Diverse Local Epidemics Reveal the Distinct Effects of Population Density, Demographics, Climate, Depletion of Susceptibles, and Intervention in the First Wave of COVID-19 in the United States

Niayesh Afshordi, Benjamin P. Holder, Mohammad Bahrami, and Daniel Lichtblau

1 Introduction

The new human coronavirus SARS-CoV-2 first reported in Wuhan Province, China in December 2019 [1, 2], reached 10,000 confirmed cases and 200 deaths due to the disease (known as COVID-19) by the end of January 2020. Although travel from China was halted by late-January, dozens of known introductions of the virus to North America occurred prior to that [3, 4], and dozens more known cases were imported to the US and Canada during February from Europe, the Middle East, and elsewhere. Community transmission of unknown origin was first detected in California on February 26, followed quickly by Washington State [5], Illinois and Florida, but only on March 7 in New York City. Retrospective genomic analyses have demonstrated that case-tracing and self-quarantine efforts were effective in preventing most known imported cases from propagating [6–8], but that the eventual outbreaks on the West Coast [5, 8, 9] and New York [7] were likely seeded by

N. Afshordi (✉)
Department of Physics and Astronomy, University of Waterloo, Waterloo, ON, Canada

Waterloo Centre for Astrophysics, University of Waterloo, Waterloo, ON, Canada

Perimeter Institute of Theoretical Physics, Waterloo, ON, Canada
e-mail: nafshordi@pitp.ca; nafshordi@perimeterinstitute.ca

B. P. Holder
Department of Physics and Astronomy, University of Waterloo, Waterloo, ON, Canada

Department of Physics, Grand Valley State University, Grand Rapids, MI, USA

M. Bahrami · D. Lichtblau
Wolfram Research Inc., Champaign, IL, USA

V. K. Murty, J. Wu (eds.), *Mathematics of Public Health*, Fields Institute Communications 85, https://doi.org/10.1007/978-3-030-85053-1_1

unknown imports in mid-February. By early March, cross-country spread was primarily due to interstate travel rather than international imports [10].

In mid-March 2020, nearly every region of the United States of America saw a period of uniform exponential growth in daily confirmed cases—signifying robust community transmission—followed by a plateau in late March, likely due to social mobility reduction. The same qualitative dynamics were seen in COVID-19 mortality counts, delayed by approximately 1 week. Although the qualitative picture was similar across locales, the quantitative aspects of localized epidemics—including initial rate of growth, infections/deaths per capita, duration of plateau, and rapidity of resolution—were quite diverse across the country. Understanding the origins of this diversity will be key to predicting how the relaxation of social distancing, annual changes in weather, and static local demographic/population characteristics will affect the resolution of the first wave of cases, and will drive coming waves, prior to the availability of a vaccine.

The exponential growth rate of a spreading epidemic is dependent on the biological features of the virus-host ecosystem—including the incubation time, susceptibility of target cells to infection, and persistence of the virus particle outside of the host—but, through its dependence on the transmission rate between hosts, it is also a function of external factors such as population density, air humidity, and the fraction of hosts that are susceptible. Initial studies have shown that SARS-CoV-2 has a larger rate of exponential growth (or, alternatively, a lower doubling time of cases)[1] than many other circulating human viruses [11]. For comparison, the pandemic influenza of 2009, which also met a largely immunologically-naive population, had a doubling time of 5–10 d [12, 13], while that of SARS-CoV-2 has been estimated at 2–5 d [14, 15] (growth rates of $\sim 0.10\ \mathrm{d}^{-1}$ vs. $\sim 0.25\ \mathrm{d}^{-1}$). It is not yet understood which factors contribute to this high level of infectiousness.

While the dynamics of an epidemic (e.g., cases over time) must be described by numerical solutions to nonlinear models, the exponential growth rate, λ, usually has a simpler dependence on external factors. Unlike case or mortality incidence numbers, the growth rate does not scale with population size. It is a directly measurable quantity from the available incidence data, unlike, e.g., the reproduction number, which requires knowledge of the serial interval distribution [16–18], something that is difficult to determine empirically [19, 20]. Yet, the growth rate contains the same threshold as the reproduction number ($\lambda = 0$ vs. $R_0 = 1$), between a spreading epidemic (or an unstable uninfected equilibrium) and a contracting one (or an equilibrium that is resistant to flare-ups). Thus, the growth rate is an informative direct measure on that space of underlying parameters.

In this work, we leverage the enormous data set of epidemics across the United States to evaluate the impact of demographics, population density and structure, weather, and non-pharmaceutical interventions (i.e., mobility restrictions) on the exponential rate of growth of COVID-19. Following a brief analysis of the initial spread in metropolitan regions, we expand the meaning of the exponential rate

[1] The doubling time is $\ln 2$ divided by the exponential growth rate.

to encompass all aspects of a local epidemic—including growth, plateau and decline—and use it as a tracer of the dynamics, where its time dependence and geographic variation are dictated solely by these external variables and per capita cumulative mortality. Finally, we use the results of that linear analysis to calibrate a new nonlinear model—a renewal equation that utilizes the excursion probability of a random walk to determine the incubation period—from which we develop local predictions about the impact of social mobility relaxation, the level of herd immunity, and the potential of rebound epidemics in the Summer and Fall of 2020. The methodology can be modified to make local predictions as the pandemic evolves.

