (4)) exists only for the initial period of the pandemic. After the initial period, however, there is little apparent relationship. California (Figure 4), for example, illustrates a one-to-one relationship between $\beta(t)$ and daily deaths in the initial period (the circled blue line) but no one-to-one relationship after this initial period

[5] The dashed lines in Figure 1 represent the 68% and 95% posterior probability intervals, which include both location uncertainty and sampling uncertainty. Most of the cross-sectional dispersion in growth rates in the figure, however, is driven by location uncertainty as sampling uncertainty within a location is very small.

[6] This result is robust to various extensions of the standard SIR model and consistent with the findings by the IHME at the University of Washington.

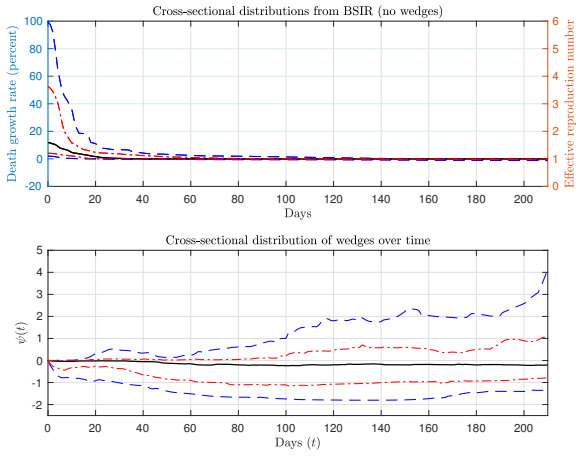

Figure 3. : Top panel: outcomes from the BSIR model without wedges. Bottom panel: wedges.

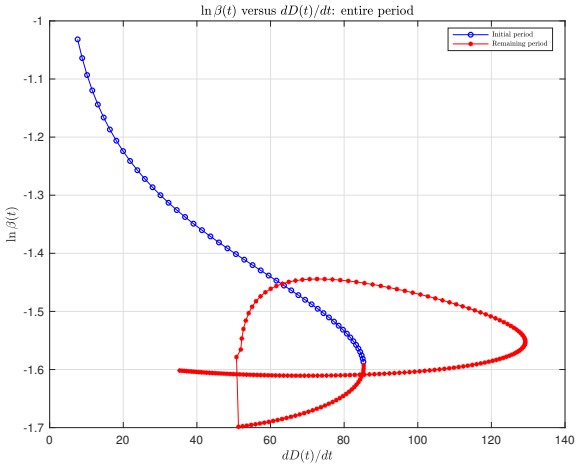

Figure 4. : Relationship between the model-implied $\ln\beta(t)$ and the data on daily deaths $\dot{D}(t)$ for California.

(the starred red line).[7] This pattern, which is suggestive of a regime shift in behavior over time in response to the deadly pandemic, is common across all the countries and U.S. states we examine. Accounting for such a regime shift in behavior in an economic model is a necessary but challenging task.

[7] Measuring $Y(t)$ with Google mobility data, we find a remarkably similar pattern: the one-to-one relationship between $\beta(t)$ and $Y(t)$ (equation (2)) or between $Y(t)$ and $\dot{D}(t)$ (equation (3)) breaks down after the initial phase of the pandemic.

IV. Going Forward

Qualitatively, the push by economists to introduce theories of behavior into epidemiological models of COVID-19 has been a big empirical success. Even a simple BSIR model matches the main features of the dynamics of the growth rate of deaths observed in many locations around the world.

As our findings demonstrate, however, much of the dynamics of the growth and level of deaths is left unexplained by the standard BSIR model. Such a model must be augmented with very large wedges to the transmission rate, while holding disease prevalence constant, to match the data on deaths. The next immediate task is to identify and estimate possible variations of behavioral parameters such as α, ν, γ, and κ across regions and over time, and to study whether these variations can account for part of the widening wedges we find. A more challenging task for future research is to assess how much of the regional and time variations in behavioral parameters is due to nonpharmaceutical policies versus changes in voluntary behavior (e.g., wearing masks and practicing social distancing) as well as versus changes in the disease itself.

REFERENCES

Atkeson, Andrew, Karen Kopecky, and Tao Zha. 2020. "Behavior and the Transmission of COVID-19." Unpublished Manuscript.

Droste, Michael, and James H. Stock. 2021. "Adapting to the COVID-19 Pandemic." American Economic Review (2021, this issue).