REOPENING UNDER COVID-19: WHAT TO WATCH FOR

Jeffrey E. Harris

Working Paper 27166
http://www.nber.org/papers/w27166

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2020

The comments of Robin Bell, Peter Boomer, Ken Boynton, Gil Brodsky, Zack Cooper, Michael Cragg, Denise Everett, Richard Florida, Daniel Geselowitz, Ray Girouard, Beatriz González López-Valcárcel, Lara Gosce, Bill Goyette, Robert Hanlon, Jarrett Harris, Bill James, Tom Kalb, Louis Kaplow, John Lowell, Mitch Metzner, Andrew Racine, Tom Reichert, Brina Sedar, Rivana Cohen Stadtlander, Pat Tracy, Heidi Wald, Judy Wieber, Delbert Yoder, and John Ziegler are gratefully acknowledged. The opinions expressed here are solely those of the author and do not represent the views of the Massachusetts Institute of Technology, Eisner Health, the National Bureau of Economic Research, or any other individual or organization. The author has received no direct or indirect remuneration for this article, and has no conflicts of interest to declare. This is the third article in a series. For the first two articles, see (Harris 2020a, b). The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Jeffrey E. Harris. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.
Reopening Under COVID-19: What to Watch For
Jeffrey E. Harris
NBER Working Paper No. 27166
May 2020
JEL No. I11,I12,I14,I18,I24

ABSTRACT

We critically analyze the currently available status indicators of the COVID-19 epidemic so that state governors will have the guideposts necessary to decide whether to further loosen or instead retighten controls on social and economic activity. Overreliance on aggregate, state-level data in Wisconsin, we find, confounds the effects of the spring primary elections and the outbreak among meat packers. Relaxed testing standards in Los Angeles may have upwardly biased the observed trend in new infection rates. Reanalysis of New Jersey data, based upon the date an ultimately fatal case first became ill rather than the date of death, reveals that deaths have already peaked in that state. Evidence from Cook County, Illinois shows that trends in the percentage of positive tests can be wholly misleading. Trends on emergency department visits for influenza-like illness, advocated by the White House Guidelines, are unlikely to be informative. Data on hospital census counts in Orange County, California suggest that healthcare system-based indicators are likely to be more reliable and informative. An analysis of cumulative infections in San Antonio, Texas, shows how mathematical models intended to guide decisions on relaxation of social distancing are severely limited by untested assumptions. Universal coronavirus testing may not on its own solve difficult problems of data interpretation and causal inference.

Jeffrey E. Harris
Department of Economics, E52-422
MIT
50 Memorial Drive
Cambridge, MA 02142
and NBER
jeffrey@mit.edu
Introduction

This article is motivated by a striking disconnect. On the one hand, various think tanks, foundations, and academic institutions have issued their own formal plans to guide the reopening of the U.S. economy in the face of the ongoing COVID-19 epidemic (Allen et al. 2020, Watson et al. 2020, Rockefeller Foundation 2020, Romer 2020). These white papers, for the most part, envision the widespread if not universal testing for active coronavirus infection as fundamental to the nation’s recovery. On the other hand, governors of numerous states throughout the country have already issued orders as to what worksites and places of congregation can reopen and under what conditions (Murphy 2020, Kemp 2020, Evers 2020a, b, Abbott 2020c, Polis 2020).

One side is abstract, conceptual, and far-sighted. The other side is concrete, microscopically focused on details, and keyed to the here and now. Both sides have nothing but the best intentions. They are, however, living in different worlds.

The gap between these two approaches may narrow once we come up with rapid, inexpensive tests for viral antigens to replace the current time-consuming, expensive tests for viral RNA, and once we decide collectively as a society how to coerce each other into being tested, traced and isolated. And, of course, everything could change for the better if we had highly effective antiviral medications and even a moderately effective vaccine, or if we jointly developed adequate herd immunity, or if the virus ultimately mutated to adapt to its human host (Taubenberger and Morens 2006, Packer 2020). Until then, we will need what economists call a second-best strategy for a judicious recovery.

Since we don’t yet have universal compulsory testing to give us a detailed, complete and timely snapshot of the epidemic, we will have to make do for now with a diverse collection of imperfect indicators. If we cannot critically analyze these status indicators in order to reliably discern the underlying trends, we will have more than a few problems trying to determine whether various state governments’ efforts to rekindle economic and social activity have been working or failing. The governors of these states will lack the guideposts necessary to decide whether to further loosen or instead retighten controls. Imagine trying to bring a plane to a soft landing when you don’t really know its altitude or velocity.

What follows is a detailed vetting of the most salient status indicators. For the most part, our analysis shies away from modeling of future trends based upon untested assumptions. Instead, we focus sharply on real historical data. Our objective is not to exhaustively cover every
state, county and city in the country. Instead, we attempt to learn lessons from selected cases. We make no pretense that these cases were drawn randomly.