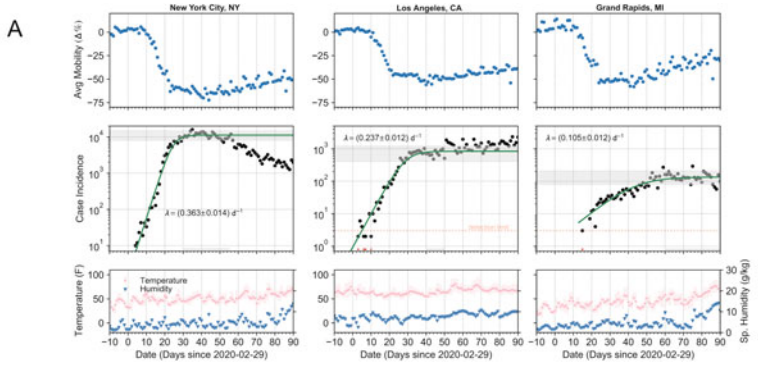
2 Results

2.1 Initial Growth of Cases in Metropolitan Regions Is Exponential with Rate Depending on Mobility, Population, Demographics, and Humidity

As an initial look at COVID-19's arrival in the United States, we considered the ∼100 most populous metropolitan regions—using maps of population density to select compact sets of counties representing each region (see [21])—and estimated the initial exponential growth rate of cases in each region. We performed a linear regression to a large set of demographic (sex, age, race) and population variables, along with weather and social mobility [22] preceding the period of growth (Fig. 1). In the best fit model ($R^2 = 0.75$, BIC $= -183$), the baseline value of the initial growth rate was $\lambda = 0.21\ \mathrm{d}^{-1}$ (doubling time of 3.3 d), with average mobility 2 weeks prior to growth being the most significant factor (Fig. 1b). Of all variables considered, only four others were significant: population density (including both *population-weighted density* (PWD)—also called the "lived population density" because it estimates the density for the average individual [23]—and *population sparsity*, γ, a measure of the difference between PWD and standard population density, see Methods), $p < 0.001$ and $p = 0.006$; specific humidity 2 weeks prior to growth, $p = 0.001$; and median age, $p = 0.04$.

While mobility reduction certainly caused the "flattening" of case incidence in every region by late-March, our results show (Fig. 1c) that it likely played a key role in reducing the *rate* of growth in Boston, Washington, DC, and Los Angeles, but was too late, with respect to the sudden appearance of the epidemic, to have such an effect in, e.g., Detroit and Cleveland. In the most extreme example, Grand Rapids, MI, seems to have benefited from a late arriving epidemic, such that its growth (with a long doubling time of 7 d) occurred almost entirely post-lockdown.

Specific humidity, a measure of absolute humidity, has been previously shown to be inversely correlated with respiratory virus transmission [24–27]. Here, we found it to be a significant factor, but weaker than population density and mobility (Fig. 1c). It could be argued that Dallas, Los Angeles, and Atlanta saw a small



B

Variable		Mean value	Coef.	Std. Err.	t	95% CI
average growth rate, λ_{avg}			0.232	0.005	43.1	[0.220, 0.242]
avg mobility, 2wk prior	–	+2.2%	0.0063	0.0006	10.8	[0.0051,0.0074]
ln[PWD]	–	8.1	0.048	0.010	4.7	[0.028,0.069]
sp. humidity, 2wk prior	–	4.9 g/kg	-0.010	0.0027	-3.7	[-0.016, -0.005]
pop. sparsity, γ	–	0.21	-0.33	0.11	-3.0	[-0.55, -0.11]
median age	–	35.7	0.0049	0.0024	2.1	[0.0002,0.0096]

C

	med. age	mob. 2wk pr.	pop. sparsity	hum. 2wk pr.	ln[pwpd]	λ_{exp}	λ_{exp}	λ_{exp} err
ny_newyorkcity	0.004	0.011	0.007	0.021	0.026	0.343	0.363	0.014
il_chicago	-0.002	0.022	0.02	0.018	0.017	0.307	0.316	0.017
mi_detroit	0.008	0.044	0.019	0.02	-0.023	0.301	0.412	0.035
oh_cleveland	0.014	0.035	0.02	0.018	-0.024	0.294	0.298	0.027
pa_philadelphia	0.004	0.01	0.015	0.014	0.012	0.286	0.291	0.009
ma_boston	0.003	-0.03	0.015	0.021	0.017	0.258	0.222	0.013
ct_hartford	0.012	-0.003	0.01	0.019	-0.014	0.255	0.287	0.019
co_denver	-0.009	0.022	-0.018	0.027	-0.002	0.252	0.24	0.014
wa_seattle	-0.003	0.02	0.002	0.007	-0.006	0.252	0.223	0.019
ma_worcester	0.009	0.003	0.001	0.024	-0.02	0.248	0.237	0.026
tx_dallas	-0.015	0.027	0.011	-0.016	0.003	0.24	0.252	0.021
ga_atlanta	-0.008	0.022	0.025	-0.006	-0.023	0.24	0.242	0.019
ca_losangeles	-0.007	0.0	-0.006	-0.016	0.027	0.229	0.237	0.012
dc_washington	-0.002	-0.047	0.008	0.003	0.012	0.206	0.213	0.009
mi_grandrapids	-0.006	-0.12	-0.001	0.015	-0.038	0.083	0.105	0.012

Fig. 1 Mobility and COVID-19 incidence data examples, and the results of linear regression to extracted initial exponential growth rates, λ_{exp}, in the top 100 metropolitan regions. (**a**) Three example cities with different initial growth rates. Data for Google mobility (blue points), daily reported cases (black points), and weather (red and blue points, bottom) are shown with a logistic fit to cases (green line). Data at or below detection limit were excluded from fits (dates marked by red points). Thin grey bars at base of cases graphs indicate region considered "flat", with right end indicating the last point used for logistic fitting; averaging over "flat" values generates the thick grey bars to guide the eye. [See Supp. Matt. in [21] for additional information and for complete data sets for all metropolitan regions.] (**b**) Weighted linear regression results in fit to λ_{exp} for all metropolitan regions. (**c**) Effect of each variable on growth rate (i.e., $\Delta\lambda$ values) for those regions with well-estimated case and death rates; white/yellow indicates a negative effect on λ, red indicates positive

benefit from higher humidity at the time of the epidemic's arrival, while the dry late-winter conditions in the Midwest and Northeast were more favorable to rapid transmission of SARS-CoV-2.

