

Original Paper

Dynamic Panel Estimate–Based Health Surveillance of SARS-CoV-2 Infection Rates to Inform Public Health Policy: Model Development and Validation

James Francis Oehmke¹, BSc, MSc, PhD; Theresa B Oehmke², BSc, MSc; Lauren Nadya Singh¹, BSc, MPH; Lori Ann Post¹, BA, MA, PhD

¹Department of Emergency Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

²Department of Civil Engineering, University of California at Berkeley, Berkeley, CA, United States

Corresponding Author:

Lori Ann Post, BA, MA, PhD

Department of Emergency Medicine

Feinberg School of Medicine

Northwestern University

420 E. Superior St

Chicago, IL, 60611

United States

Phone: 1 203 980 7107

Email: lori.post@northwestern.edu

Abstract

Background: SARS-CoV-2, the novel coronavirus that causes COVID-19, is a global pandemic with higher mortality and morbidity than any other virus in the last 100 years. Without public health surveillance, policy makers cannot know where and how the disease is accelerating, decelerating, and shifting. Unfortunately, existing models of COVID-19 contagion rely on parameters such as the basic reproduction number and use static statistical methods that do not capture all the relevant dynamics needed for surveillance. Existing surveillance methods use data that are subject to significant measurement error and other contaminants.

Objective: The aim of this study is to provide a proof of concept of the creation of surveillance metrics that correct for measurement error and data contamination to determine when it is safe to ease pandemic restrictions. We applied state-of-the-art statistical modeling to existing internet data to derive the best available estimates of the state-level dynamics of COVID-19 infection in the United States.

Methods: Dynamic panel data (DPD) models were estimated with the Arellano-Bond estimator using the generalized method of moments. This statistical technique enables control of various deficiencies in a data set. The validity of the model and statistical technique was tested.

Results: A Wald chi-square test of the explanatory power of the statistical approach indicated that it is valid ($\chi^2_{10}=1489.84$, $P<.001$), and a Sargan chi-square test indicated that the model identification is valid ($\chi^2_{946}=935.52$, $P=.59$). The 7-day persistence rate for the week of June 27 to July 3 was 0.5188 ($P<.001$), meaning that every 10,000 new cases in the prior week were associated with 5188 cases 7 days later. For the week of July 4 to 10, the 7-day persistence rate increased by 0.2691 ($P=.003$), indicating that every 10,000 new cases in the prior week were associated with 7879 new cases 7 days later. Applied to the reported number of cases, these results indicate an increase of almost 100 additional new cases per day per state for the week of July 4-10. This signifies an increase in the reproduction parameter in the contagion models and corroborates the hypothesis that economic reopening without applying best public health practices is associated with a resurgence of the pandemic.

Conclusions: DPD models successfully correct for measurement error and data contamination and are useful to derive surveillance metrics. The opening of America involves two certainties: the country will be COVID-19-free only when there is an effective vaccine, and the “social” end of the pandemic will occur before the “medical” end. Therefore, improved surveillance metrics are needed to inform leaders of how to open sections of the United States more safely. DPD models can inform this reopening in combination with the extraction of COVID-19 data from existing websites.

KEYWORDS

COVID-19; models; surveillance; COVID-19 surveillance system; dynamic panel data; infectious disease modeling; reopening America; COVID-19 guidelines; COVID-19 health policy

Introduction

Background

The SARS-CoV-2 pandemic is unprecedented [1,2], with high mortality and morbidity of the virus due to its rapid spread worldwide [3,4]. Without an effective vaccine [5-7], countries are at risk for continued spread [8]. Without good health surveillance, public health leaders are unaware of where and how the disease is spreading. Effective surveillance can inform the safe reopening of economies [9-22] by geographical region [23]. To that end, we submit this proof of concept of the creation of surveillance metrics that correct for measurement error and data contamination. This study applies state-of-the-art statistical modeling to existing data mined from the internet to derive the best available estimates of the state-level dynamics of COVID-19 infection to determine if the sustained decline in SARS-CoV-2 infection that is necessary to reopen is occurring or, conversely, if reopening without applying best public health practices is resulting in a resurgence of SARS-CoV-2.

Public health surveillance is defined as the “ongoing systematic collection, analyses, and interpretation of outcome-specific data for use in the planning, implementation and evaluation of public health practice [18].” Unfortunately, existing surveillance methods suffer from undercounts, bias, and error, and they mostly include more severe cases [24-32]. Research has confirmed that best practices for containment of the COVID-19 pandemic include closing borders between countries [33,34], extreme quarantine measures [35-37], social isolation at home [38], social distancing [39], hand hygiene [40-42], crowd control [43], and wearing a mask in public [44,45]; however, health surveillance must inform where and when to employ these best practices. Due to delays in reporting of new cases, deaths, and testing [46-48], these decisions are made based on partial evidence. Existing models of COVID-19 contagion rely on parameters such as the basic reproduction number (R_0), which are difficult to measure in real time, and they use static statistical methods that do not capture all of the relevant dynamics [49], such as varying specificity and sensitivity of diagnostic testing or asymptomatic individuals who are never tested and are unwittingly carrying SARS-CoV-2 [25,50]. The epidemiological definition of R_0 is the average number of people who contract a disease from a contagious person. It applies specifically to a population of people who were previously free of infection and were not vaccinated [51]. Existing surveillance systems use data that are subject to significant measurement error and other contaminants [52,53]. Moreover, timely information is needed to improve statistical methods that extract information from data sets posted on websites [54-56].

The conventional approach to modeling the spread of diseases such as COVID-19 is to posit an underlying contagion model [57] and then to seek accurate direct measurement of the model

parameters, such as reproduction rates or other parameters; these measurements are sometimes inferred through deaths, hospitalizations, and caseloads [58], and they often involve labor-intensive methods that rely on contact tracing to determine the spread of the disease among a sample population [54,59-61]. For viral epidemics with an incubation period of up to 14 days [62], weeks if not months are required to generate accurate parameter estimates, even for simple contagion models. For example, early estimates of COVID-19 were estimated using methods developed by Lipsitch [63] applied to data from contact tracing in Wuhan and Italy; however, the statistical properties were weak [64-70]. For example, Zhao [65] estimated the serial interval distribution and R_0 based on only six pairs of cases [71]. These models also rely on underlying assumptions about immunity, common propensity for infection, and well-mixed populations, among others. Improvements in these models typically focus on relaxing these assumptions, such as disaggregating the population by geography and modeling within-geography and cross-geography personal interactions [3]. Martcheva [76] provides an excellent dynamic analysis of a wide variety of contagion models and their possible dynamics [72-77]. Unfortunately, they provide limited options for the statistical inference of parameter values from actual data [76]. The objective of this study is to derive surveillance metrics using methods that control for data limitations and contamination.

Methods

Model Development

In contrast to previous studies, we used an empirical approach that focuses on statistical modelling of widely available empirical data, such as the number of confirmed cases or the number of tests, which can inform estimates of the current values of critical parameters such as the infection rate or reproduction rate. We explicitly recognized that the data generating process for the reported data contains an underlying contagion component; a politico-economic component, such as availability of accurate test kits; a social component, such as how strongly people adhere to social distancing measures, mask requirements, and shelter-in-place policies; and a sometimes inaccurate data reporting process that may obscure the underlying contagion process. Therefore, we sought to develop a statistical approach that can provide meaningful information despite the complex and sometimes obfuscating data generation process. Our approach is consistent with the principles of evidence-based medicine, including controlling for complex pathways that may include socioeconomic factors such as mediating variables and policy recommendations, and “based on the best available knowledge, derived from diverse sources and methods [5].”

There are two primary advantages to this empirical approach. First, we can apply the empirical model relatively quickly to a

short data set. This advantage stems from the panel nature of the model. We used US states as the cross-sectional variable; therefore, one week of data from 52 states and territories (including Puerto Rico and the District of Columbia) provides a reasonable sample size. In addition to enabling parameter estimation early in a pandemic, using this property, we tested to see if a shift had occurred in the infection or reproduction rates of the contagion process in the past week (ie, whether there is statistical evidence that reopening is associated with an acceleration in the number of cases).