**Incidence of Infection Based on Partial Voluntary Testing: Wisconsin**

Figure 1 below shows the daily incidence of newly diagnosed COVID-19 infections in all counties of the state of Wisconsin combined from March 15 through May 4, 2020. The incidence, measured in terms of numbers of cases per 100,000 population, is rendered on a logarithmic scale so that a straight line would represent exponential growth (Harris 2020a). The incidence data are calculated as the numbers of positive tests for coronavirus infection (Wisconsin Department of Health Services 2020) divided by the state’s residential population (Wisconsin Department of Administration 2020). The numbers of positive tests in turn are derived solely from individuals who voluntarily sought testing.

![Graph showing daily incidence of new COVID-19 cases per 100,000 population in Wisconsin counties from March 15 to May 4, 2020. The graph includes markers for key events such as spring primary elections, “Badger Bounces Back” order, and “Safer At Home” order.](image)

*Figure 1. Daily Incidence of New COVID-19 Cases per 100,000 Population, All Wisconsin Counties, March 15 – May 4, 2020*

The arrows point to a few of the potentially relevant events during the time period covered in Figure 1. On April 20, 2020, Gov. Tony Evers announced his Badger Bounces Back plan to reopen the Wisconsin economy (Evers 2020a). Following the most recent federal guidelines (White House 2020), Evers’ plan continued the state’s Safer At Home restrictions,
issued March 24 (Palm 2020a), which kept non-essential businesses closed until certain criteria were met, including a sustained 14-day decline in documented cases. Golf courses, however, were allowed open, and exterior lawn care was permitted (Evers 2020b). A week later, the state’s secretary of health services announced an Interim Order to Turn the Dial, allowing non-essential businesses to make curbside drop-offs and opening up outdoor recreational rentals and self-service car washes, so long as social distancing measures remained in place (Palm 2020b).

The observation period covered Figure 1 also includes the April 7 spring primary elections, when the number of polling places in Milwaukee was reduced from 180 to five, and where some voters had to wait up to two and half hours to cast their ballots (Rakich 2020). “Now, over two weeks later,” wrote the newly elected justice to the Wisconsin Supreme Court in a recent opinion article, “we have an uptick in Covid-19 cases, especially in dense urban centers like Milwaukee and Waukesha, where few polling places were open and citizens were forced to stand in long lines to cast a ballot.” (Karofsky 2020). Press reports have quoted Milwaukee’s health officials that from 7 to 26 new cases of coronavirus infection appeared to be related to in-person voting at the primaries (Associated Press 2020, Spicuzza 2020).

What inferences can we draw, if any, about the trends in Figure 1 and their relationship to the highlighted events? After an upsurge in cases during the weeks of March 15 and March 22, the incidence of new infections in Figure 1 appears to have leveled off. More recently, however, new daily cases have been increasing statewide. On April 14, there were 127 new infections, yielding a statewide incidence of 2.17 cases per 100,000 population. By April 25, the number of new infections had increased to 331, yielding a rate of 5.66 per 100,000.

With an incubation period from initial infection to onset of symptoms averaging 5 days (Linton et al. 2020, Li et al. 2020), and with the additional delay between the onset of symptoms and the performance of a diagnostic test, no one is going to attribute the recent doubling of daily coronavirus cases to the governor’s April 20 Badger Bounces Back order or the April 27 Interim Order to Turn the Dial. But the data do raise the question whether the original Safer at Home restrictions of March 24 were enough. They certainly do not satisfy the original requirement for a sustained 14-day decline in documented cases.

Figure 2 below shows the incidence of new coronavirus cases in three Wisconsin counties, Milwaukee County (the orange points), Waukesha County (the yellow points), and Brown County (the purple points), which houses the city of Green Bay. Once again, we would
not expect to see a rise in incidence during the week immediately following the primaries. Without definitive data on delays between symptoms and testing, it is difficult to determine whether an increase in cases attributable to the primary elections would have occurred 1 or 2 weeks later. We obtained similar results with data from the City of Milwaukee (Mukai 2020).

Figure 2. Daily Incidence of New COVID-19 Cases per 100,000 Population, Milwaukee, Waukesha, and Brown Counties, Wisconsin, March 15 – May 4, 2020

Table 1 shows the average daily incidence by week, where “week 0” started on primary day, Tuesday April 7. In Milwaukee and Waukesha counties, the incidence does not increase above baseline until the third week starting April 28. This is certainly not strong evidence in favor of a large effect of the primary elections on the incidence of new cases. It does not, however, rule out a smaller effect detectable only through detailed case finding.