2.2 Exponential Growth Rate of Mortality as a Dynamical, Pan-Epidemic, Measure

In the remainder of this report, we consider the exponential rate of growth (or decay) in local confirmed deaths due to COVID-19. The statistics of mortality is poorer compared to reported cases, but it is much less dependent on unknown factors such as the criteria for testing, local policies, test kit availability, and asymptomatic individuals [28]. Although there is clear evidence that a large fraction of COVID-19 mortality is missed in the official counts (e.g., [29, 30]), mortality is likely less

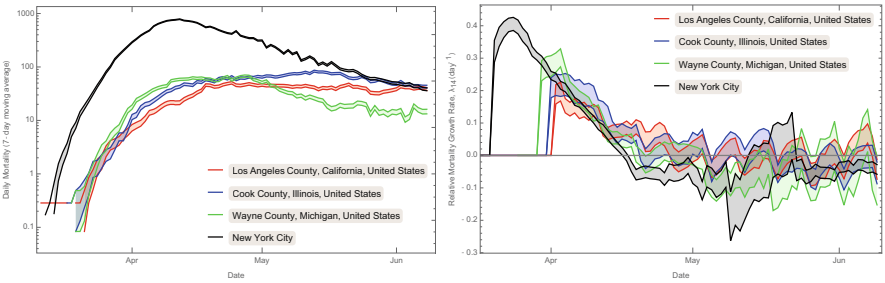


Fig. 2 COVID-19 mortality incidence (7-day rolling average, left) and exponential growth rate (λ_{14}, determined by regression of the logged mortality data over 14-day windows, right) for the four US counties with >2400 confirmed COVID-19 reported deaths (as of 8th June, 2020)

susceptible to rapid changes in reporting, and, as long as the number of reported deaths is a monotonic function of the actual number of deaths (e.g., a constant fraction, say 50%), the sign of the exponential growth rate will be unchanged, which is the crucial measure of the success in pandemic management.

To minimize the impact of weekly changes, such as weekend reporting lulls, data dumps, and mobility changes from working days to weekends, we calculate the regression of $\ln\left[\text{Mortality}\right]$ over a 14-day interval, and assign this value, $\lambda_{14}(t)$, and its standard error to the last day of the interval. Since only the data for distinct 2-week periods are independent, we multiply the regression errors by $\sqrt{14}$ to account for correlations between the daily estimates. Together with a "rolling average" of the mortality, this time-dependent measure of the exponential growth rate provides, at any day, the most up-to-date information on the progression of the epidemic (Fig. 2).

In the following section, we consider a linear fit to λ_{14}, to determine the statistically-significant external (non-biological) factors influencing the dynamics of local exponential growth and decline of the epidemic. We then develop a first-principles model for λ_{14} that allows for extrapolation of these dependencies to predict the impact of future changes in social mobility and climate.

2.3 Epidemic Mortality Data Explained by Mobility, Population, Demographics, Depletion of Susceptible Population and Weather, Throughout the First Wave of COVID-19

We considered a spatio-temporal dataset containing 3933 estimates of the exponential growth measure, λ_{14}, covering the 3 month period of 8 March 2020–8 June 2020 in the 187 US counties for which information on COVID-19 mortality and all potential driving factors, below, were available (the main barrier was social mobility information, which limited us to a set of counties that included 69% of

Table 1 Joint linear fit to $\lambda_{14}(t)$ data (Top). Any dependence with t-statistic below 2.5σ is considered not statistically significant. Joint linear fit to $\lambda_{14}(t)$, including only statistically significant dependencies (Bottom). For all coefficients, the population-weighted baseline is subtracted from the linear variable

	Estimate	Std. err	t-Statistic
Joint fit to all potential drivers			
Baseline mortality growth rate λ_{14}	0.195	0.011	17.2
COVID death fraction	−59.4	6.1	−9.7
Social mobility (2wks prior)	0.00238	0.00028	8.5
ln(population weighted density) 8.24	0.0412	0.0058	7.1
Social mobility (4wks prior)	0.00122	0.00019	6.6
Population sparsity 0.188	−0.249	0.063	−3.9
log(annual death) 4.04	−0.0301	0.0091	−3.3
Median age 37.47	0.0038	0.0012	3.0
People per household 2.76	0.023	0.014	1.6
Specific humidity (2wks prior) 5.92 g/kg	−0.0033	0.0031	−1.1
Temperature (2wks prior) 13.11 C	−0.00083	0.0013	−0.6
Temperature (4wks prior) 11.60 C	−0.00060	0.0014	−0.4
Specific humidity (4wks prior) 5.53 g/kg	0.00058	0.0032	0.2
Joint fit to statistically significant drivers			
Baseline mortality growth rate λ_{14}	0.198	0.011	18.7
COVID death fraction	−56.7	5.9	−9.7
Social mobility (2wks prior)	0.00236	0.00027	8.8
Social mobility (4wks prior)	0.00131	0.00017	7.6
ln(population weighted density) 8.24	0.0413	0.0058	7.2
Population sparsity 0.188	−0.260	0.061	−4.3
Specific humidity (2wks prior) 5.92 g/kg	−0.0047	0.0011	−4.1
log(annual death) 4.04	−0.0324	0.0088	−3.7
Median age 37.48	0.0040	0.0012	3.3

US mortality). A joint, simultaneous, linear fit of these data to 12 potential driving factors (Table 1) revealed only 7 factors with *independent* statistical significance. Re-fitting only to these variables returned the optimal fit for the considered factors (BIC $= -5951$; $R^2 = 0.674$).