The second advantage of our approach is that it directly measures and informs policy-relevant variables. For example, the White House issued guidance on reopening the US economy that depends on a decrease in the documented number of cases and in the proportion of positive test results over a 14-day period, among other criteria and considerations [23,78-83]. As noted above, the number and proportion of positive test results are the outcomes of a data generating process that includes not only the underlying contagion process but a multitude of mediating factors as well as idiosyncrasies of the data collection and a delayed reporting process. We specifically modeled the number of positive test results in our empirical model, which provides evidence of direct use in policy dialogue.

Herein, we proceed with a brief discussion of the contagion models that informed our selection of an empirical model. We describe the basic dynamic panel data (DPD) approach and its advantages for analyzing the current pandemic. We obtained results that validate the model specification, which is a necessary and important step in the development of a surveillance system [9-11,14,15,18,20]. We then used the validated model to interrogate our research question: is reopening associated with increased infection transmission and a re-emergence of the pandemic? We approached this research question by statistically testing whether R-type contagion parameters and, specifically, the daily and weekly persistence increased during the weeks of June 27-July 3 and July 4-10, 2020.

Representing Contagion as a DPD Model

Transmission models are typically population-based differential equations of the form $dY/dt = f(Y,X)$, where Y is a vector of a population or subpopulation characteristic of interest, such as the number of exposed or infected individuals; X is a vector of mediating factors (often omitted); and f is a transition function. For empirical purposes, we will use difference equations because the data come in discrete time periods, specifically days. For example, the sizes of the susceptible, infected, and recovered populations in the susceptible-infected-recovered (SIR) model in difference equation form are:

$$\begin{aligned} \Delta(S_{it} - S_{(it-1)}) &= -(\beta S_{(it-1)} I_{(it-1)})/N_{(it-1)} \\ \Delta(I_{it} - I_{(it-1)}) &= (\beta S_{(it-1)} I_{(it-1)})/N_{(it-1)} - [\gamma R_{(it-1)} \\ &+ \gamma D_{(it-1)}] I_{(it-1)} \\ \Delta(R_{it} - R_{(it-1)}) &= \gamma R_{(it-1)} \\ \Delta(D_{it} - D_{(it-1)}) &= \gamma D_{(it-1)} \end{aligned} \quad (1)$$

where S , I , and R are the sizes of the susceptible, infected, and recovered populations, respectively; D is the number of deaths due to SARS-CoV-2; N is the size of the total population ($S + I + R + D$); and the subscripts denote the time period. The first line represents the change in the susceptible population, which

decreases when a susceptible individual becomes infected. This occurs when the susceptible individual interacts with another individual who is infected, in which case the virus is transmitted to the susceptible individual with probability I/N . The number of infected individuals increases by the number of newly infected individuals and decreases by the number of previously infected individuals who either recovered or died. The γ parameters are the probability of recovering or dying. β and the γ are the unknown parameters of the model. Calibration of contagion models requires estimation of the true parameter values.

The availability of state-level data suggests that Equation 1 can be rewritten in panel regression form as

$$\begin{aligned} \Delta(S_{it} - S_{(it-1)}) &= -(\beta S_{(it-1)} I_{(it-1)})/N_{(it-1)} \\ &+ \varepsilon_{Sit} \\ \Delta(I_{it} - I_{(it-1)}) &= (\beta S_{(it-1)} I_{(it-1)})/N_{(it-1)} - [\gamma R_{(it-1)} \\ &+ \gamma D_{(it-1)}] I_{(it-1)} + \varepsilon_{Iit} \\ \Delta(R_{it} - R_{(it-1)}) &= \gamma R_{(it-1)} + \varepsilon_{Rit} \\ \Delta(D_{it} - D_{(it-1)}) &= \gamma D_{(it-1)} + \varepsilon_{Dit} \end{aligned} \quad (2)$$

The additional index i refers to the state; therefore, I_{it} represents the number of infected people in state i at time t . Consistent with the panel data specifications, we added a state-specific “fixed effect” to each of the equations, γ_i , which represents time-invariant state characteristics such as population rate. The ε_{it} represent error terms.

We apply the dynamic panel data approach to the number of positive test results per day as reported on internet sites. To avoid imposing too much specificity, we allowed for some flexibility in the functional form by including the number of tests both linearly and quadratically and as a proportion of the population:

$$\begin{aligned} (3) \quad P_{it} &= \beta_1 P_{it-1} + \beta_2 I_{6,27} P_{it-1} + \beta_3 I_{7,04} P_{it-1} + \beta_4 P_{it-7} + \beta_5 I_{6,27} P_{it-7} + \beta_6 I_{7,04} P_{it-7} + \\ &\gamma_{0,i} + \gamma_1 T_{it} + \gamma_2 T_{it}^2 + \gamma_3 \frac{T_{it}}{Pop_i} + \gamma_4 I_{7,04} + \varepsilon_{it} \end{aligned}$$

where P_{it} is the number of new positive test results and T_{it} is the number of tests administered in state i on day t ; $I_{6,27}$ and $I_{7,04}$ are indicator variables for the time periods from June 27-July 3 and July 4-10, 2020, respectively (latest available data at the time of analysis); and Pop_i is the population of state i (assumed to be constant during the sample). Equation 3 is readily interpretable. The terms containing a β parameter represent the dynamic component of the model. The first term on the right side represents a day-to-day persistence effect (ie, every new case the previous day is a risk factor that contributes β_1 new cases to the current day’s caseload). The next two terms allow for shifts in this risk factor (additions or subtractions) for the weeks beginning June 27 and July 4. Analogously, the next three terms represent a 7-day persistence effect and shifts in that effect for the weeks beginning June 27 and July 4. The 7-day persistence effect is the approximate modal time between viral contraction and the appearance of symptoms; therefore, it is related to the reproduction rate (R parameter) in structural contagion models. The final five terms of Equation 3 contain all the contemporaneous effects in the model (the nonhomogeneous component of the difference equation), as in, all the time subscripts occur contemporaneously at time t except for the state fixed effects, which by definition do not change

over time. The first of these terms represents state-specific effects, which are an important control variable in the panel models. The next two terms are linear and quadratic terms of the number of tests administered, while the third term is the number of tests per person. The next three terms represent the effects of the number of tests administered. The fourth term allows for a shift or discontinuity in the level of new infections for the week of July 4-10 because of increasing concern that the pandemic has re-emerged, particularly in the previous 7 days. We would associate a positive shift with an underlying increase in infection rates. The final term is an error term that represents all types of measurement errors.

Data Sources

Case and test data, including the total number of tests administered and the number of positive results, were taken from the COVID Tracking Project [84], which compiles data from multiple sources. Data were accessed from GitHub [85] after 6 PM on July 10, 2020, so that the data would be complete for that day. Population estimates were derived from the 2019 annual state estimates from the US Census Bureau [86].

Estimation

There are three problems with the specification of Equation 3 for estimation purposes. First, the inclusion of lagged dependent variables on the right side means that the errors are autocorrelated and that the usual exogeneity restrictions are violated; therefore, least squares estimates are inappropriate. Second, some variables are omitted, such as all the variables represented in extensions of the SIR model, and other variables that represent socioeconomic factors influencing the contagion, testing, and reporting processes may also have been omitted. Third, our data set has a relatively short time duration, and the asymptotic properties of fixed-effects or random-effects panel data estimators such as statistical efficiency or normality apply as $t \rightarrow \infty$. Use of these estimators with small values of t creates a small-sample problem with unknown or undesirable estimator properties. We applied the Arellano-Bond approach [87,88], which has improved properties for small samples and is appropriate for application to data sets with a small t and large i .

