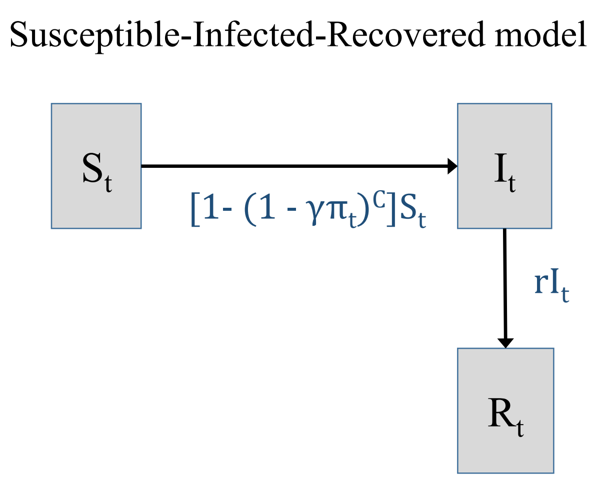
CHAPTER 15: PREDICTING COVID

Lockdown 2020

As of March 21, 2020, only about 500 people had died from COVID[1] (Admittedly, attributing deaths to COVID is a tricky business) in the US. Over 58,000 people had died from other causes since January 1, 2020. Yet, around that time, many governors took the drastic step of ordering "lockdowns", closing non-essential businesses and placing restrictions on essential ones. Predictions from models of infectious disease transmission played a key role in the response. The situation wasn't bad in mid-March, but models showed that in the absence of a lockdown, hospitals would be overwhelmed and millions would die. Since then, models have been used to address a variety of questions about re-opening, whether to allow in-person school and college, masks, and testing.

Example of an SIR model

The Susceptible-Infectious-Recovered model is a workhorse of infectious disease epidemiology. The model follows the population over time. Time periods can represent days, weeks, or months, depending on the disease in question. In the model, the population can be divided into three groups: people who are at risk of getting infected (susceptible), people who are infected, and people who have recovered from infection and are immune. S_t, I_t, and R_t represent the number of people in each state at time t.



[1] Centers for Disease Control and Prevention. Daily Updates of Totals by Week and State. Provisional Death Counts for Coronavirus Disease 2019 (COVID-19) https://www.cdc.gov/nchs/nvss/vsrr/covid19/index.htm

Note that this one of many possible forms of the model.

The total population size is

$$S_t + I_t + R_t \qquad [1]$$

The prevalence of infection at time t is

$$\pi_t = \frac{I_t}{S_t+I_t+R_t} \qquad [2]$$

The probability that an uninfected person becomes infected at time t is given by this gnarly equation:

$$\theta_t = 1 - (1 - \gamma\pi_t)^C, \qquad [3]$$

where γ is the probability that you get sick if you come in contact with an infected person (the transmission rate) and C is the number of people you come in contact with.

The movement of people between states over time is described by a series of equations. Usually, these are written as differential equations, but we will write them in a simpler form.

$$S_{t+1} = S_t - \theta_t S_t \qquad [4.1]$$

$$I_{t+1} = I_t + \theta_t S_t - rI_t \qquad [4.2]$$

$$R_{t+1} = R_t + rI_t \qquad [4.3]$$

The last term in 4.2 and 4.3 represents recovery, where r is the recovery rate, i.e. the share of infected people who recover every period.

Let's unpack equation 3. Suppose you come in contact with a random person. What is the probability you get infected? It is the probability your contact is infected multiplied by the probability that you catch the disease from an infected person: $\gamma\pi_t$.

Now suppose you come in contact with C people? What is the probability that you get infected? I learned to think about probabilities of this sort using the example of torpedoes. Suppose there are three torpedoes heading for the boat you are on. The boat will sink if any one of the torpedoes hits it. The probability that any one of the torpedoes hits the boat is 0.25 (or 25%). What is the probability that the boat sinks? It is the probability that any one of the torpedoes hits. There are multiple combinations (e.g., the first torpedo hits the boat but not the other two, the first two hit but not the third, the probability that the first and third miss but the second hits). To calculate the probability that at least one hits, you could calculate the probabilities of all these combinations and add them together. Or, you could just calculate one minus the probability that none hit.

Here's how to do it. The probability that one torpedo doesn't hit is 1 – 0.25. The probability that three independent events occur (i.e., three torpedoes not hitting) is the probabilities multiplied together. So the probability that none of the torpedoes hit is

$$(1 - 0.25) \times (1 - 0.25) \times (1 - 0.25) = (1 - 0.25)^3 = 0.422.$$

The probability that at least one of the torpedoes hits is one minus the probability that none of them hit:

$$0.578 = 1 - (1 - 0.25)^3,$$

Better jump off that boat.