We found, not surprisingly, that higher population density, median age, and social mobility correlated with positive exponential growth, while population sparsity, specific humidity, and susceptible depletion correlated with exponentially declining mortality. Notably the coefficients for each of these quantities was in the 95% confidence intervals of those found in the analysis of metropolitan regions (and vice versa). Possibly the most surprising dependency was the negative correlation, at $\simeq -3.7\sigma$ between λ_{14} and the *total* number of annual deaths in the county. In fact, this correlation was marginally more significant than a correlation with log(population), which was -3.3σ. One possible interpretation of this negative correlation is that the number of annual death is a proxy for the number of potential

outbreak clusters. The larger the number of clusters, the longer it might take for the epidemic to spread across their network, which would (at least initially) slow down the onset of the epidemic.

2.4 Nonlinear Model

To obtain more predictive results, we developed a mechanistic nonlinear model for infection (see [21] for details). We followed the standard analogy to chemical reaction kinetics (infection rate is proportional to the product of susceptible and infectious densities), but defined the generation interval (approximately the incubation period) through the excursion probability in a 1D random walk, modulated by an exponential rate of exit from the infected class. This approach resulted in a *renewal equation* [19, 31, 32], with a distribution of generation intervals that is more realistic than that of standard SIR/SEIR models, and which could be solved formally (in terms of the Lambert W function) for the growth rate in terms of the infection parameters:

$$\lambda = \frac{1}{2\tau}\left[\mathrm{W}\left(\sqrt{\frac{\beta\mathcal{S}\tau}{2}}\right)\right]^2 - d \tag{1}$$

The model has four key dependencies, which we describe here, along with our assumptions about their own dependence on population, demographic, and climate variables. As mortality (on which our estimate of growth rate is based) lags infection (on which the renewal equation is based), we imposed a fixed time shift of Δt for time-dependent variables:

1. We assumed that the susceptible population, which feeds new infections and drives the growth, is actually a sub-population of the community, consisting of highly-mobile and frequently interacting individuals, and that most deaths occurred in separate sub-population of largely immobile non-interacting individuals. Under these assumptions, we found (see Supp. Mat. in [21]) that the susceptible density, $S(t)$, could be estimated from the cumulative per capita death fraction, f_D, as:

$$S(t - \Delta t) = S(0)\exp\left[-C_D\, f_D(t)\right] \quad (f_D = D_{\mathrm{tot}}/N)\ ,$$

   where D_{tot} is the cumulative mortality count, N is the initial population, and the initial density is $S(0) = k$ PWD.
2. We assumed that the logarithm of the "rate constant" for infection, β, depended linearly on social mobility, m, specific humidity, h, population sparsity, γ, and

total annual death, A_D, as:

$$\begin{aligned}\ln\left[\beta\left(\mathcal{M},\mathcal{H},\gamma,A_D\right)\right] &= \ln\left[\beta_0\right] \\ &+ C_{\mathcal{M}}\left(\mathcal{M}-\bar{\mathcal{M}}\right)+C_{\mathcal{H}}\left(\mathcal{H}-\bar{\mathcal{H}}\right) \\ &+ C_{\gamma}\left(\gamma-\bar{\gamma}\right)+C_{A_D}\left(A_D-\bar{A}_D\right)\end{aligned} \tag{2}$$

where a barred variable represents the (population-weighted) average value over all US counties, and where the mobility and humidity factors were time-shifted with respect to the growth rate estimation window: $\mathcal{M} = m\left(t-\Delta t\right)$ and $\mathcal{H} = h\left(t-\Delta t\right)$.

3. The characteristic time scale to infectiousness, τ, is intrinsic to the biology and therefore we assumed it would depend only on the median age of the population, A. We assumed a power law dependence:

$$\tau = \tau_0\left(\frac{A}{A_0}\right)^{C_A} \tag{3}$$

where we fixed the pivot age, A_0, to minimize the error in τ_0.

4. The exponential rate of exit from the infected class, d, was assumed constant, since we found no significant dependence for it on other factors in our analysis of US mortality. From the properties of the Lambert W function, when the infection rate or susceptibility density approach zero—through mobility restrictions or susceptible depletion—the growth rate will tend to $\lambda \approx -d$, its minimum value.

With these parameterizations, we performed a nonlinear regression to $\lambda_{14}(t)$ using the entire set of US county mortality incidence time series (Table 2).

Table 2 Best-fit parameters for the nonlinear model using parametrization defined in the text

Parameter	Best-fit ± Std. err	Description
$\tau = \tau_0(\text{Median Age}/26.2\text{ years})^{C_A}$		Time from exposure to contagiousness
τ_0(day)	160 ± 58	Normalization
C_A	-2.26 ± 0.95	Age dependence
d^{-1}(day)	17.6 ± 2.2	Time from exposure to quarantine/recovery
C_D	3460 ± 610	Conversion constant, $f_D \to f_I$
β: Eq. (2)		Rate constant for infection
$\ln\left[k\beta_0\tau_0^{-2}(\text{m}^2/\text{day}^3)\right]$	0.37 ± 1.25	Normalization
$100C_{\mathcal{M}}$	8.08 ± 1.76	Dependence on social mobility
$C_{\mathcal{H}}$	-0.154 ± 0.055	Dependence on specific humidity
C_{γ}	-5.52 ± 2.35	Dependence on population sparsity
C_{A_D}	-1.05 ± 0.25	Dependence on total annual deaths

Compared to the linear model of the previous section (Table 1b), the fit improved by 7.6σ (BIC $= -6008$; $R^2 = 0.724$), despite both having 9 free parameters. Through the estimated parameter values, the model makes predictions for an individual's probability of becoming infectious, and the distributions of incubation period and generation interval, all as a function of the median age of the population (see Supplementary Material in [21]).