Table 1. Average Daily Incidence by Week, Selected Counties, Wisconsin

<table>
<thead>
<tr>
<th>Week</th>
<th>Milwaukee</th>
<th>Waukesha</th>
<th>Brown</th>
<th>All Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7.4</td>
<td>2.0</td>
<td>2.1</td>
<td>1.4</td>
</tr>
<tr>
<td>1</td>
<td>6.8</td>
<td>1.8</td>
<td>11.3</td>
<td>1.2</td>
</tr>
<tr>
<td>2</td>
<td>7.3</td>
<td>1.5</td>
<td>30.9</td>
<td>1.6</td>
</tr>
<tr>
<td>3</td>
<td>9.4</td>
<td>1.9</td>
<td>32.4</td>
<td>3.0</td>
</tr>
</tbody>
</table>
The data for Brown County both Figure 2 and Table 1, by contrast, show a marked increase in the incidence of new cases. The Brown County data are in fact quantitatively responsible for the apparent increase in the overall statewide incidence seen at the same time in Figure 1. The cumulative total of 971 cases through April 29 is attributable mostly to outbreaks of COVID-19 infection at three meat packing facilities. At one plant, there were reportedly 262 cases of coronavirus infection among employees and 86 cases of secondary spread among the employees’ contacts (Amundsen 2020). On April 28, two days after the plant had voluntarily closed, the president issued an executive order, under the authority of the Defense Production Act, declaring that “[s]uch closures threaten the continued functioning of the national meat and poultry supply chain, undermining critical infrastructure during the national emergency.” (Trump 2020)

Our examination of the data for Wisconsin has so far taught us three important lessons.

(1) We need to be very cautious about relying on state-level data to make policy decisions about reopening. In this case, state-level data obscured critical events occurring at the micro or county level.

(2) Even if state governments are granted the power to manipulate the levers of renormalization, they need to take account of strategic interventions at the federal executive level.

(3) It may be difficult to make causal inferences about the consequences of state policies to relax or tighten social distancing rules.

Would we have drawn different conclusions if we had data from universal testing of new coronavirus infections? Researchers suspect that the actual incidence rates are from twice to ten times those shown in Figures 1 and 2 (Sutton et al. 2020). The main reason for the undercount is presumably that many people have had infections so mild that they did not seek medical attention. But Figures 1 and 2 would remain valid guides to the actual trend in COVID-19 infection rates so long as undercount factor remained unchanged (Harris 2020a). We’ll investigate the validity of this key assumption below.

**Deaths Attributable to COVID-19: Los Angeles County**

Figure 3 below shows the numbers of coronavirus-positive cases (pink points) and COVID-19 attributable deaths (mango points) in Los Angeles County during March 1 – April
30, 2020. Since there is only one geographic area with a constant population, we have not converted the counts into incidence rates. The horizontal axis ticks off the date that the test was performed or the date of death (Los Angeles County Department of Public Health 2020). This dating convention differs from Figures 1 and 2 above, where the horizontal axis marked the date the report was received. The convention is more accurate in pinning down the test date, but it has the disadvantage that the more recent counts remain unreliable until all the reports are in.

While the rapid increase in positive cases during the first week in March has slowed, new cases in Figure 3 still appear to be increasing. The counts of COVID-19 attributable deaths, on the other hand, appear to be flat. The Los Angeles County Department of Public Health originally discouraged healthcare providers from ordering tests on presumptively infected but uncomplicated patients. As tests have become increasingly available, the Department has relaxed its position and, in fact, will now test anyone with or without any symptoms.

The apparent continuing rise in positive cases may thus be no more than an artifact of increased testing. But does that mean the data points showing COVID-19 attributable deaths are
more reliable? Figure 4 below addresses this question. In the figure, we have duplicated the plot of Figure 3, but with one critical modification. As indicated by the leftward pointing gray arrow, we’ve shifted the data series on deaths backward by 16 days. By one estimate, that’s the average time from the onset of symptoms until a patient dies of complications (Muzimoto and Chowell 2020). Instead of the mango data points (which we’ve partially erased), we now have light gray data points to indicate when those patients who will ultimately succumb to the virus first got sick. The light gray data points are almost perfectly parallel to the pink data points on coronavirus-positive cases. The near-constant vertical distance between the two plots indicates a near-constant case fatality rate of 9 percent. For a related approach, see (Baud et al. 2020).

![Figure 4. Same Plot as Figure 3, But with COVID-19 Attributable Deaths Backed Up by 16 Days.](image)

We’ve now learned the following additional lessons.

(4) Data from partial voluntary testing can give a misleading picture of the state of the epidemic if standards for testing have been recently changing.

(5) Data on case fatality rates can serve as an indicator of the degree of penetration of voluntary testing. When case fatality rates are unusually high, testing penetration is likely to be low.
What To Watch For

Jeffrey E. Harris

12-May-2020

(6) While data on deaths attributable to COVID-19 may avoid this bias, they lag data on testing-based incidence by about 16 days.

(7) The inherent lag in COVID-19 attributable deaths will prevail even if we had near-perfect data on incidence derived from universal testing.

(8) With an incubation period of about 5 days from infection to symptoms, the data derived from symptom-based testing offer at least a reasonably rapid indicator of what’s happening in real time. But when we add another 16 days from symptoms to death, we’re talking about a three-week total lag time from infection to death.