In the case of COVID, each contact is like a torpedo. You have to calculate the probability that an uninfected person becomes infected from any of his C contacts.

I simulated the model using the following assumptions:

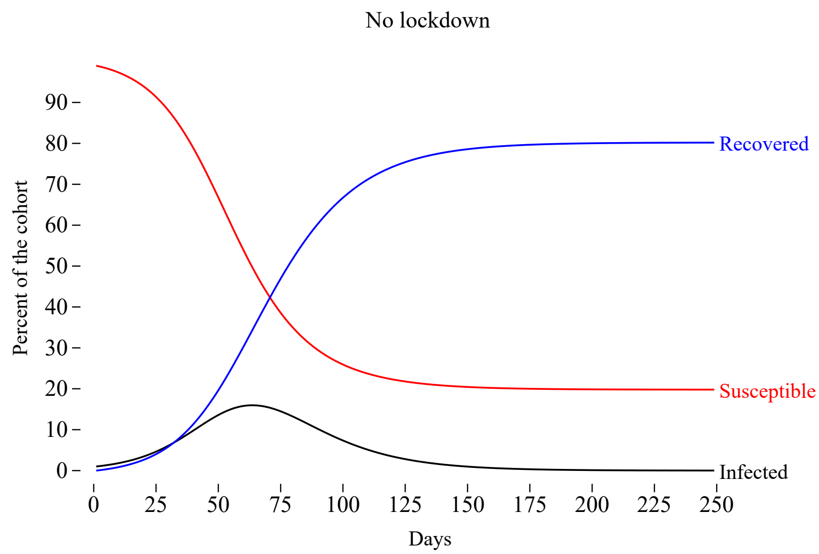
$r = 1/14$ (i.e., the average duration of infection is 14 days)
$C = 2.14$ (contacts per day)
$\gamma = 0.0666$ (transmission rate)

I selected C and γ so that an infected person infects two other people:

$2.14 \times 0.0666 \times 14$ (days) $= 2$.

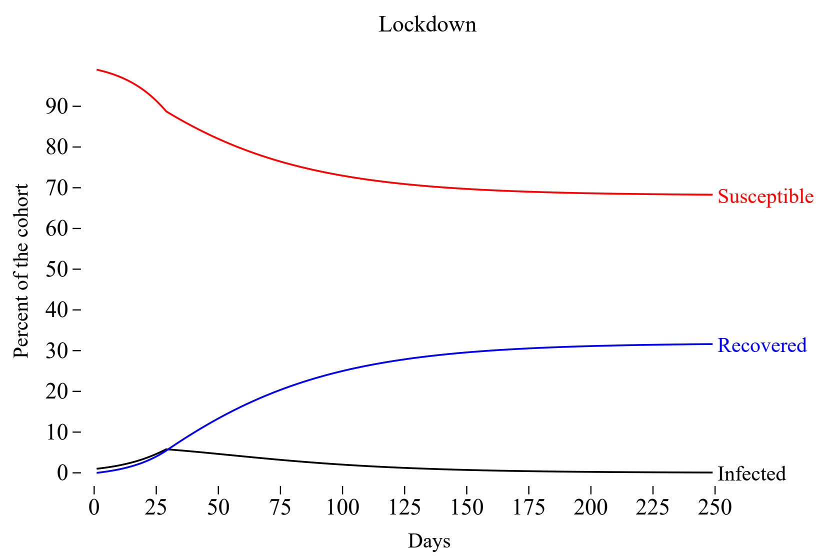
Two was a reasonable assumption of the R_0 value for COVID (although some recent estimates place the value higher).