The model was very well fit to the mortality growth rate measurements for counties with a high mortality (Fig. 3). More quantitatively, the scatter of measured

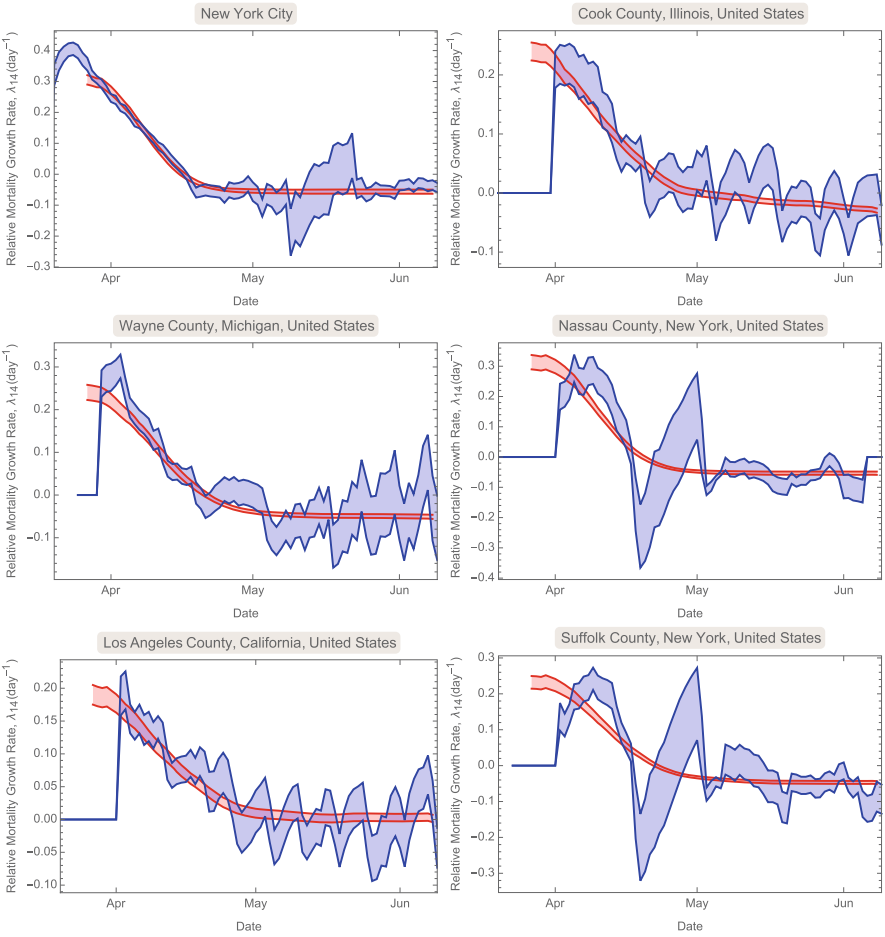


Fig. 3 Nonlinear model prediction (Eq. 1, red) for the actual (blue) mortality growth rate, in the six counties with highest reported death. Bands show 1-σ confidence region for both the model mean and the λ_{14} value

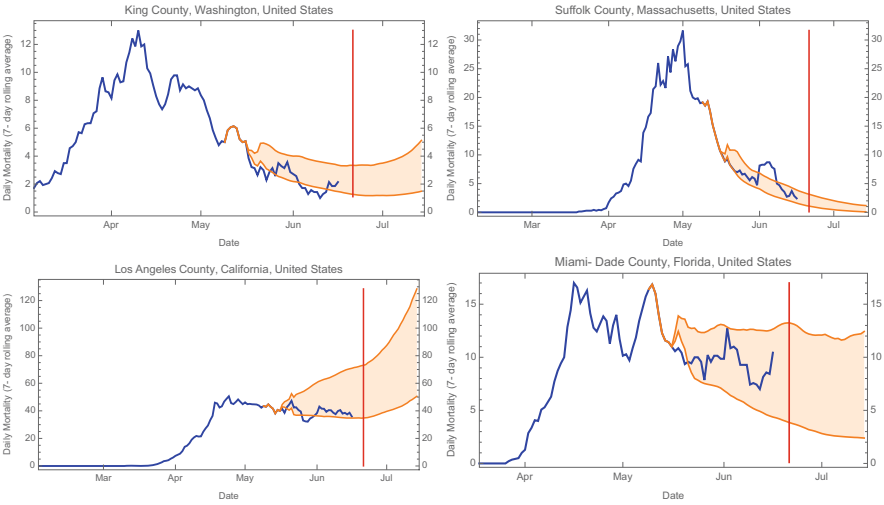


Fig. 4 Forecasts of COVID-19 mortality (orange)—based on the best-fit nonlinear model to data prior to May 16th, 2020—versus actual reported mortality (blue) for 4 large US counties. The 68% confidence range (orange regions) were determined from 100 random 60-day long simulations (see Supplementary Methods in [21]). The vertical red lines indicate June 21st. Forecasts for most US counties can be found at our online dashboard: https://wolfr.am/COVID19Dash

growth rates around the best-fit model predictions was (on average) only 13% larger than the measurement errors, independent of the population of the county.[2]

Importantly, when the model was calibrated on only a subset of the data—e.g., all but the final month for which mobility data is available—its 68% confidence prediction for the remaining data was accurate (Fig. 4) given the known mobility and weather data for that final month. This suggests that the model, once calibrated on the first wave of COVID-19 infections, can make reliable predictions about the ongoing epidemic, and future waves, in the United States.

2.5 Predictions for Relaxed Mobility Restrictions, the Onset of Summer, and the Potential Second Wave

Possibly the most pressing question for the management of COVID-19 in a particular community is the combination of circumstances at which the virus fails to propagate, i.e., at which the growth rate, estimated here by λ_{14}, becomes negative (or, equivalently, the reproduction number R_t falls below one). In the absence of mobility restrictions this is informally called the threshold for "herd immunity,"

[2] See [21] for more detailed discussion of Error Diagnostics.

which is usually achieved by mass vaccination (e.g., [33, 34]). Without a vaccine, however, ongoing infections and death will deplete the susceptible population and thus decrease transmission. Varying the parameters of the nonlinear model individually about their Spring 2020 population-weighted mean values (Fig. 5) suggests that this threshold will be very much dependent on the specific demographics, geography, and weather in the community, but it also shows that reductions in social mobility can significantly reduce transmission prior the onset of herd immunity.