Deaths Attributable to COVID-19: New Jersey

We turn to the problem faced by Gov. Philip D. Murphy of New Jersey, who on Wednesday, April 29, 2020 reopened state parks to fishing, hunting and hiking, while keeping picnic areas and playgrounds closed and limiting parking to 50% of maximum capacity (Murphy 2020). State golf courses were likewise reopened so long as minimum social distancing policies were observed. Yet the following day, the New York Times relayed, “New Jersey reported 460 new virus-related deaths on Thursday, more than any other state in the nation. … The increase came in a week when Mr. Murphy, encouraged by other measures that showed New Jersey making progress in fighting the virus, began to sketch out how the state might reopen in the weeks ahead.” (Badger et al. 2020)

![New reported cases by day in New Jersey](image)

**Figure 5. Screen Shot from New York Times, New Jersey Coronavirus Map and Case Count, May 1, 2020**
Figure 5 above shows a screen shot of the *New York Times* data on daily cases and deaths in New Jersey as of April 30, based upon the date of report of each event (New York Times 2020). The bars show the daily numbers, with the solid connected line segments showing the 7-day moving averages. While the counts of voluntary testing-based infections are falling, the number of reported deaths is still surging. If thousands of New Jerseyans go back to their state parks and golf courses, are we risking thousands of additional deaths and tens of thousands of new infections before we can belatedly step on the brakes?

Figure 6 below resolves the apparent contradiction. The data, published by the Communicable Disease Service of the New Jersey Department of Health (New Jersey Department of Health 2020), show the numbers of newly reported COVID-19 infections according to the date the test was performed. Likewise, the counts of COVID-19 attributable deaths are related to the date the test was performed on the same patients. Put differently, New Jersey’s Communicable Disease Service has already been backed up each death to the date when the patient got sick, just as we did in Figure 4 for Los Angeles. As before, we’ve omitted the most recent week because counts based on the date of testing are subject to reporting delays.

*Figure 6. COVID-19 Cases and Deaths According to Date of Testing, New Jersey, March 1 – April 19, 2020.*
Figure 6 makes clear that the numbers of patients with fatal COVID-19 infections in New Jersey had already peaked and turned downward about three weeks before Gov. Murphy reopened the state’s parks and golf courses. As in Figure 4, the vertical distance between the two data series reflects the case fatality rate, with a large gap reflecting lower fatality. In fact, the case fatality rate dropped from 10.6 percent for the week starting March 1 to 4.4 percent for the week starting April 12, once again suggesting that the apparent stagnation of the case counts resulted from enhanced testing of less seriously ill individuals.

The analysis of Figure 6 does not eliminate the problem that COVID-19 attributable deaths are a lagged status variable. In order to confirm that fatal COVID-19 infections actually peaked around the first of April, we still had to wait until the end of the month for all the death reports to come in. Still, the figure teaches us another important lesson.

(9) *For the purposes of deciding whether to relax or tighten controls, we should be thinking not only about dates of death, but also about the dates that ultimately fatal cases first became sick.*

It is arguable, nonetheless, that the graphs of COVID-19 attributable deaths by date of death are more relevant to strategic decision making. While we defer a full treatment of this point to a later section, here is the basic argument. Strategic decision makers need to be concerned not only about overall trends in the path of the epidemic, but also on the real-time capacity of the healthcare system. Since resource requirements are likely to be maximal as the patient approaches death, graphs of COVID-19 attributable deaths by date of death inform us about the demands on those resources. Put differently, if all 460 deaths reported in New Jersey on April 30 had occurred among patients admitted to the same medical center, the center’s resources would have surely been overwhelmed, no matter when they were first admitted.

**Deaths Attributable to COVID-19: New York City**

In Figure 7 below, we inquire: Are we accounting for all COVID-19 attributable deaths? The data points, derived from the New York City Department of Health database (Montesano 2020b), show numbers of deaths confirmed by coronavirus testing (orange points) and the combined numbers of confirmed and probable deaths attributable to COVID-19 (purple points). In view of current limitations on data availability, the horizontal axis reverts to measuring the date of death rather than the date the patient was first tested.
Figure 7 strongly suggests that deaths confirmed by coronavirus testing may significantly understate the total number of deaths attributable to COVID-19. The possible but not confirmed deaths are likely to include patients who tested negative or were never tested, but who succumbed to acute adult respiratory distress syndrome, cardiogenic shock, cytokine storm, overwhelming acute kidney injury, and strokes from massive coagulopathy. At least from the last week of March onward, the inclusion of probable cases increases the total death counts by about one third.

![Figure 7. Confirmed and Combined Confirmed and Probable COVID-19 Deaths by Date of Death, New York City, March 8 – April 25, 2020.](image)

Whether the undercount reflects narrow limits on the availability of coronavirus testing or broader limits on our healthcare system remains unclear. Still, the lessons of Figure 7 are clear.

(10) No matter what definition is used, deaths from COVID-19 reached their peak in New York City during the second week of April, and are now declining.

(11) The actual demands on our healthcare system to identify and adequately treat seriously ill COVID-19 patients are likely to have been seriously underestimated.
COVID-19 Hospitalization Rates: New York City

Figure 8 below shows the daily counts of COVID-19 cases, hospitalizations and deaths in New York City in relation to the applicable date of testing, hospitalization or death (Montesano 2020a) from March 1 through April 25, 2020. Once again, to avoid undercounts due to reporting delays, provisional counts after April 25 are excluded. As we would expect, the trend in deaths attributable to COVID-19, when recorded by date of death, lags the trend in test-confirmed cases.