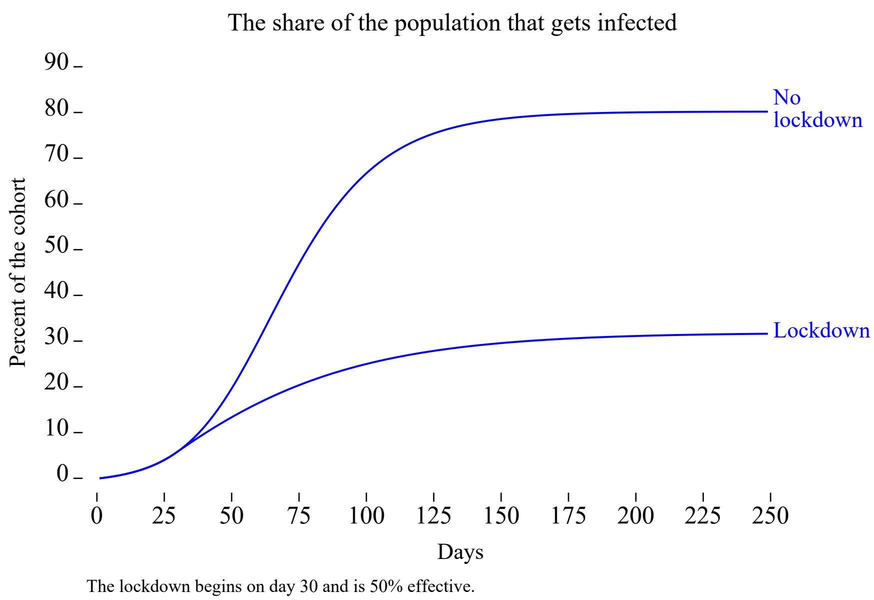
The graph below shows the share of the population in each of the three states over time.



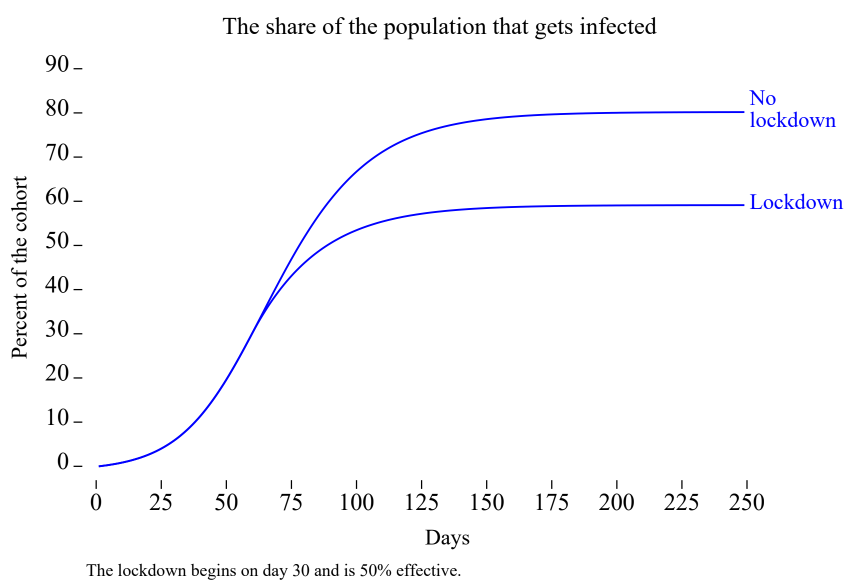
In this case, herd immunity is reached once 80% of the population has become infected. Herd immunity occurs when enough people are immune that the infection no longer propagates itself through the community.

How would a "lockdown" affect outcomes? We can model a lockdown by reducing the number of contacts, C. Lockdowns aren't perfectly effective. Essential workers still have contacts with co-workers and customers, people have contacts with others in their households, and some people are non-compliant. Let's assume a lockdown is 50% effective, meaning that under a lockdown, the number of contacts is cut in half. In this case, a lockdown imposed on day 30 means that only 32% of the population gets infected.

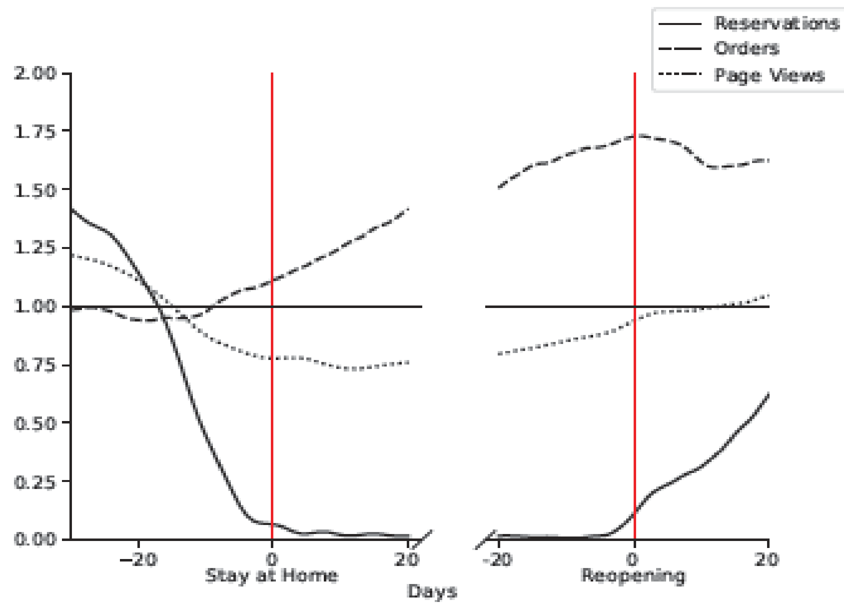




The timing of the lockdown is important. Here is what happens if the lockdown begins on day 60 instead of 30.