Fortunately, DPD methods can be used to specifically resolve these statistical problems [89-95]. DPD models allow direct estimation of difference equations with panel data, which resolves multiple problems that appear in the COVID-19 data [96]. The technique we used was developed by Arellano and Bond [87], who applied a generalized method of moments (GMM) approach to a dynamic formulation of employment equations, such as the influence of employment levels in a previous period on employment levels in the current period [97-99]. The basic concept translates to the COVID-19 pandemic in the sense that the number of infections in the current period is a function of lagged infection numbers and other variables. In addition, the DPD removes the individual state effects by first differencing the model. Regressions that include a lagged value of the dependent variable violate the exogeneity restrictions for ordinary least squares and panel estimators such as fixed or random effect models because the lagged dependent variable will be correlated with the error term. DPD model

estimation is an application of Hansen's GMM approach to difference equations estimated from panel data [97,100-102]. The GMM approach solves the endogeneity problem [103,104]. Rather than minimizing a loss function such as the sum of squared errors or maximizing a distribution-specific likelihood function, the GMM approach focuses on the identification of restrictions, including exogeneity restrictions. In an estimable model, there are more identifying restrictions than parameters, and the GMM selects the parameter values that come closest to satisfying the overidentifying restrictions [105]. In our application, we used 10 explanatory variables as defined in Equation 3 and 940 overidentifying restrictions (ie, the same order of magnitude as the sample size $n=1040$); therefore, the degrees of freedom were more than sufficient for statistical inference. The GMM procedure requires a set of instrumental variables; in the case of DPDs, the instruments include lags and/or lag differences in the Y variables. These instruments help resolve the endogeneity problem as well as the omitted-variables problem. In addition to addressing the theoretical concerns inherent in the estimation of any difference equation model, the DPD approach addresses multiple statistical issues that are likely to occur in COVID-19 data.

First, the GMM approach is asymptotically efficient; however, it also has good small sample properties, including samples with a large cross-section and a small number of time periods [102]. This is especially important for statistical analysis early in pandemics, when data are not available for a long period of time, as well as for our testing of whether changes in the transmission rate (that may have occurred 1 to 2 weeks ago) have affected the number of positive test results in the past week.

Second, this approach is robust to omitted variables because of its reliance on identifying restrictions and instrumental variables. This is important because we estimate a relatively sparse model that does not include direct controls for mediating factors, data collection issues, or reporting idiosyncrasies.

Third, the approach includes statistical testing of the overidentifying restrictions (ie, whether the empirical model and estimation technique are statistically valid). For this test, we used the Sargan chi-square test.

Fourth, this approach corrects for autocorrelation.

A significant drawback to DPD methods is that they are computationally complex and become very time- and resource-intensive as the number of observations grows.

We used the Arellano-Bond estimation technique developed specifically for DPD applications. We implemented the Arellano-Bond technique using the *xtabond* command in Stata 16.1 (StataCorp LLC).

Model Validation

To validate the significance of the regression, we used a Wald chi-square statistic to test the null hypothesis that the independent variables did not explain the dependent variable (standard goodness-of-fit measures such as R^2 are uninformative in models with a lagged dependent variable). To test the appropriateness of the model, we applied the Sargan chi-square

test. This is a test of the null hypothesis that the (over)identifying restrictions of the model are statistically met; heuristically, this null hypothesis means that the model and estimation procedure are valid. We used $\alpha \leq 5\%$ for tests of statistical significance.

Model Parameters

We report the point estimates and the P values for all model parameters in Equation 3 as well as additional statistical test results and P values for combinations of parameters when of interest. Of interest are the null hypotheses: $\beta_2 = 0$, $\beta_3 = 0$, $\beta_5 = 0$, and $\beta_6 = 0$. These hypotheses jointly represent the hypothesis that there has been no change in the persistence of the pandemic (ie, the number of new COVID-19 cases over the past two weeks has remained relatively constant). We interpreted rejection of one or more of the hypotheses as evidence that the pandemic is evolving differently, with positive parameter values associated with greater persistence and a re-emergence of the pandemic.

Surveillance Reporting

We translated the estimation results into a surveillance reporting context. The dynamic component (Equation 3) is presented in terms of the persistence rate per 100,000 cases, defined as the number of new COVID-19 cases in every 100,000 cases that remained constant, and this component was applied to the

reported infection numbers to determine its effect on the number of cases per state per day. The contemporaneous component was applied to the reported infection numbers to determine its effect on the number of cases per state per day. The two effects were added to obtain a modeled total number of cases per state per day, and this number was multiplied by 52 to obtain a national figure (including the District of Columbia and Puerto Rico but excluding other territories).

Results

Data

The internet data mining effort resulted in a panel (longitudinal data set) with 52 “panels” (50 states, the District of Columbia, and Puerto Rico) using observations from June 13 through July 10, 2020. Before the analysis, outlying and negative values were crosschecked with other reputable COVID-19 data tracking websites, including USA Facts [106] and the Johns Hopkins Coronavirus Resource Center [107]. The data set has $m = 52 \times 28 = 1456$ observations. Because the model requires 8 days of observations to account for various lags and differencing, the model estimation uses $n = 52 \times 20 = 1040$ observations.

Estimation Results

We present the estimation results in Table 1.

Table 1. Arellano-Bond dynamic panel data modeling of the number of daily infections by state from March 20 to July 10, 2020.

Estimation	Coefficient	P value
Variables		
Lagged daily positive cases	0.0630	.31
Lagged daily positive shift, June 27-July 03	0.0977	.14
Lagged daily positive shift, July 04-10	-0.1727	.009
Seven-day lagged daily positive cases	0.5188	<.001
Seven-day lagged daily positive shift, June 27-July 03	0.0118	.90
Seven-day lagged daily positive shift, July 04-10	0.2691	.002
Constant	17.7791	.68
Daily tests	0.0520	<.001
Daily tests squared	-1.54×10^{-7}	.002
Daily tests / population	-86,527	<.001
Fitness measurements		
Wald test of regression significance (χ^2_{10})	1489.84	<.001
Sargan test of overidentifying restrictions (χ^2_{946})	935.52	.59
Test of lagged daily positive cases + shift July 04-10 = 0 (χ^2_1)	-9.92	.002

Model Validation

To examine the model fit, we applied a Wald chi-square test of the null hypothesis that there is no explanatory power in the explanatory variables. The model was statistically significant ($\chi^2_{10}=1489.84$, $P<.001$). The Sargan chi-square test failed to reject the null hypothesis of valid overidentifying restrictions ($\chi^2_{946}=935.52$, $P=.593$).

Model Parameter Estimates

The coefficient on the lagged dependent variable of the number of daily cases that tested positive on the previous day was positive and statistically significant (0.0630, $P<.001$). The shift values for this parameter for the weeks beginning June 27 and July 4, 2020, are 0.0977 ($P=.138$) and -0.1727 ($P=.009$), respectively. The effective parameter value for the week of July 4 is $0.0630 - 0.1727 = -0.1097$ ($P=.002$).

The coefficient on the 7-day lagged dependent variable, the number of daily cases that tested positive 7 days earlier, was positive and statistically significant (0.5188, $P<.001$). The shift values for this parameter for the weeks beginning June 27 and July 4, 2020, are 0.0118 ($P=.897$) and 0.2691 ($P=.002$), respectively. The effective parameter value for the week of July 4 is $0.5188 + 0.2691 = 0.7879$ ($P<.001$).

The coefficient on the linear term in the number of daily tests administered was positive and statistically significant (0.0520, $P<.001$), and the coefficient on the quadratic term was negative and statistically significant ($-1.54e-07$, $P=.002$). The coefficient on the number of daily tests per person was negative and statistically significant ($-86,527$, $P<.001$).

Surveillance Results

Table 2 translates the statistical results into a user-friendly, intuitive surveillance reporting template. The first two rows are the reported number of cases and tests, respectively. The third

row is the estimated 1-day persistence rate, as in, the number of cases estimated on the current day for every 10,000 cases the previous day. The fourth row is the 7-day persistence rate (ie, the estimated number of cases on the current day for every 10,000 cases 7 days prior). The fifth row is the estimated dynamic component of the model in terms of the number of cases per state per day. This was determined by applying the persistence rates from rows 3 and 4 to the average reported number of cases and adding the effects. The sixth row is the estimated contemporaneous component of the model in terms of number of cases per state per day. The seventh row sums the dynamic and contemporaneous effects to obtain the total estimated effect, as in, the estimated number of new positive test results per state per day. The first column contains the described information as state averages for the period of June 27 to July 03, 2020. The second column contains the information for the United States in aggregate. The third and fourth columns show the same data as the first two columns but for the period of July 4 to 10, 2020.