![Figure 8. Daily Cases, Hospitalizations and Confirmed Deaths from COVID-19, March 1 – April 25, 2020.](image)

Our focus here is on hospitalizations (the green points). The excess of hospitalizations over positive tests during the first week in March suggests the possibility that hospital admission was delayed among some individuals who were actually infected in the month of February. There is also an expected time lag between the onset of initial symptoms leading an individual to get tested and subsequent hospitalization for worsening illness. This time delay may have changed during the course of the month of March. Still, the curve of hospitalizations reaches a peak during the first week of April.
Figure 8 communicates three important messages.

(12) As a potential status variable, daily hospitalizations gauge the incidence of seriously ill cases that place significant demands on healthcare resources.

(13) The time lag between new infection and hospitalization appears to be significantly shorter than the lag between infection and death.

(14) The peak incidence of hospitalizations could serve as an indicator of peak demands on healthcare resources so long as hospital admissions are not rationed as a result of bed unavailability or personnel constraints.

Daily Hospital Census: Orange County, California

Focusing further on the demand for healthcare resources, Figure 9 shows the current daily census attributable to COVID-19 in a universe of 25 reporting hospitals in Orange County, California (Orange County Health Care Agency 2020). The number of hospitals reporting on any given day varied from 19 to 25, with an average of 23. Accordingly, the data points in Figure 9 below show the scaled-up census for all 25 hospitals, based upon the assumption that each nonreporting hospital had the same census as the average of the reporting hospitals. The best-fitting regression line covering the points from April 10 onward gives a doubling time of the total COVID-19 census in a bit more than a month. The results are basically the same without the imputation for nonreporting hospitals.

The findings in Figure 9 are important and ominous. As the weather has improved in Southern California, people have been flocking to some beaches in Orange County, culminating in the arrival of an estimated 40 thousand beachgoers at Newport Beach on the April 25-26 weekend (Lozano 2020, Baxter, Wigglesworth, and Chang 2020, Andone and Vercammen 2020, Connelly and Kopetman 2020). In response, Gov. Gavin Newsom of California shut down beaches in Orange County on April 30 (Badger et al. 2020). At this juncture, we cannot attribute the increasing census of COVID-19 cases in Orange County hospitals to the recent influx of thousands of beachgoers. Press reports suggest, in fact, that many of the attendees were from neighboring cities and counties, including Los Angeles, Santa Monica, Laguna Beach, Malibu, and Venice, where beaches remained closed. If they came down with COVID-19 and became seriously ill, they might not end up in an Orange County hospital.

Still, Figure 9 teaches us another important lesson.

(15) Highly focused measures of intensive healthcare demands in micro areas may be critical status variables for the implementation of effective reopening strategies.

Percent Positive Tests Among Those Tested: Cook County, Illinois

We move on to another measure based on voluntary testing, namely, the percentage of positive COVID-19 tests among those tested. This indicator is featured in the White House Guidelines (White House 2020), which specifically recommend reliance on a “Downward trajectory of positive tests as a percent of total tests within a 14-day period (flat or increasing volume of tests).” Why the volume of tests needs to be flat or increasing remains unclear. It is well acknowledged that voluntary testing results are based on a self-selected sample. At least during the initial response to the outbreak, healthcare providers on the front lines were routinely counseling patients with fever, cough, shortness of breath and loss of taste to stay home and not get tested so long as they were stable. With expanding availability of tests even to those with minimal or no symptoms, the percentage positive would tend to be increasingly biased downward, thus giving the false impression of a favorable trend in disease rates.
Figure 10 illustrates the point, using data from Cook County, Illinois, which includes Chicago (Illinois Department of Public Health 2020). During the observation period, the volume of tests was steadily increasing, from 224 on March 15, to 5,402 on April 5, to 19,417 by May 3. Yet the percent tested positive, as indicated by the sky-blue data points has been declining since the first week in April. That’s certainly more than 14 days. Yet the daily incidence of newly diagnosed COVID-19 cases per 100,000 population, as indicated by the orange data points, has continued to increase. It would seem that the percent tested positive is a wholly misleading indicator of the epidemic trend.

![Figure 10](image)

Figure 10. Percent Tested Positive and Incidence of New COVID-19 Cases per 100,000 Population, Cook County, Illinois, March 10 – May 8, 2020

Figure 10 points to the following conclusion.

(16) Trends in the percentage of positive tests among persons who have voluntarily tested are uninformative, and can in fact be misleading.

Emergency Department Visits for Influenza-Like Illnesses: New York City

We focus next on emergency department visits for influenza-like illnesses. We include this indicator primarily because a “downward trajectory of influenza-like illnesses (ILI) reported
within a 14-day period” was at the top of the list of “gating criteria” featured in the White House Guidelines (White House 2020), and not because it appears to be particularly informative.

Figure 11 superimposes two trends. The first trend (the dark red line segments) shows the percentage of visits to New York City emergency departments for ILI from the start of the annual flu season in October 2019 through mid-March 2020 (New York Department of Health and Mental Hygiene 2020). One can see the wave of emergency department (ED) visits taking off after Thanksgiving and peaking in January 2020. This wave represented the flu season for this past year. Just when it appeared that the seasonal wave was coming to an end, however, a new wave of ED visits took off on March 1. This new upstroke, which reached 10 percent of ED visits (as gauged on the left-hand axis) was undoubtedly powered by the rapidly emerging COVID-19 epidemic.