Lockdowns are, fortunately, rare. But that makes it difficult to evaluate their impact. Data from Yelp[2] shows that the number of restaurant reservations plummeted before state lockdowns went into effect. This result suggests that lockdowns don't have much of an effect over and above the precautions that people take on their own. But trends around the removal of lockdowns suggest that they affect behavior. Additional lockdowns may be less effective than the initial lockdown as the shock of the pandemic wares off.



Yelp Outcomes

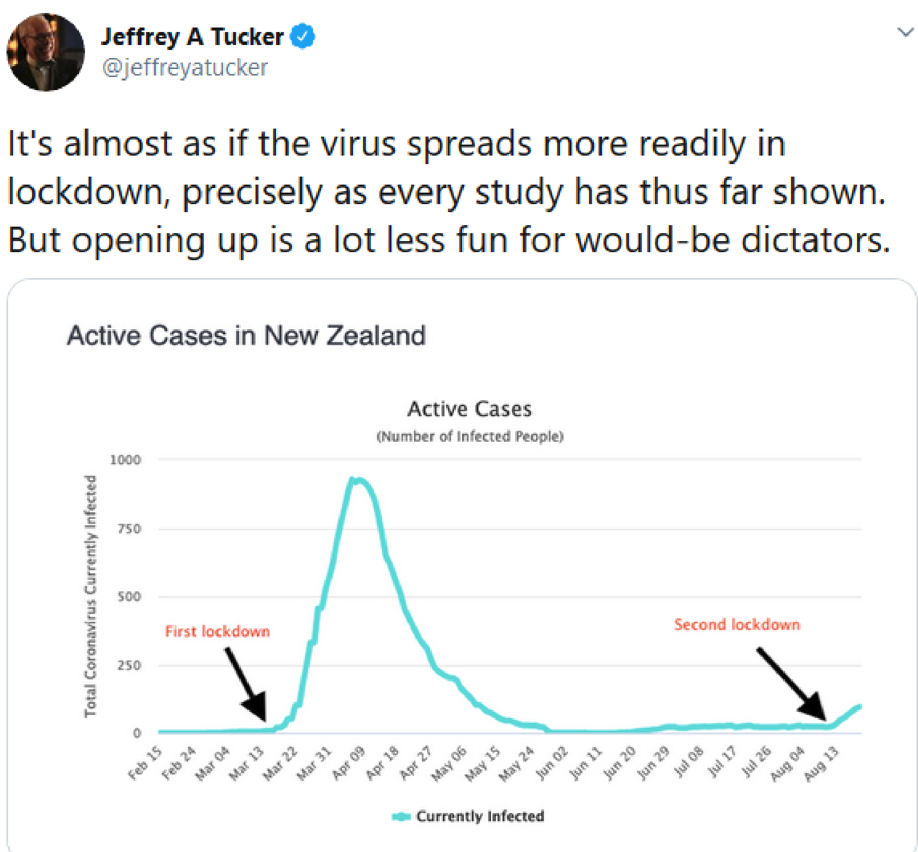
[2] Glaeser EL, G Zhe Jin, BT Leyden, M Luca. Learning from deregulation: the asymmetric impact of lockdown and reopening on risky behavior during COVID-19. National Bureau of Economic Research working paper 27650.

An alternative to SIR models

One approach (SIR models) is to develop a mathematical model to represent transmission of an infectious agent based on the number of contacts between people and the probability that the pathogen is transmitted during a contact between two people. The growth (or decline) of an epidemic depends on the underlying assumptions about contacts, transmissibility, and the duration of time during which people are infectious (i.e., able to transmit the pathogen).