Table 2. Dynamic panel data estimation results for the United States from June 27 to July 10, 2020.

Variable	June 27-July 3		July 4-10	
	State average	National average	State average	National average
Daily average number of cases for the week	909	47,278	1048	54,491
Daily average number of tests for the week	12,281	638,619	12,630	656,741
Estimated daily persistence rate (per 10,000 cases)	1607	1607	-1106	-1106
Estimated 7-day persistence rate (per 10,000 cases)	5306	5306	7816	7816
Estimated dynamic component (number of cases per day)	499	25,968	595	30,923
Estimated contemporaneous component (number of cases per day)	466	24,254	490	25,464
Total number of estimated cases per day	966	50,221	1084	56,387

Discussion

Principal Findings

Our primary findings are that the 7-day persistence rate is statistically significant and important in magnitude and that the 7-day persistence rate increased by almost 50% from the week of June 27-July 3 to the week of July 4-10 (Table 1). The increase in the 7-day persistence translates into an increase from 5306 new cases per 10,000 cases 7 days prior to 7816 new cases per 10,000 cases (Table 2). On average, this resulted in 95 new cases per state per day. Coupled with a modest increase in the contemporaneous component, the combined result is an estimated increase of 118 new cases per state per day or 6166 new cases nationally per day. The increase in the number of new cases per day is indicative of a shift in the underlying contagion transmission and corroborative of the statement that reopening the US economy has increased the contagion reproduction rate.

The coefficients on the daily lagged dependent variable are small in magnitude and do not indicate strong day-to-day persistence. The negative estimated daily persistence rate for the week of July 4 is indicative of a daily “snaggle-tooth” pattern in the number of daily cases at the state level. This simply

indicates that a low number of cases on one day is offset by a high number of cases the next day, probably due to reporting delays and differential testing periods; this pattern appears slightly in the US aggregate data and is strongly evident in the California data. Other states exhibited different snaggle-tooth patterns, including high-incidence states such as Florida, Texas, and Georgia.

The contemporaneous component of the model contributed positively to the number of new daily cases but did not change significantly over the sample period.

Limitations

While DPD is useful in deriving dynamic estimates of the rate of transmission of COVID-19, static numbers using traditional surveillance tools must also be included to obtain a complete understanding of the pandemic.

Conclusions

The DPD model is a statistically validated analysis of reported COVID-19 data and an important addition to the epidemiological toolkit for understanding the progression of the pandemic. It is important to recognize that this is a supplementary tool that does not replace detailed contagion modeling with detailed and specific data for accurate

representation of contagion model parameters. However, there are four salient advantages of the DPD approach. First, this approach enables statistically efficient extraction of information from existing data sets, including statistical validation of results; therefore, it is applicable to the most commonly tracked and reported data in the current pandemic. Second, the tool could be applied relatively quickly after the pandemic started because of its ability to model reported data rather than detailed contact tracing data, which is largely unavailable to date. That is, changes in the evolution of the pandemic can be confirmed much more quickly using panel data than using aggregate data. Third, this approach informs real-time policy decisions, including decisions based on commonly reported data, such as reopening state economies. Fourth, the model results can help inform the parameterization of more traditional contagion models.

This model is consistent in that it shows a higher reproduction rate during the most recent 7 days; this confirms that in general, normal operation should not be resumed in the United States. Rather, empirically validated public health guidelines such as

wearing masks, social distancing, social isolation, hand washing, and avoidance of social gatherings should be immediately adopted to reduce the contagion. In fact, White House guidelines recommend 14 sustained days of reduced COVID-19–related deaths, new infection cases, and proportions of positive test results prior to reopening. That threshold has not been met. While these findings reflect the national average, it is possible that some areas within the United States meet the White House guidelines, even though reopening is contraindicated in general.

The opening of America involves two certainties. First, the United States will be COVID-19–free only when there is an effective vaccine. While scientists are working at unprecedented speed worldwide to develop a SARS-CoV-2 vaccine [6,108-113], realistically, it will be necessary to rely on best public health practices to minimize COVID-19 infection and mortality for at least one more year [110,114-116]. Second, the “social” end of the pandemic will occur before the “medical” end [117]; therefore, improved surveillance metrics are needed to inform health policy on opening sections of America more safely.

Acknowledgments

The opinions expressed herein are those of the author(s) and do not necessarily reflect the views of the US Agency for International Development.

Conflicts of Interest

None declared.

References

1. Kumari T, Shukla V. COVID-19: Towards Confronting an Unprecedented Pandemic. *IJBI* 2020;02(01):01-10. [doi: [10.46505/ijbi.2020.2101](https://doi.org/10.46505/ijbi.2020.2101)]
2. Baker S, Bloom N, Davis S, Kost K, Sammon M, Viratyosin T. The unprecedented stock market impact of COVID-19. *NBER Working Papers* 2020 Apr:1-24. [doi: [10.3386/w26945](https://doi.org/10.3386/w26945)]
3. Lai C, Shih T, Ko W, Tang H, Hsueh P. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *Int J Antimicrob Agents* 2020 Mar;55(3):105924 [FREE Full text] [doi: [10.1016/j.ijantimicag.2020.105924](https://doi.org/10.1016/j.ijantimicag.2020.105924)] [Medline: [32081636](https://pubmed.ncbi.nlm.nih.gov/32081636/)]
4. Petrosillo N, Viceconte G, Ergonul O, Ippolito G, Petersen E. COVID-19, SARS and MERS: are they closely related? *Clin Microbiol Infect* 2020 Jun;26(6):729-734 [FREE Full text] [doi: [10.1016/j.cmi.2020.03.026](https://doi.org/10.1016/j.cmi.2020.03.026)] [Medline: [32234451](https://pubmed.ncbi.nlm.nih.gov/32234451/)]
5. Collins FS, Stoffels P. Accelerating COVID-19 Therapeutic Interventions and Vaccines (ACTIV): An Unprecedented Partnership for Unprecedented Times. *JAMA* 2020 Jun 23;323(24):2455-2457. [doi: [10.1001/jama.2020.8920](https://doi.org/10.1001/jama.2020.8920)] [Medline: [32421150](https://pubmed.ncbi.nlm.nih.gov/32421150/)]
6. Corey L, Mascola JR, Fauci AS, Collins FS. A strategic approach to COVID-19 vaccine R&D. *Science* 2020 May 29;368(6494):948-950. [doi: [10.1126/science.abc5312](https://doi.org/10.1126/science.abc5312)] [Medline: [32393526](https://pubmed.ncbi.nlm.nih.gov/32393526/)]
7. Cohen J. Vaccine designers take first shots at COVID-19. *Science* 2020 Apr 03;368(6486):14-16. [doi: [10.1126/science.368.6486.14](https://doi.org/10.1126/science.368.6486.14)] [Medline: [32241928](https://pubmed.ncbi.nlm.nih.gov/32241928/)]
8. Shrivastava S, Shrivastava P. World Health Organization Declares Very High Global Risk of Transmission of Corona Virus Disease 2019. *Prim Health Care* 2020 May 25;20(2):339 [FREE Full text]
9. Thacker S, Berkelman R. Public health surveillance in the United States. *Epidemiol Rev* 1988;10(1):164-190. [doi: [10.1093/oxfordjournals.epirev.a036021](https://doi.org/10.1093/oxfordjournals.epirev.a036021)] [Medline: [3066626](https://pubmed.ncbi.nlm.nih.gov/3066626/)]
10. Thacker SB, Berkelman RL, Stroup DF. The Science of Public Health Surveillance. *J Public Health Pol* 1989;10(2):187. [doi: [10.2307/3342679](https://doi.org/10.2307/3342679)]
11. Thacker SB, Stroup DF. Public Health Surveillance and Health Services Research. In: Armenian HK, Shapiro S, editors. *Epidemiology and Health Services*. New York, NY: Oxford University Press; 1998:61-82.
12. Nsubuga P, White ME, Thacker SB, Anderson MA, Blount SB, Broome CV, et al. Public Health Surveillance: A Tool for Targeting and Monitoring Interventions. In: Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M, Evans DB, et al, editors. *Disease Control Priorities in Developing Countries*. 2nd edition. Washington, DC: The International Bank for Reconstruction and Development/The World Bank; 2006.