The problem with this impressive graph is that a new wave of COVID-19 could come at any time. If it happened to arrive during the regular seasonal wave of influenza next fall or winter, we won’t have any idea from the ILI data alone whether it was COVID-19 or the flu.

What’s worse, the timing of waves of both seasonal and pandemic influenza is quite varied. The second trend in Figure 10 (the lavender area bounded above by black line segments) shows the death rates for the 1918 influenza pandemic of 1918–1919 (as gauged on the right-hand axis). (See Figure 1 of (Taubenberger and Morens 2006).) The first wave (not shown in Figure 11 above) came out of nowhere during July 1918. The second wave (as shown in Figure 11) picked up steam in October 1919, with a peak mortality about 5 times that of the first wave. Then, after a lull in January 1919, the third wave came rolling in, peaking in March 1919.
One could argue, of course, that surveillance data on ILI could be used in combination the other indicators reviewed above. All of these other status variables are specific to COVID-19. So why not just rely upon them alone?

(17) *Trends in emergency department visits for influenza-like illnesses, while a basic staple of public health reporting, are unlikely to serve as reliable status indicators of COVID-19 resurgence or decline.*

**New COVID-19 Cases: San Antonio**

In an executive order on March 19, 2020, Texas Gov. Greg Abbott declared that “every person in Texas shall avoid social gatherings in groups of 10 or more people.” (Abbott 2020a) In the same order, Abbott mandated that “people shall avoid eating or drinking at bars, restaurants, and food courts, or visiting gyms or massage parlors,” with an exception for “drive-thru, pickup, or delivery options.” (Abbott 2020a) The order was continued on March 31, barring attendance at nursing homes and keeping schools closed to in-classroom instruction (Abbott 2020b). In an April 27, 2020 partial reversal of his original orders, Gov. Abbott allowed non-essential retail establishments, movie theaters, and shopping malls to reopen for in-store services provided that the stores operate at no more than 25 percent occupancy (Abbott 2020d). (p. 3)

Figure 12 below shows the daily incidence of new COVID-19 infections per 100,000 population in the City of San Antonio, Texas, from March 19 – May 4, 2020 (City of San Antonio 2020). Superimposed on the raw data points is the *centered, 7-day moving average.* Figure 12 shows a rapid uptake of new COVID-19 infections during the second half of March, followed by an apparently abrupt leveling off on April 6. Once again, given the lag between initial infection and subsequent testing, that looks a bit too soon to attribute the change to the governor’s March 31 renewal of his order. Figure 13, also shown below, provides exactly the same plot with a superimposed *left-sided, 7-day moving average.* The trend in new COVID-19 infections now appears to be leveling off is now on April 9. If we were to take any steps to smooth or otherwise digest the raw data in Figures 12 and 13, the most appropriate method would be to aggregate the rates by week.

---

* The centered 7-day moving average at discrete time $t$ equals $\left( x_{t-3} + x_{t-2} + x_{t-1} + x_t + x_{t+1} + x_{t+2} + x_{t+3} \right) / 7$.

† The leftsided 7-day moving average at discrete time $t$ equals $\left( x_{t-6} + x_{t-5} + x_{t-4} + x_{t-3} + x_{t-2} + x_{t-1} + x_t \right) / 7$. 
Figure 11. Daily Incidence of New COVID-19 Cases per 100,000 Population, San Antonio, March 19 – May 4, 2020 With Superimposed Centered, 7-Day Moving Average

Figure 13. Daily Incidence of New COVID-19 Cases per 100,000 Population, San Antonio, March 19 – May 4, 2020 With Superimposed Left-sided, 7-Day Moving Average
In view of the significant variability in the daily diagnosis counts, it is difficult to discern when the incidence curve flattened or whether it has in fact started to resume an upward trend. The use of moving averages doesn’t really add any information to the raw data, but it can create the false impression of breaks in the data or reversals of trend that simply aren’t there.

(18) Moving averages may make noisy data plots look smoother, but they are to be avoided. Moving averages are no substitute for additional, more reliable data on trends in COVID-19 incidence.

**Cumulative COVID-19 Cases: San Antonio**

Figure 14 shows another way to filter out the noise in the San Antonio data. Graphed here are the cumulative number of infections to date. The sky-blue points show the raw data. The dark red curve is an application of a classical SIR mathematical model of the spread of an epidemic, where $S$ is for “susceptible,” $I$ is for “infective,” and $R$ is for “resistant.” (Kermack and McKendrick 1991, Harris 2020a)

The details of the SIR model are given in the Appendix. In this particular application of the model, which we describe as “unrestricted,” almost everyone is the population starts out as susceptible (S), a handful start out as infective (I), and no one is naturally resistant (R). As the susceptible and infective persons repeatedly engage in contact with each other, the former become infected by the latter. Over time, the I’s lose their infectivity and convert to R’s, either by getting better or dying. By selecting the appropriate mathematical constants for this model application, we were able to track the cumulative incidence for about the first four weeks of the city’s epidemic.