To determine the threshold for herd immunity in the absence or presence of social mobility restrictions, we considered the "average US county" (i.e., a region with population-weighted average characteristics), and examined the dependence of the growth rate on the cumulative mortality. We found that in the absence of social distancing, a COVID-19 mortality rate of 0.13% (or 1300 per million population) would bring the growth rate to zero. However, changing the population density of this average county shows that the threshold can vary widely (Fig. 5).

Examination of specific counties showed that the mortality level corresponding to herd immunity varies from 10 to 2500 per million people (Fig. 6). At the current levels of reported COVID-19 mortality, we found that, as of June 22nd, 2020, only 128 ± 55 out of 3142 counties (inhabiting $9.4 \pm 2.1\%$ of US population) have surpassed this threshold at 68% confidence level (Fig. 7). Notably, New York City, with the highest reported per capita mortality (2700 per million) has achieved mobility-independent herd immunity at the 10σ confidence level, according to the model (Fig. 8). A few other large-population counties in New England, New Jersey, Michigan, Louisiana, Georgia and Mississippi that have been hard hit by the pandemic also appear to be at or close to the herd immunity threshold. This is not the case for most of the United States, however (Fig. 7). Nationwide, we predict that COVID-19 herd immunity would only occur after a death toll of $340,000 \pm 61,000$, or 1058 ± 190 per million of population.

We found that the approach to the herd immunity threshold is not direct, and that social mobility restrictions and other non-pharmaceutical interventions must be applied carefully to avoid excess mortality beyond the threshold. In the absence of social distancing interventions, a typical epidemic will "overshoot" the herd immunity limit (e.g., [35, 36]) by up to 300%, due to ongoing infections (Fig. 8). At the other extreme, a very strict "shelter in place" order would simply delay the onset of the epidemic; but if lifted (see Figs. 8 and 9), the epidemic would again overshoot the herd immunity threshold. A modest level of social distancing, however—e.g., a 33% mobility reduction for the average US county—could lead to fatalities "only" at the level of herd immunity. Naturally, communities with higher population density or other risk factors (see Fig. 5), would require more extreme measures to achieve the same.

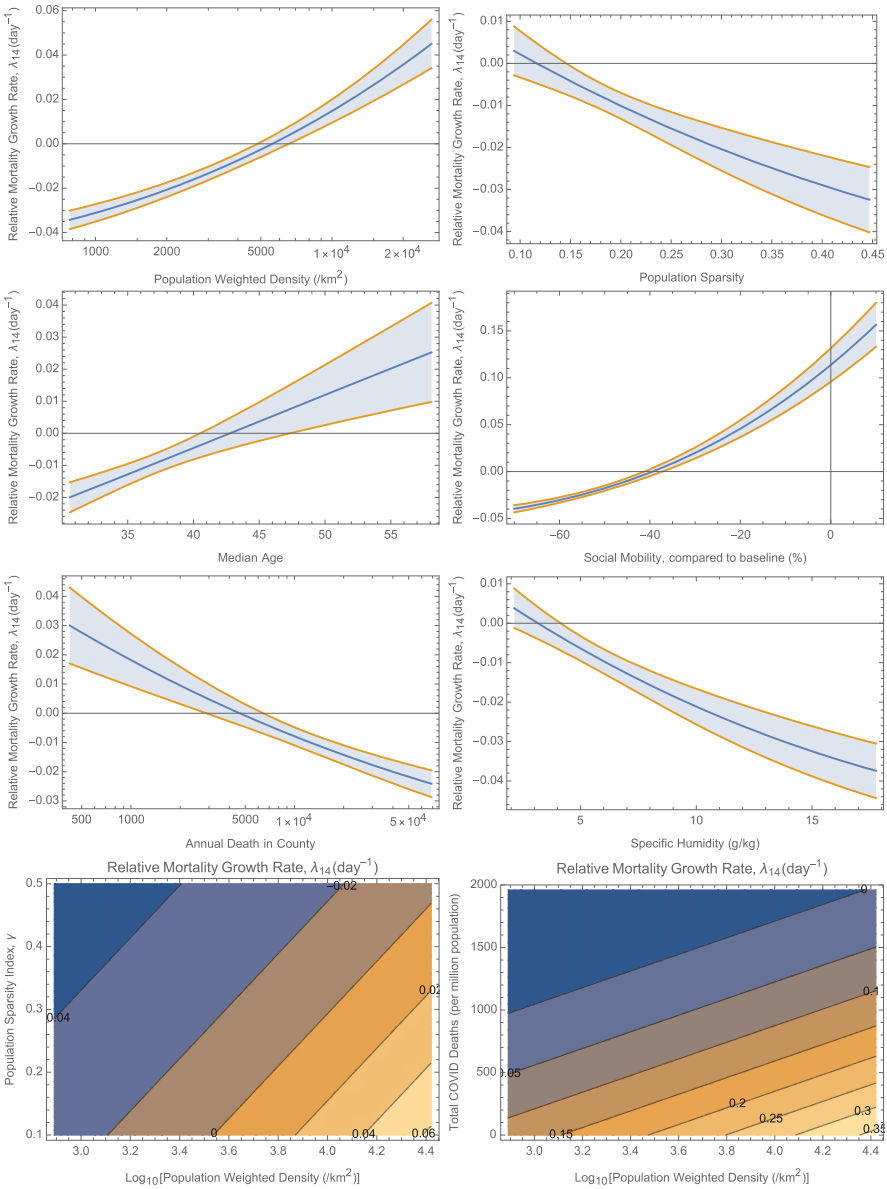


Fig. 5 Dependence and 68% confidence bands of the mortality growth rate—as specified by the nonlinear model (Eq. 1)—on various parameters for an "average county." All parameters not being varied are fixed at their population-weighted mean values (as of 8th June, 2020): $\log_{10}[\text{PWD/km}^{-2})] = 3.58$, population sparsity = 0.188, COVID death fraction $= 5.1 \times 10^{-4}$ (510 deaths/million population), Median age = 37.5 yr, log(annual death) = 4.04, social mobility $\bar{\mathcal{M}} = -44\%$, and specific humidity $\bar{\mathcal{H}} = 5.7$ g/kg