13. Lee LM, Teutsch SM, Thacker SB, Louis ME. New York, NY: Oxford University Press; 2010.
14. Thacker S, Qualters J, Lee L, Centers for Disease Control and Prevention. Public health surveillance in the United States: evolution and challenges. *MMWR Suppl* 2012 Jul 27;61(3):3-9. [Medline: [22832990](#)]
15. Thacker S, Stroup D. Future directions for comprehensive public health surveillance and health information systems in the United States. *Am J Epidemiol* 1994 Sep 01;140(5):383-397. [doi: [10.1093/oxfordjournals.aje.a117261](#)] [Medline: [8067331](#)]
16. Calain P. Exploring the international arena of global public health surveillance. *Health Policy Plan* 2007 Jan 08;22(1):2-12. [doi: [10.1093/heapol/czl034](#)] [Medline: [17237489](#)]
17. Lee LM, Thacker SB. Public health surveillance and knowing about health in the context of growing sources of health data. *Am J Prev Med* 2011 Dec;41(6):636-640. [doi: [10.1016/j.amepre.2011.08.015](#)] [Medline: [22099242](#)]
18. Teutsch S, Churchill R, editors. Principles and Practice of Public Health Surveillance, Second Edition. In: Oxford University Press, USA. New York, NY: Oxford University Press; Oct 05, 2000.
19. Romaguera R, German R, Klaucke D, Teutsch S, Churchill R. Evaluating Public Health Surveillance. In: Teutsch S, Churchill R, editors. Principles and Practice of Public Health Surveillance, Second Edition. New York, NY: Oxford University Press; 2000:176-193.
20. Teutsch S. Considerations in Planning a Surveillance System. New York, NY: Oxford University Press; 2010:18-28.
21. Baker S, Bloom N, Davis S, Kost K, Sammon M, Viratyosin T. The Unprecedented Stock Market Reaction to COVID-19. *Covid Economics, Vetted and Real-Time Papers* 2020 Mar 03;1(3):33-42. [doi: [10.3386/w26945](#)]
22. Yu J, Pang Y. 'Is it wise to open up America again now?' Learning from government responses to COVID-19 and economy reopen points in three Asia-Pacific countries. *Br J Health Care Manag* 2020 Jul 02;26(7):1-5. [doi: [10.12968/bjhc.2020.0071](#)]
23. Harris J. Reopening Under COVID-19: What to Watch For. *NBER Working Papers* 2020 May:898-2937. [doi: [10.3386/w27166](#)]
24. Weinberger DM, Chen J, Cohen T, Crawford FW, Mostashari F, Olson D, et al. Estimation of Excess Deaths Associated With the COVID-19 Pandemic in the United States, March to May 2020. *JAMA Intern Med* 2020 Jul 01:online. [doi: [10.1001/jamainternmed.2020.3391](#)] [Medline: [32609310](#)]
25. Jewell N, Lewnard J, Jewell B. Caution Warranted: Using the Institute for Health Metrics and Evaluation Model for Predicting the Course of the COVID-19 Pandemic. *Ann Intern Med* 2020 Aug 04;173(3):226-227. [doi: [10.7326/m20-1565](#)]
26. Fielding-Miller R, Sundaram M, Brouwer K. Social determinants of COVID-19 mortality at the county level. *medRxiv* 2020 Jul 01:preprint [FREE Full text] [doi: [10.1101/2020.05.03.20089698](#)] [Medline: [32637976](#)]
27. Faust JS, Del Rio C. Assessment of Deaths From COVID-19 and From Seasonal Influenza. *JAMA Intern Med* 2020 Aug 01;180(8):1045-1046. [doi: [10.1001/jamainternmed.2020.2306](#)] [Medline: [32407441](#)]
28. Wetzler HP, Wetzler EA. COVID-19 Excess Deaths in the United States, New York City, and Michigan During April 2020. *medRxiv* 2020 May 19:preprint. [doi: [10.1101/2020.04.02.20051532](#)]
29. Hart OE, Halden RU. Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: Feasibility, economy, opportunities and challenges. *Sci Total Environ* 2020 Aug 15;730:138875 [FREE Full text] [doi: [10.1016/j.scitotenv.2020.138875](#)] [Medline: [32371231](#)]
30. Quast T, Anandel R. Excess mortality and potential undercounting of COVID-19 deaths by demographic group in Ohio. *medRxiv* 2020 Jul 17:preprint. [doi: [10.1101/2020.06.28.20141655](#)]
31. Ciminelli G, Garcia-Mandicó S. Covid-19 in Italy: an analysis of death registry data. *VOXEU, Centre for Economic Policy Research*. 2020 Apr 22. URL: <https://voxeu.org/article/covid-19-italy-analysis-death-registry-data> [accessed 2020-09-15]
32. Williamson GD, Weatherby Hudson G. A monitoring system for detecting aberrations in public health surveillance reports. *Statist Med* 1999 Dec 15;18(23):3283-3298. [doi: [10.1002/\(sici\)1097-0258\(19991215\)18:23<3283::aid-sim316>3.0.co;2-z](#)]
33. Yang Y, Peng F, Wang R, Yange M, Guan K, Jiang T, et al. The deadly coronaviruses: The 2003 SARS pandemic and the 2020 novel coronavirus epidemic in China. *J Autoimmun* 2020 May;109:102434 [FREE Full text] [doi: [10.1016/j.jaut.2020.102434](#)] [Medline: [32143990](#)]
34. Ng Y, Li Z, Chua YX, Chaw WL, Zhao Z, Er B, et al. Evaluation of the Effectiveness of Surveillance and Containment Measures for the First 100 Patients with COVID-19 in Singapore - January 2-February 29, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Mar 20;69(11):307-311 [FREE Full text] [doi: [10.15585/mmwr.mm6911e1](#)] [Medline: [32191691](#)]
35. Pastor C. Sentiment Analysis on Synchronous Online Delivery of Instruction due to Extreme Community Quarantine in the Philippines caused by Covid-19 Pandemic. *AJMS* 2020 Mar 20;3(1):1-6 [FREE Full text]
36. Piguillem F, Shi L. Optimal Covid-19 Quarantine and Testing Policies: CEPR Discussion Paper No. DP14613. SSRN 2020 Apr:1-42.
37. Sjödin H, Wilder-Smith A, Osman S, Farooq Z, Rocklöv J. Only strict quarantine measures can curb the coronavirus disease (COVID-19) outbreak in Italy, 2020. *Eurosurveillance* 2020 Apr 02;25(13):2000280. [doi: [10.2807/1560-7917.es.2020.25.13.2000280](#)]
38. Grant A. The reproduction number R for COVID-19 in England: Why hasn't "lockdown" been more effective? *medRxiv* 2020:preprint. [doi: [10.1101/2020.07.02.20144840](#)]
39. Lewnard JA, Lo NC. Scientific and ethical basis for social-distancing interventions against COVID-19. *Lancet Infect Dis* 2020 Jun;20(6):631-633. [doi: [10.1016/s1473-3099\(20\)30190-0](#)]