At some point around April 12, the data points on cumulative infections start to deviate from the unrestricted model. To accommodate this deviation, we’ve modified the unrestricted model to take account of social distancing. That’s the solid black curve in the figure. The basic idea is that every day, an additional fraction of the susceptible population avoids contact with others, including the infectives. This modification appears to fit the data points quite well.

After April 27, however, when Gov. Abbott partially releases the restrictions in his original order, a third alternative model takes over, as indicated by the purple curve. The basic idea is that some of the susceptible people who had reduced contact with others begin comingle with others.

As a general matter, the underlying concepts of the three related models are entirely reasonable. And, in fact, a more sophisticated version of the basic SIR model that also takes into account the presence of asymptomatic carriers has been applied to the San Antonio data (Gutierrez 2020). But there are two critical problems that give us pause when we seek to rely upon such modeling exercises.

First, it’s easy to fit a mathematical model to the cumulative counts of the COVID-19 cases. That’s because the cumulative cases just keep going up. Fitting a model to the underlying incidence data, which jump up and down in Figures 12 and 13, is a lot harder. The apparent fit to the cumulative counts makes us overconfident that the model is correct.

Second, all the action is in the purple curve. This curve is generated under the critical assumption that a specific percentage of susceptible people who had isolated themselves will abandon their efforts. But that specific percentage is an untested number. We can make the purple curve look better or worse simply by changing it. What appears to be a model driven by the facts is in fact a model based on facts plus one untested assumption.
(19) **Graphs of cumulative cases (or cumulative deaths) provide little or no guidance in the decision to release or tighten social distancing measures. These graphs always go up, and never down.**

(20) **Because graphs of cumulative cases go up smoothly, it’s too easy to fit them to mathematical models. The apparent fit gives us false confidence in our model.**

(21) **When mathematical models are extended to predict future responses to releasing or tightening restrictions, it is critical that we make explicit their untested assumptions.**

**Discussion**

We summarize here what we’ve learned about those status indicators we can reliably monitor as we judiciously loosen controls on economic activity.

Reports of the incidence of new coronavirus infections based solely upon partial, voluntary testing can still provide a reliable picture of the recent path of the epidemic. Increasingly relaxed standards of eligibility for testing, however, can artificially raise the counts of positive cases and thus give the false impression of an upward or sustained trend in infection rates. The observation of a progressively declining case fatality rate can provide a clue to the presence of this potential bias. While some have appropriately proposed representative sampling (Kaplow 2020) or pooled testing (Lakdawalla et al. 2020) in the absence of universal testing, the detailed mechanics of obtaining and properly weighting representative samples have yet to be fully worked out.

Like all status indicators, reports of the incidence of new infections at the aggregate geographic level can obscure critical trends at the micro level, including counties, neighborhoods, and even individual firms and residences. Rules for reopening will need to be adaptable to micro-level events, especially isolated outbreaks in residential facilities, retail establishments serving the public, and firms requiring high concentrations of workers. Examples include the meatpacking firms in Brown County, Wisconsin, discussed here, and the outbreak of frontline transit workers in New York City, discussed in (Harris 2020b).

Counts of deaths from COVID-19 suffer from the serious limitation that these reports lag behind the incidence of new infections by about three weeks. Reliance on trends in the numbers of deaths can result in an entirely misleading interpretation of the effects of the recent tightening or relaxation of controls on social and economic activity. This inherent lag in death rates would
prevail even if we had reliable data on the incidence of infection from near-universal testing. For the purposes of deciding whether to relax or tighten controls, we need to look at the dates that individuals with fatal infections became ill, and not the dates when they died.

Deaths from COVID-19 are undoubtedly being underdiagnosed, especially in the absence of universal testing and full knowledge of the multi-organ system manifestations of severe disease. Still, there is no evidence that incomplete ascertainment of death has seriously biased our analyses of trends in numbers of deaths.

Data on the percentage of positive tests among all persons tested are uninformative and, in fact, can be outright misleading. Trends in emergency department visits for influenza-like illnesses, while a basic staple of public health reporting, are unlikely to serve as reliable status indicators of COVID-19 resurgence or decline. There is just too much potential confounding with other viral illnesses, including influenza. Even seasonal influenza does not keep its appointments on a fixed calendar, and pandemic illnesses – which will remain inevitable – don’t arrive at pre-appointed times.

Hospital-based measures of COVID-19 morbidity are likely to be superior, more stable indicators of underlying trends, and thus more reliable status variables for determining whether and when to relax or tighten economic controls. Reports on the census of hospital cases are also sensitive indicators of the demand for constrained medical resources.

Techniques to smooth trends, such as moving averages, can give the false impression of bends, breaks or abrupt changes in the data. Some of the noise in the data is undoubtedly due to sampling error, but some of the day-to-day variability is inherent in the system of measurement. Even with universal testing for COVID-19 infection, not everybody is going to get tested on the same day.