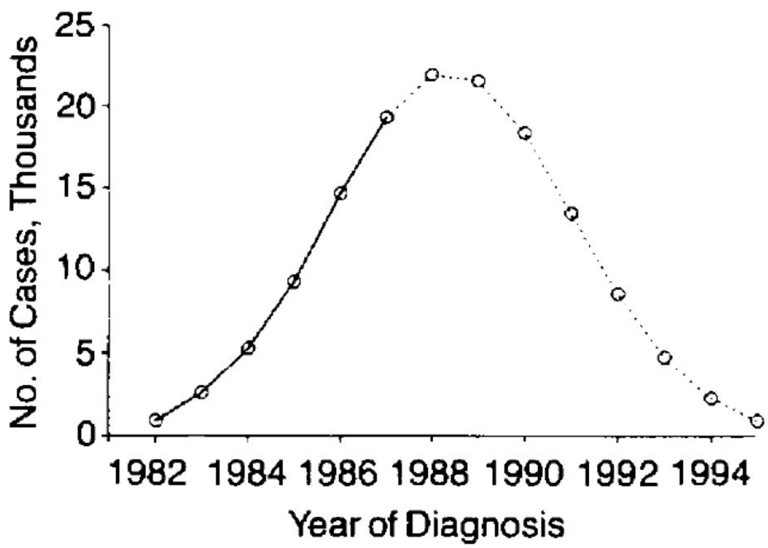
A second approach is to project the growth or decline based on historic data about the path of the epidemic in other regions. For example, we know that epidemics usually follow a bell-shaped pattern. So if we know the peak of the bell shape in another region, and have some data about the left side of the bell shape in our own region, we can project the trajectory going forward. It doesn't require us to make any assumptions about transmission rates, contacts, etc.

Can you guess the problem with the analysis below???



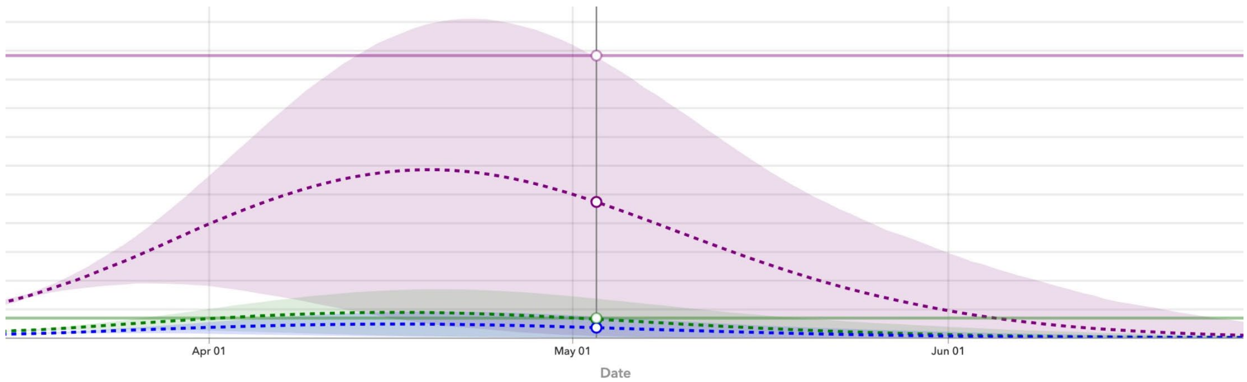
The University of Washington's Institute for Health Metrics and Evaluation developed a model for predicting COVID outcomes that took a different approach from the S-I-R models that dominate much of infectious disease modeling. (Reportedly the model now incorporates some elements of the S-I-R model.)

The model is based on the observation, known as Farr's law, that the share of the population infected with a novel pathogen at a point in time looks like a bell curve or normal distribution. Here is a graph from a study that projected the incidence of AIDS.[3]



[3] Bregman DJ, Langmuir AD. Farr's Law Applied to AIDS Projections. *Journal of the American Medical Association* 1990;263(11):1522–1525.

The early versions of the IHME model fit a normal curve based on early infection data and assumptions about the trajectory of the disease based on data from China and Italy.



IHME model projections[4]
Pink line: All beds. Green line: ICU beds. Blue line: ventilators.

Despite being panned by many infectious disease experts, the IHME model was widely quoted in the media and Trump administration officials in the pandemic's early days. Here is how an article in *Vox*[5] explained the IHME model's influence.

> Some of the factors that make the IHME model unreliable at predicting the virus may have gotten people to pay attention to it. For one thing, it's more simplistic compared to other models. That means it can be applied in ways more complicated models could not, such as providing state-level projections (something state officials really wanted), which other modelers acknowledged that they didn't have enough data to offer.
>
> Meanwhile, its narrow confidence intervals for state-by-state estimates meant it had quotable (and optimistic) topline numbers. A confidence interval represents a range of numbers wherein the model is very confident the true value will lie. A narrow range that gives "an appearance of certainty is seductive when the world is desperate to know what lies ahead," a criticism of the IHME model published in the Annals of Internal Medicine argued. But the numbers and precise curves the IHME is publishing "suggests greater precision than the model is able to offer."