40. Cavanagh G, Wambier CG. Rational hand hygiene during the coronavirus 2019 (COVID-19) pandemic. *J Am Acad Dermatol* 2020 Jun;82(6):e211 [FREE Full text] [doi: [10.1016/j.jaad.2020.03.090](https://doi.org/10.1016/j.jaad.2020.03.090)] [Medline: [32268168](https://pubmed.ncbi.nlm.nih.gov/32268168/)]
41. Water, sanitation, hygiene and waste management for COVID-19: technical brief, 03 March 2020. World Health Organization. 2020 Mar 03. URL: <https://apps.who.int/iris/handle/10665/331305> [accessed 2020-09-15]
42. Abtahi-Naeini B. Frequent handwashing amidst the COVID-19 outbreak: prevention of hand irritant contact dermatitis and other considerations. *Health Sci Rep* 2020 Jun 27;3(2):e163 [FREE Full text] [doi: [10.1002/hsr2.163](https://doi.org/10.1002/hsr2.163)] [Medline: [32346616](https://pubmed.ncbi.nlm.nih.gov/32346616/)]
43. Tabari P, Amini M, Moghadami M, Moosavi M. International Public Health Responses to COVID-19 Outbreak: A Rapid Review. *Iran J Med Sci* 2020 May;45(3):157-169 [FREE Full text] [doi: [10.30476/ijms.2020.85810.1537](https://doi.org/10.30476/ijms.2020.85810.1537)] [Medline: [32546882](https://pubmed.ncbi.nlm.nih.gov/32546882/)]
44. Doung-ngern P, Suphanchaimat R, Panjangampatthana A, Janekrongtham C, Ruampoom D, Daochaeng N, et al. Associations between Wearing Masks, Washing Hands, and Social Distancing Practices, and Risk of COVID-19 Infection in Public: A Cohort-Based Case-Control Study in Thailand. *SSRN Journal* 2020 Aug 14;preprint. [doi: [10.2139/ssrn.3633162](https://doi.org/10.2139/ssrn.3633162)]
45. Infection prevention and control guidance for long-term care facilities in the context of COVID-19: interim guidance, 21 March 2020. World Health Organization. 2020 Mar 21. URL: <https://apps.who.int/iris/handle/10665/331508> [accessed 2020-09-15]
46. Zhong B, Luo W, Li H, Zhang Q, Liu X, Li W, et al. Knowledge, attitudes, and practices towards COVID-19 among Chinese residents during the rapid rise period of the COVID-19 outbreak: a quick online cross-sectional survey. *Int J Biol Sci* 2020;16(10):1745-1752 [FREE Full text] [doi: [10.7150/ijbs.45221](https://doi.org/10.7150/ijbs.45221)] [Medline: [32226294](https://pubmed.ncbi.nlm.nih.gov/32226294/)]
47. Zhang L, Liu Y. Potential interventions for novel coronavirus in China: A systematic review. *J Med Virol* 2020 May 03;92(5):479-490 [FREE Full text] [doi: [10.1002/jmv.25707](https://doi.org/10.1002/jmv.25707)] [Medline: [32052466](https://pubmed.ncbi.nlm.nih.gov/32052466/)]
48. Stoecklin S, Rolland P, Silue Y. First cases of coronavirus disease 2019 (COVID-19) in France: surveillance, investigations and control measures, January 2020. *Eurosurveillance* 2020 Jan;25(6):2000094. [doi: [10.2807/1560-7917.es.2020.25.6.2000094](https://doi.org/10.2807/1560-7917.es.2020.25.6.2000094)]
49. Tan SXD, Chen L. Real-Time Differential Epidemic Analysis and Prediction for COVID-19 Pandemic. *arXiv*. 2020 Apr 15. URL: <https://arxiv.org/abs/2004.06888> [accessed 2020-09-15]
50. Gandhi M, Yokoe D, Havlir D. Asymptomatic Transmission, the Achilles' Heel of Current Strategies to Control Covid-19. *N Engl J Med* 2020 May 28;382(22):2158-2160. [doi: [10.1056/nejme2009758](https://doi.org/10.1056/nejme2009758)]
51. Cespedes MDS, Souza JCRPD. Coronavirus: a clinical update of Covid-19. *Rev Assoc Med Bras (1992)* 2020 May 15;66(2):116-123. [doi: [10.1590/1806-9282.66.2.116](https://doi.org/10.1590/1806-9282.66.2.116)] [Medline: [32428144](https://pubmed.ncbi.nlm.nih.gov/32428144/)]
52. Calvo R, Deterding S, Ryan R. Health surveillance during covid-19 pandemic. *BMJ* 2020 Apr 06;369:m1373. [doi: [10.1136/bmj.m1373](https://doi.org/10.1136/bmj.m1373)] [Medline: [32253180](https://pubmed.ncbi.nlm.nih.gov/32253180/)]
53. Chiolerio A, Buckeridge D. Glossary for public health surveillance in the age of data science. *J Epidemiol Community Health* 2020 Jun 24;74(7):612-616. [doi: [10.1136/jech-2018-211654](https://doi.org/10.1136/jech-2018-211654)] [Medline: [32332114](https://pubmed.ncbi.nlm.nih.gov/32332114/)]
54. Desjardins M, Hohl A, Delmelle E. Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. *Appl Geogr* 2020 May;118:102202 [FREE Full text] [doi: [10.1016/j.apgeog.2020.102202](https://doi.org/10.1016/j.apgeog.2020.102202)] [Medline: [32287518](https://pubmed.ncbi.nlm.nih.gov/32287518/)]
55. Stokes EK, Zambrano LD, Anderson KN, Marder EP, Raz KM, El Burai Felix S, et al. Coronavirus Disease 2019 Case Surveillance - United States, January 22-May 30, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Jun 19;69(24):759-765 [FREE Full text] [doi: [10.15585/mmwr.mm6924e2](https://doi.org/10.15585/mmwr.mm6924e2)] [Medline: [32555134](https://pubmed.ncbi.nlm.nih.gov/32555134/)]
56. Barros JM, Duggan J, Rebholz-Schuhmann D. The Application of Internet-Based Sources for Public Health Surveillance (Infoveillance): Systematic Review. *J Med Internet Res* 2020 Mar 13;22(3):e13680 [FREE Full text] [doi: [10.2196/13680](https://doi.org/10.2196/13680)] [Medline: [32167477](https://pubmed.ncbi.nlm.nih.gov/32167477/)]
57. Iacopini I, Petri G, Barrat A, Latora V. Simplicial models of social contagion. *Nat Commun* 2019 Jun 06;10(1):2485 [FREE Full text] [doi: [10.1038/s41467-019-10431-6](https://doi.org/10.1038/s41467-019-10431-6)] [Medline: [31171784](https://pubmed.ncbi.nlm.nih.gov/31171784/)]
58. Stock JH. Data gaps and the policy response to the novel coronavirus. *NBER Working Papers* 2020 Mar:1-12. [doi: [10.3386/w26902](https://doi.org/10.3386/w26902)]
59. Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z, et al. Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation and Diagnosis for COVID-19. *IEEE Rev Biomed Eng* 2020 Apr 16:1-1. [doi: [10.1109/rbme.2020.2987975](https://doi.org/10.1109/rbme.2020.2987975)]
60. Naudé W. Artificial Intelligence against COVID-19: An early review. *IZA Institute of Labor Economics*. 2020 Apr. URL: <http://ftp.iza.org/dp13110.pdf> [accessed 2020-09-15]
61. Tang Z, Zhao W, Xie X, Zhong Z, Shi F, Liu L, et al. Severity assessment of coronavirus disease 2019 (COVID-19) using quantitative features from chest CT images. *arXiv*. 2020 Mar 26. URL: <https://arxiv.org/abs/2003.11988> [accessed 2020-09-15]
62. Interim Clinical Guidance for Management of Patients with Confirmed Coronavirus Disease (COVID-19). US Centers for Disease Control and Prevention. 2020. URL: <https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html> [accessed 2020-07-25]
63. Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, et al. Transmission dynamics and control of severe acute respiratory syndrome. *Science* 2003 Jun 20;300(5627):1966-1970. [doi: [10.1126/science.1086616](https://doi.org/10.1126/science.1086616)] [Medline: [12766207](https://pubmed.ncbi.nlm.nih.gov/12766207/)]