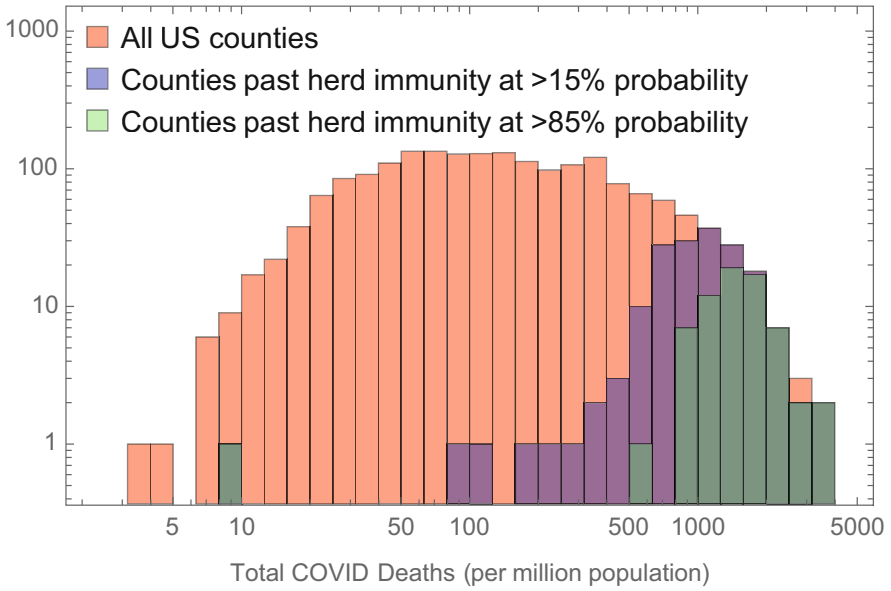


Fig. 6 Histogram of reported COVID-19 deaths per million for all US counties, showing the proportion that have passed "herd immunity" threshold, according to fit of the nonlinear model

Avoiding the level of mortality required for herd immunity will require long-lasting and effective non-pharmaceutical options, until a vaccine is available. The universal use of face masks has been suggested for reducing the transmission of SARS-CoV-2, with a recent meta-analysis [37] suggesting that masks can suppress the rate of infection by a factor of 0.07–0.34 (95% CI), or equivalently $\Delta \ln(\text{transmission}) = -1.9 \pm 0.4$ (at 1σ). Using our model's dependence of the infection rate constant on mobility, this would correspond to an equivalent social mobility reduction of $\Delta\bar{\mathcal{M}}_{\text{mask}} \simeq -24\% \pm 9\%$. Warmer, more humid weather has also considered a factor that could slow the epidemic (e.g., [38–40]). Annual changes in specific humidity are $\Delta\bar{\mathcal{H}} \simeq 6$ g/kg (Figure 10b in [21]), which can be translated in our model to an effective mobility decrease of $\Delta\bar{\mathcal{M}}_{\text{summer}} \simeq -12\% \pm 5\%$. Combining these two effects could, in this simple analysis, yield a modestly

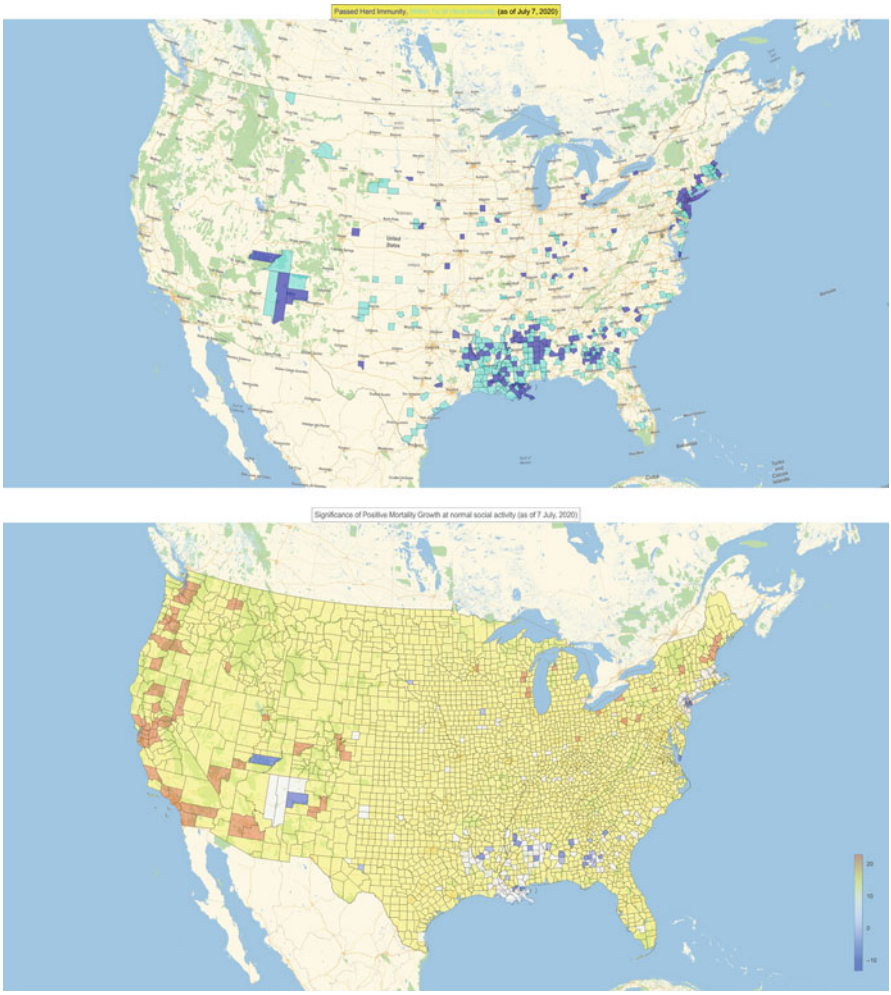


Fig. 7 Top: United States counties that have passed (blue), or are within (cyan), the threshold for "herd immunity" at the 1-σ level, as predicted by the nonlinear model. Bottom: Predicted confidence in the growth of COVID-19 outbreak (defined as predicted daily growth rate divided by its uncertainty), for all counties should they return today to their baseline (pre-COVID) social mobility. Counties that have approached the threshold of herd immunity have lower growth rates due to the depletion of susceptible individuals