An article in the *Annals of Internal Medicine*[6] described the IHME model as a "statistical model with no epidemiologic basis." It helped that the IHME model gave optimistic forecasts compared to other models (for example, at one point the model projected that there would

[4] Hospital resource use.
https://pbs.twimg.com/media/EULx0s0UwAAGoqX?format=jpg&name=4096x4096
[5] Piper K. This coronavirus model keeps being wrong. Why are we still listening to it? *Vox*. May 2, 2020.

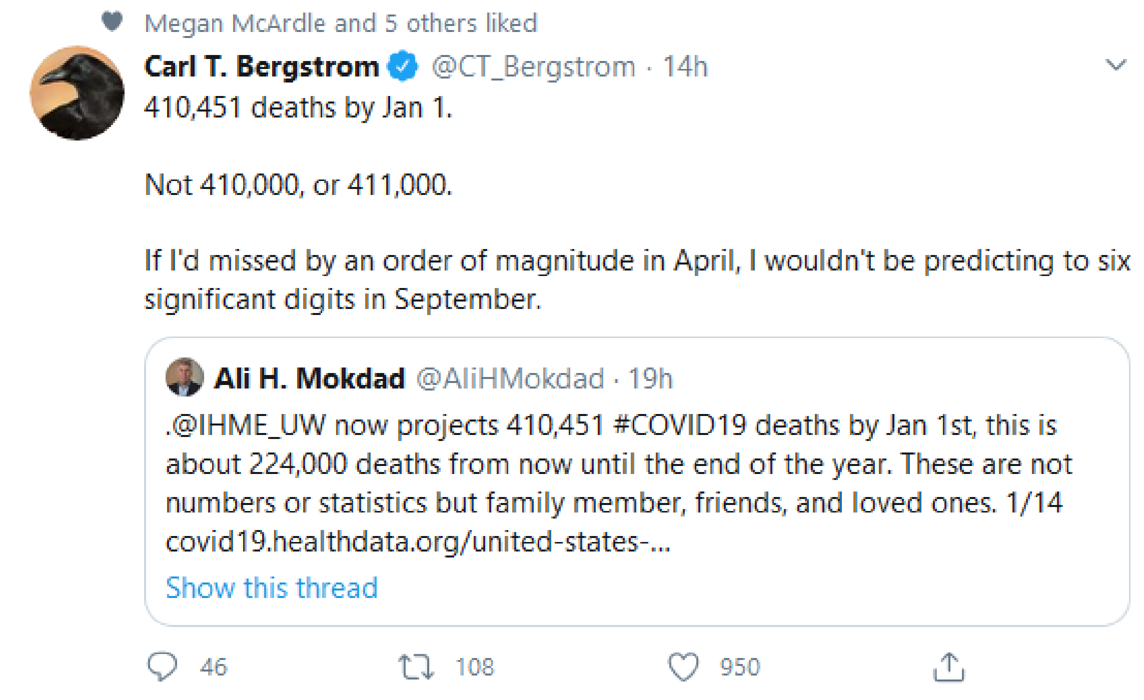
[6] Jewell NP, Lewnard JA, Jewell BL. Caution Warranted: Using the Institute for Health Metrics and Evaluation Model for Predicting the Course of the COVID-19 Pandemic. *Annals of Internal Medicine* 2020;173(3):226-227.

be only 60,000 deaths), which allowed the Trump administration to claim that Democrats were overreacting to COVID.

The IHME's director argued that most other models do not report results that are useful to policymakers[7]: "We're willing to make a forecast. Most academics want to hedge their bets and not be found to ever be wrong…That's not useful for a planner — you can't go to a hospital and say you might need 1,000 ventilators, or you might need 5,000."

To be fair, IHME was not the only modeling outfit making incorrect projections. At the other end of the spectrum, Michael Osterholm, director of the University of Minnesota's Center for Infectious Disease Research and Policy projected on the Joe Rogan podcast[8] that "conservatively", COVID will result in 480,000 deaths and 48,000,000 hospitalizations (the actual number, as of September 4, 2020, was less than 2 million.).

[7] Cancryn A. How overly optimistic modeling distorted Trump team's coronavirus response. *Politico* April 24, 2020.
[8] DeRensis, Hunter. Scientist: 480,000 People Could Die Due to Coronavirus, 48 Million Hospitalizations. *The National Interest* March 12, 2020.