64. Lipsitch M, Swerdlow DL, Finelli L. Defining the Epidemiology of Covid-19 — Studies Needed. *N Engl J Med* 2020 Mar 26;382(13):1194-1196. [doi: [10.1056/nejmp2002125](https://doi.org/10.1056/nejmp2002125)]
65. Zhao S, Gao D, Zhuang Z. Estimating the serial interval of the novel coronavirus disease (COVID-19): A statistical analysis using the public data in Hong Kong from January 16 to February 15, 2020. *medRxiv* 2020 Feb 25:preprint. [doi: [10.1101/2020.02.21.20026559](https://doi.org/10.1101/2020.02.21.20026559)]
66. Gatto M, Bertuzzo E, Mari L, Miccoli S, Carraro L, Casagrandi R, et al. Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proc Natl Acad Sci USA* 2020 May 12;117(19):10484-10491 [FREE Full text] [doi: [10.1073/pnas.2004978117](https://doi.org/10.1073/pnas.2004978117)] [Medline: [32327608](https://pubmed.ncbi.nlm.nih.gov/32327608/)]
67. You C, Deng Y, Hu W, Sun J, Lin Q, Zhou F, et al. Estimation of the time-varying reproduction number of COVID-19 outbreak in China. *Int J Hyg Environ Health* 2020 Jul;228:113555 [FREE Full text] [doi: [10.1016/j.ijheh.2020.113555](https://doi.org/10.1016/j.ijheh.2020.113555)] [Medline: [32460229](https://pubmed.ncbi.nlm.nih.gov/32460229/)]
68. Sanche S, Lin YT, Xu C, Romero-Severson E, Hengartner N, Ke R. High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2. *Emerg Infect Dis* 2020 Jul;26(7):1470-1477 [FREE Full text] [doi: [10.3201/eid2607.200282](https://doi.org/10.3201/eid2607.200282)] [Medline: [32255761](https://pubmed.ncbi.nlm.nih.gov/32255761/)]
69. Giordano G, Blanchini F, Bruno R, Colaneri P, Di Filippo A, Di Matteo A, et al. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nat Med* 2020 Jun 22;26(6):855-860 [FREE Full text] [doi: [10.1038/s41591-020-0883-7](https://doi.org/10.1038/s41591-020-0883-7)] [Medline: [32322102](https://pubmed.ncbi.nlm.nih.gov/32322102/)]
70. Peng L, Yang W, Zhang D, Zhuge C, Hong L. Epidemic analysis of COVID-19 in China by dynamical modeling. *medRxiv* 2020 Feb 18:preprint. [doi: [10.1101/2020.02.16.20023465](https://doi.org/10.1101/2020.02.16.20023465)]
71. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. *Int J Infect Dis* 2020 Apr;93:284-286 [FREE Full text] [doi: [10.1016/j.ijid.2020.02.060](https://doi.org/10.1016/j.ijid.2020.02.060)] [Medline: [32145466](https://pubmed.ncbi.nlm.nih.gov/32145466/)]
72. Sha A, Samanta S, Martcheva M, Chattopadhyay J. Backward bifurcation, oscillations and chaos in an eco-epidemiological model with fear effect. *J Biol Dyn* 2019 Dec 03;13(1):301-327. [doi: [10.1080/17513758.2019.1593525](https://doi.org/10.1080/17513758.2019.1593525)] [Medline: [31046638](https://pubmed.ncbi.nlm.nih.gov/31046638/)]
73. Chen L. Persistent spatial patterns of interacting contagions. *Phys Rev E* 2019 Feb 14;99(2):022308. [doi: [10.1103/physreve.99.022308](https://doi.org/10.1103/physreve.99.022308)]
74. Kwuimy CAK, Nazari F, Jiao X, Rohani P, Nataraj C. Nonlinear dynamic analysis of an epidemiological model for COVID-19 including public behavior and government action. *Nonlinear Dyn* 2020 Jul 16:1-15 [FREE Full text] [doi: [10.1007/s11071-020-05815-z](https://doi.org/10.1007/s11071-020-05815-z)] [Medline: [32836814](https://pubmed.ncbi.nlm.nih.gov/32836814/)]
75. Li X, Yang J, Martcheva M. Age-Structured Epidemic Models. In: *Age Structured Epidemic Modeling. Interdisciplinary Applied Mathematics*, vol 52. Cham, Switzerland: Springer; May 29, 2020:23-67.
76. Martcheva M. *An Introduction to Mathematical Epidemiology*. Cham, Switzerland: Springer; 2015.
77. Martcheva M, Castillo-Chavez C. Diseases with chronic stage in a population with varying size. *Mathematical Biosciences* 2003 Mar;182(1):1-25. [doi: [10.1016/s0025-5564\(02\)00184-0](https://doi.org/10.1016/s0025-5564(02)00184-0)]
78. Wiley L. Public health law and science in the community mitigation strategy for Covid-19. *J Law Biosci* 2020;7(1):Isaa019 [FREE Full text] [doi: [10.1093/jlb/Isaa019](https://doi.org/10.1093/jlb/Isaa019)] [Medline: [32728464](https://pubmed.ncbi.nlm.nih.gov/32728464/)]
79. Lambert C, Stoicu S, Nestsiarovich A, Kumar P, Lauve N, Sloane L, et al. 2020-04-17/18 Daily UNM Global Health COVID-19 Briefing. University of New Mexico Digital Repository. 2020 Apr 24. URL: https://digitalrepository.unm.edu/hsc_covid19_briefings/10/ [accessed 2020-09-15]
80. Alhaery M, Suh E. A COVID-19 Reopening Readiness Index: The Key to Opening up the Economy. SSRN 2020 Jun 07:preprint. [doi: [10.1101/2020.05.22.20110577](https://doi.org/10.1101/2020.05.22.20110577)]
81. Ganz S. Does your state have 14-days of declining COVID-19 cases? Finding straightforward answers to a surprisingly tricky question. American Enterprise Institute. 2020 May 11. URL: https://www.jstor.org/stable/resrep24601?seq=2#metadata_info_tab_contents [accessed 2020-09-15]
82. Baqaee D, Farhi E, Mina M, Stock J. Reopening Scenarios. NBER Working Papers 2020 May:1-42. [doi: [10.3386/w27244](https://doi.org/10.3386/w27244)]
83. Khan Z, Van BF, Hussain F. A predictive model for Covid-19 spread applied to six US states. *arXiv*. 2020 Jun 10. URL: <https://arxiv.org/abs/2006.05955> [accessed 2020-09-15]
84. The COVID Tracking Project. The Atlantic. 2020. URL: <https://covidtracking.com/> [accessed 2020-07-10]
85. COVID 19 Tracking 2020. GitHub. 2020. URL: <https://github.com/CSSEGISandData/COVID-19> [accessed 2020-09-15]
86. 2019 National and State Population Estimates. United States Census Bureau. 2019 Dec 30. URL: <https://www.census.gov/newsroom/press-kits/2019/national-state-estimates.html> [accessed 2020-08-15]
87. Arellano M, Bond S. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Rev Econ Stud* 1991 Apr;58(2):277. [doi: [10.2307/2297968](https://doi.org/10.2307/2297968)]
88. Montes-Rojas G, Sosa-Escudero W, Zincenko F. Level-Based Estimation of Dynamic Panel Models. *J Econom Method* 2019 Aug 01;9(1):1-23. [doi: [10.1515/jem-2018-0015](https://doi.org/10.1515/jem-2018-0015)]
89. Fritsch M. On GMM estimation of linear dynamic panel data models. *Passauer Diskussionspapiere-Betriebswirtschaftliche Reihe No. B-36-19*. 2019 Sep 20. URL: <https://www.econstor.eu/bitstream/10419/204581/1/1678187852.pdf> [accessed 2020-09-15]
90. Liu L, Moon HR, Schorfheide F. Forecasting With Dynamic Panel Data Models. *Econometrica* 2020;88(1):171-201. [doi: [10.3982/ecta14952](https://doi.org/10.3982/ecta14952)]