Cumulative rates of infection or death have little informative value. These numbers always go up. It is easy to fit a predictive model to cumulative rates, and the apparently good fit gives a false impression of reliability. The predictive value of such a model rests on the validity of its underlying untested assumptions, and not on its fit to historical data.

Even if we had perfect data on the incidence of new infections through mandatory, universal testing, we would still need to confront the difficulty of drawing causal connections between the incidence of infection and policy measures to tighten or loosen controls on social and economic activity. In some cases, the temporal and geographic patterns may be strong, as in
a recent study of the relation between subway traffic and the unique surge and subsequent decline of coronavirus infections in New York City during the month of March 2020 (Harris 2020b). And in other cases, such as our analysis of the data from Wisconsin here, one may be able to isolate an outbreak at the micro level as the underlying basis of a statewide trend. But in other cases, we may end up with perfect data on trends in incidence but no way to clearly disentangle their causes.

The underlying purpose of universal testing is not simply to monitor trends, which we might very well be able to do with representative testing or testing of pooled samples, but also to track contacts and isolate those affected. When one thinks about the hundreds of contacts just one subway rider would make on a single trip into work, such a task seems all the more daunting.

Appendix: SIR Models

We follow the basic notation in (Harris 2020a) with slight modifications. The time course of an epidemic can be described by a set of coupled differential equations. Let $S(t)$ denote the number of susceptible individuals, $I(t)$ denote the number of infective individuals, and $R(t)$ denote the number of resistant individuals at time $t \geq 0$. All individuals in the population are in one of these three states. In the basic version of the model, which we adopt here, the population is assumed closed. Without loss of generality, we specify $S(t) + I(t) + R(t) = 1$, so that each of the state variables is expressed as a fraction of the total population. Thus, $\dot{S} + \dot{I} + \dot{R} = 0$, where we have used the notation $\dot{S} = \frac{dS(t)}{dt}$ for the first derivative.

In the unrestricted version of the SIR model, the rate of new infections per unit time is assumed to be proportional to the number of interactions between susceptible and infective individuals, that is, $-\dot{S} = \alpha SI$, where $\alpha > 0$ is a constant. Once an individual is infected, he is infective and remains infective until he becomes resistant, either through recovery or death. Infective individuals are assumed to become resistant at a constant proportional rate, that is, $\dot{R} = \beta I$, where $\beta > 0$ is also a constant. Since our population is closed, we have

$\dot{I} = -\dot{S} - \dot{R} = \alpha SI - \beta I$.

To fit this unrestricted model to the data from San Antonio in Figure 13, we converted the basic differential equations into difference equations in discrete time format. That is,
What To Watch For  

Jeffrey E. Harris  

12-May-2020

The black curve in Figure 13 is based upon the parameters \( I_0 = \frac{1}{N} \), where the population \( N \) of San Antonio is taken as 1,493,000. That is, at the start of the epidemic, one person is infected. In addition, we take \( \alpha = 0.162 \) and \( \beta = 0.05 \). As explained in (Harris 2020a), these parameters give a basic reproductive number at the start of the epidemic equal to \( R_0 = \frac{\alpha S_0}{\beta} = 3.24 \). The value of \( \beta = 0.05 \) implies a mean duration of infectivity of 20 days, considerably larger than some estimates, but more in line with other data (Wölfel et al. 2020, Xiao, Tong, and Zhang 2020).

In the social distancing variant of the SIR model, we add another state variable, denoted by \( D \), which represents those susceptible individuals who do not come into contact with infective individuals. This is a feature of many variant SIR models in the literature (Greenhalgh 1988, Yi et al. 2009, Hansen and Day 2011). Initially, the model starts as the unrestricted model. Then, starting at time \( t = t_0 \), the differential equation governing susceptible individuals switches to \( \dot{S} = -\alpha SI - \gamma S \), while the corresponding equation for individuals who avoid contact is \( \dot{D} = \gamma S \), where \( \gamma > 0 \). Once again simulating the discrete version of this model, we obtain the black curve in Figure 13 with \( \gamma = 0.03 \) and \( t_0 \) corresponding to March 30, 2020. From that date onward, we’re assuming that the population of self-isolating individuals is growing at 3 percent daily.

In the SIR variant with released restrictions, we simulate the social distance variant, but then modify the parameters at \( t_i \) corresponding to April 27, 2020. From that point onward, \( \dot{D} = -\delta D \), where \( \delta = 0.09 \) and \( \dot{S} = -\alpha SI + \delta D \). That is, we’re assuming that population of self-isolating individuals is declining at 9 percent daily.

References


Baxter, Kevin, Alex Wigglesworth, and Cindy Chang. 2020. *Despite coronavirus, many flock to Southern California beaches to beat the heat.*


Kaplow, Louis. 2020. If We Can’t Test Everyone for Coronavirus, This Is the Next Best Thing. 


Montesano, M. 2020a. *case-hosp-death.csv (Comma Separated Value file).*

Montesano, M. 2020b. *probable-confirmed-dod.csv (Comma Separated Value file).*

Mukai, Rachel. 2020. *COVID-19 Cases by Date: Positive Labs by Date.*

Murphy, Philip D. 2020. *Executive Order No. 133.*