Megan McArdle and 5 others liked

Carl T. Bergstrom @CT_Bergstrom · 14h

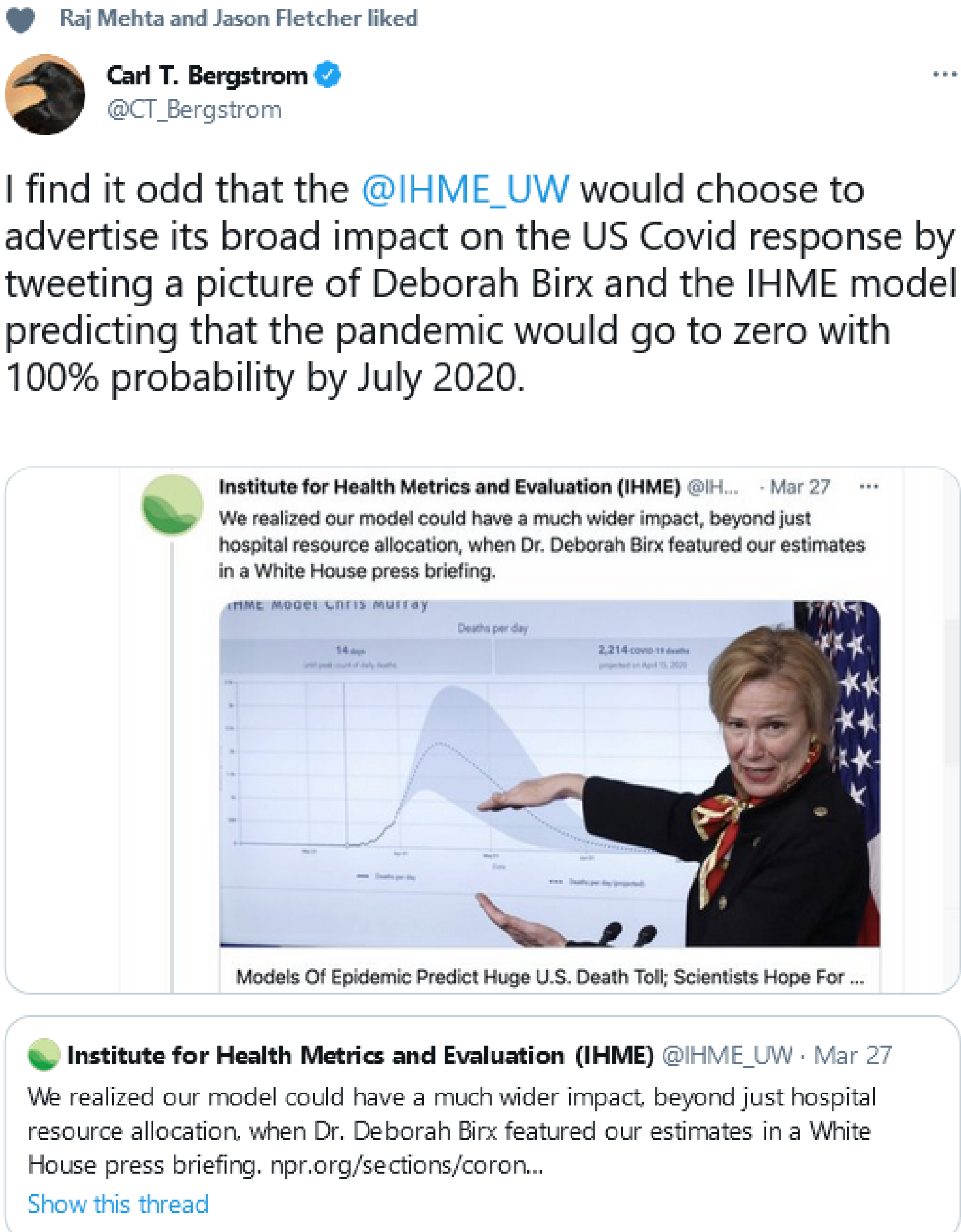
410,451 deaths by Jan 1.

Not 410,000, or 411,000.

If I'd missed by an order of magnitude in April, I wouldn't be predicting to six significant digits in September.

> **Ali H. Mokdad** @AliHMokdad · 19h
>
> .@IHME_UW now projects 410,451 #COVID19 deaths by Jan 1st, this is about 224,000 deaths from now until the end of the year. These are not numbers or statistics but family member, friends, and loved ones. 1/14 covid19.healthdata.org/united-states-...
>
> Show this thread