91. Kock AB, Tang H. Uniform inference in high-dimensional dynamic panel data models with approximately sparse fixed effects. *Econom Theory* 2018 May 24;35(2):295-359. [doi: [10.1017/s0266466618000087](https://doi.org/10.1017/s0266466618000087)]
92. Bera AK, Doğan O, Taşpınar S, Leiluo Y. Robust LM tests for spatial dynamic panel data models. *Regional Science and Urban Economics* 2019 May;76:47-66. [doi: [10.1016/j.regsciurbeco.2018.08.001](https://doi.org/10.1016/j.regsciurbeco.2018.08.001)]
93. Judson RA, Owen AL. Estimating dynamic panel data models: a guide for macroeconomists. *Econ Lett* 1999 Oct;65(1):9-15. [doi: [10.1016/s0165-1765\(99\)00130-5](https://doi.org/10.1016/s0165-1765(99)00130-5)]
94. Bowsher CG. On testing overidentifying restrictions in dynamic panel data models. *Econ Lett* 2002 Oct;77(2):211-220. [doi: [10.1016/s0165-1765\(02\)00130-1](https://doi.org/10.1016/s0165-1765(02)00130-1)]
95. Blundell R, Bond S, Windmeijer F. Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator. *Adv Econ* 2020 Jul 01;15:online. [doi: [10.1016/S0731-9053\(00\)15003-0](https://doi.org/10.1016/S0731-9053(00)15003-0)]
96. Anser MK, Yousaf Z, Khan MA, Nassani AA, Alotaibi SM, Qazi Abro MM, et al. Does communicable diseases (including COVID-19) may increase global poverty risk? A cloud on the horizon. *Environ Res* 2020 Aug;187:109668 [FREE Full text] [doi: [10.1016/j.envres.2020.109668](https://doi.org/10.1016/j.envres.2020.109668)] [Medline: [32422482](https://pubmed.ncbi.nlm.nih.gov/32422482/)]
97. Ogaki M. Generalized method of moments: Econometric applications. In: Maddala GS, Rao CR, Vinod HD, editors. *Handbook of Statistics Volume 11*. Amsterdam, Netherlands: Elsevier; Nov 01, 1993:455-488.
98. Roodman D. How to do Xtabond2: An Introduction to Difference and System GMM in Stata. *Stata J* 2018 Nov 19;9(1):86-136. [doi: [10.1177/1536867x0900900106](https://doi.org/10.1177/1536867x0900900106)]
99. Brown BW, Newey WK. Generalized Method of Moments, Efficient Bootstrapping, and Improved Inference. *J Bus Econ Stat* 2002 Oct;20(4):507-517. [doi: [10.1198/073500102288618649](https://doi.org/10.1198/073500102288618649)]
100. Hansen BE, West KD. Generalized Method of Moments and Macroeconomics. *J Bus Econ Stat* 2002 Oct;20(4):460-469. [doi: [10.1198/073500102288618603](https://doi.org/10.1198/073500102288618603)]
101. Stock JH, Wright JH, Yogo M. A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *J Bus Econ Stat* 2002 Oct;20(4):518-529. [doi: [10.1198/073500102288618658](https://doi.org/10.1198/073500102288618658)]
102. Hansen LP. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica* 1982 Jul;50(4):1029-1054. [doi: [10.2307/1912775](https://doi.org/10.2307/1912775)]
103. Windmeijer FAG, Santos Silva JMC. Endogeneity in count data models: an application to demand for health care. *J Appl Econ* 1997 May;12(3):281-294. [doi: [10.1002/\(sici\)1099-1255\(199705\)12:3<281::aid-jae436>3.0.co;2-1](https://doi.org/10.1002/(sici)1099-1255(199705)12:3<281::aid-jae436>3.0.co;2-1)]
104. Lewbel A. Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models. *J Bus Econ Stat* 2012 Feb 22;30(1):67-80. [doi: [10.1080/07350015.2012.643126](https://doi.org/10.1080/07350015.2012.643126)]
105. Newey WK. Generalized method of moments specification testing. *J Econom* 1985 Sep;29(3):229-256. [doi: [10.1016/0304-4076\(85\)90154-x](https://doi.org/10.1016/0304-4076(85)90154-x)]
106. USA Facts. URL: <https://usafacts.org/issues/coronavirus/> [accessed 2020-09-16]
107. COVID-19 Dashboard. Johns Hopkins Coronavirus Resource Center. URL: <https://coronavirus.jhu.edu/map.html> [accessed 2020-09-16]
108. Lurie N, Saville M, Hatchett R, Halton J. Developing Covid-19 Vaccines at Pandemic Speed. *N Engl J Med* 2020 May 21;382(21):1969-1973. [doi: [10.1056/nejmp2005630](https://doi.org/10.1056/nejmp2005630)]
109. Shoenfeld Y. Corona (COVID-19) time musings: Our involvement in COVID-19 pathogenesis, diagnosis, treatment and vaccine planning. *Autoimmun Rev* 2020 Jun;19(6):102538 [FREE Full text] [doi: [10.1016/j.autrev.2020.102538](https://doi.org/10.1016/j.autrev.2020.102538)] [Medline: [32268212](https://pubmed.ncbi.nlm.nih.gov/32268212/)]
110. Yamey G, Schäferhoff M, Hatchett R, Pate M, Zhao F, McDade KK. Ensuring global access to COVID-19 vaccines. *Lancet* 2020 May;395(10234):1405-1406. [doi: [10.1016/s0140-6736\(20\)30763-7](https://doi.org/10.1016/s0140-6736(20)30763-7)]
111. Wu S. Progress and Concept for COVID-19 Vaccine Development. *Biotechnol J* 2020 Jun;15(6):e2000147 [FREE Full text] [doi: [10.1002/biot.202000147](https://doi.org/10.1002/biot.202000147)] [Medline: [32304139](https://pubmed.ncbi.nlm.nih.gov/32304139/)]
112. Thanh Le T, Andreadakis Z, Kumar A, Gómez Román R, Tollefsen S, Saville M, et al. The COVID-19 vaccine development landscape. *Nat Rev Drug Discov* 2020 May;19(5):305-306. [doi: [10.1038/d41573-020-00073-5](https://doi.org/10.1038/d41573-020-00073-5)] [Medline: [32273591](https://pubmed.ncbi.nlm.nih.gov/32273591/)]
113. Cohen J. U.S. 'Warp Speed' vaccine effort comes out of the shadows. *Science* 2020 May 15;368(6492):692-693. [doi: [10.1126/science.368.6492.692](https://doi.org/10.1126/science.368.6492.692)] [Medline: [32409451](https://pubmed.ncbi.nlm.nih.gov/32409451/)]
114. Graham BS. Rapid COVID-19 vaccine development. *Science* 2020 May 29;368(6494):945-946. [doi: [10.1126/science.abb8923](https://doi.org/10.1126/science.abb8923)] [Medline: [32385100](https://pubmed.ncbi.nlm.nih.gov/32385100/)]
115. Schaffer DeRoo S, Pudalov NJ, Fu LY. Planning for a COVID-19 Vaccination Program. *JAMA* 2020 Jun 23;323(24):2458-2459. [doi: [10.1001/jama.2020.8711](https://doi.org/10.1001/jama.2020.8711)] [Medline: [32421155](https://pubmed.ncbi.nlm.nih.gov/32421155/)]
116. Smiai H, Mirimi M, Najar M, Boukhatem N. COVID-19 update : what's going on? *J Mater Environ Sci* 2020 May 13;11(6):877-884.
117. Post LA. When will the Covid-19 pandemic be over? In: Post LA, editor. *Buehler Bulletin*. Evanston, IL: Buehler Center Media; 2020.

Abbreviations

DPD: dynamic panel data

GMM: generalized method of moments

R₀: basic reproduction number

SIR: susceptible-infected-recovered

Edited by G Eysenbach, G Fagherazzi; submitted 01.06.20; peer-reviewed by J Marquard, M Adenomou, C Moss, G Onel; comments to author 09.07.20; revised version received 19.07.20; accepted 09.09.20; published 22.09.20

Please cite as:

Oehmke JF, Oehmke TB, Singh LN, Post LA

Dynamic Panel Estimate–Based Health Surveillance of SARS-CoV-2 Infection Rates to Inform Public Health Policy: Model Development and Validation

J Med Internet Res 2020;22(9):e20924

URL: <http://www.jmir.org/2020/9/e20924/>

doi: [10.2196/20924](https://doi.org/10.2196/20924)

PMID: [32915762](https://pubmed.ncbi.nlm.nih.gov/32915762/)

©James Francis Oehmke, Theresa B Oehmke, Lauren Nadya Singh, Lori Ann Post. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 22.09.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.