46 108 950



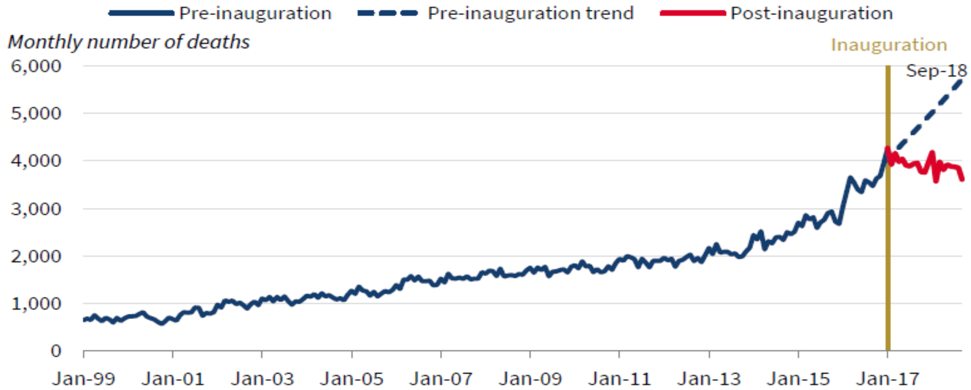
Here is another example of a projection based off of past trends.[9] Do you think this is a good analysis? Why or why not?

[9] The Role of Opioid Prices in the Evolving Opioid Crisis. The Council of Economic Advisors. April 2019.

Box 1. The Slowing of the Opioid Epidemic under the Trump Administration

When President Trump took office in January 2017, monthly overdose deaths involving opioids had reached an all-time recorded high, a 21 percent increase from the average number of monthly deaths in 2016. The total number of opioid deaths grew again in 2017 (47,600) compared to the previous year (42,249 in 2016). Fortunately, the growth in opioid deaths may have finally begun to reverse. The rising overdose death toll through September 2018, the latest month for which provisional data is available, has flattened compared to previous trends (see figure i). The opioid epidemic remains at crisis levels, but the dramatic growth of the epidemic seems to be slowing down.

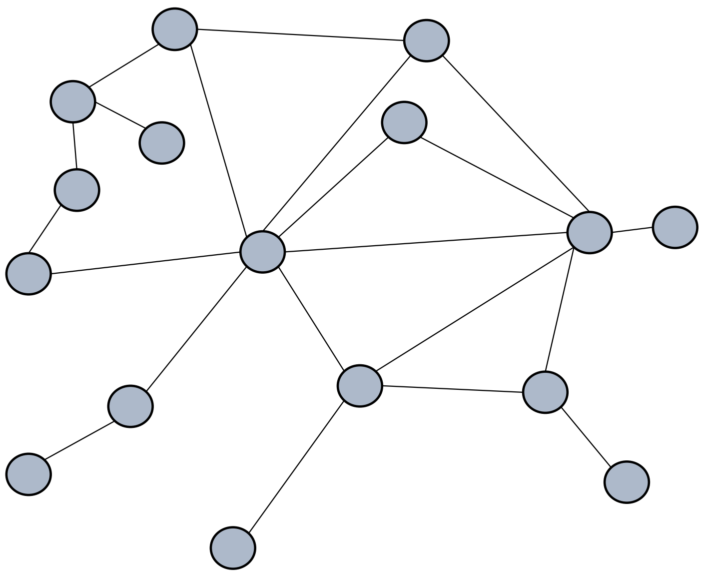
Figure i. Monthly Opioid-Involved Overdose Deaths, 1999–2018



Sources: Centers for Disease Control and Prevention; Ahmad et al. (2019); CEA calculations.

Note: Data from before January 2018 are compiled from the CDC WONDER database, and monthly data beginning in January 2018 are calculated using the provisional reported number of deaths from the CDC which is available through September 2018 as of April 25th, 2019. Pre-inauguration trend is calculated for the compound annual growth rate on a sample period from January 1999 through January 2017, with forecasted levels reconstructed from projected rates.

Network models



Additional resources

Tufekci, Zeynep. Don't Believe the COVID-19 Models. The Atlantic April 2, 2020.

SEIRS+ Model Framework with nice diagrams: https://github.com/ryansmcgee/seirsplus

A visual simulation of network models

https://www.washingtonpost.com/graphics/2020/health/coronavirus-herd-immunity-simulation-vaccine/?hpid=hp_hp-banner-main_herdimmunity-1155am%3Ahomepage%2Fstory-ans&itid=hp_hp-banner-main_herdimmunity-1155am%3Ahomepage%2Fstory-ans

And Why outbreaks like coronavirus spread exponentially, and how to "flatten the curve" - Washington Post

A compilation of predictions from various COVID models
https://projects.fivethirtyeight.com/covid-forecasts/

Podcasts on COVID models:

Ellie Murray and Lucy D'Agostino McGowan Coronavirus Conversations 2 | Episode 10
https://casualinfer.libsyn.com/coronavirus-conversations-2

Sensationalist Science: https://twitter.com/SenSciPod/status/1286779649813385216