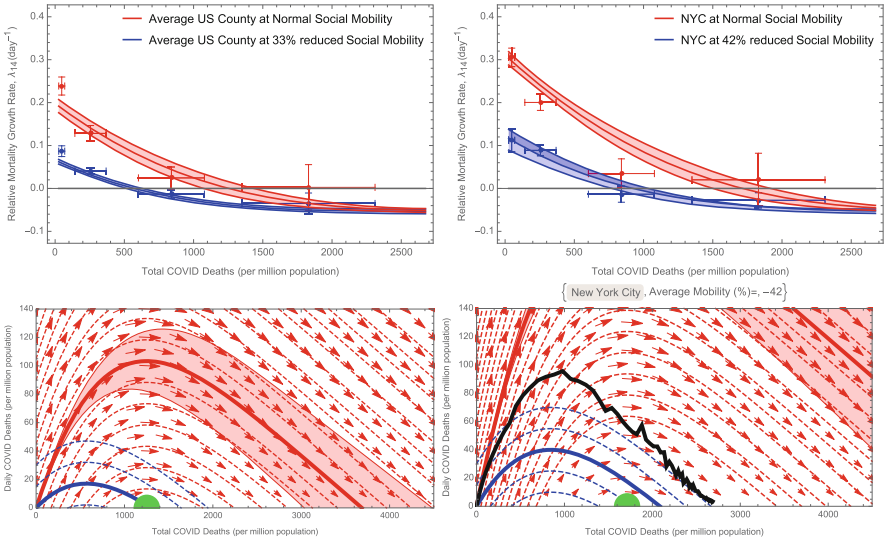


Fig. 8 Nonlinear model prediction of the exponential growth rate, λ_{14}, vs. cumulative COVID-19 mortality (top panels), assuming baseline social mobility, $\bar{\mathcal{M}} = 0$, in the "average US county" (see caption of Fig. 5) on the left, and New York City, on the right. The curves show 68% predictions for the nonlinear model (Table 2), while the points with errorbars are linear fits to all the data within bins of death fraction. The threshold for "herd immunity" ($\lambda_{14} = 0$) is reached at a mortality of approximately 1300 (1700) per million for an average county (NYC), but this would be higher in counties with more unfavorable values of the drivers. The eventual mortality burden of the average county will be determined by its path through a "phase space" of Daily vs. Total Mortality (bottom panel). An epidemic without intervention (red curves, with the particular trajectory starting at zero death shown in bold) will pass the threshold for herd immunity (1300 deaths per million; note that at zero daily deaths this is a fixed point) and continue to three times that value due to ongoing infections. A modest 33% reduction in social mobility (blue curves), however, leads to mortality at "only" the herd immunity level (the green disk). The black curve on the bottom right panel shows the 7-day rolling average of reported mortality for NYC, which appears to have "overshot" the "herd immunity threshold"

effective defense for the summer months: $\Delta\bar{\mathcal{M}}_{\text{mask+ summer}} \simeq -37\% \pm 10\%$. Therefore, this could be a reasonable strategy for most communities to manage the COVID-19 epidemic at the aforementioned -33% level of mobility needed to arrive at herd immunity with the least excess death. More stringent measures would be required to keep mortality below that level. Of course, this general prescription would need to be fine-tuned for the specific conditions of each community.

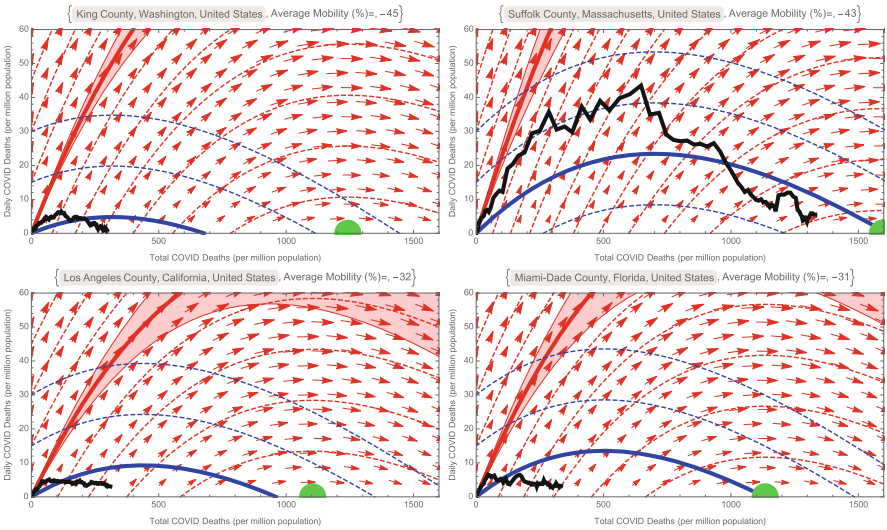


Fig. 9 Epidemic Phase Portraits for the same four counties as in Fig. 4, similar to the Phase portrait in Fig. 8. The blue curves are for the county's average Social Mobility during Feb. 15 through June 12, 2020, while red curves/arrows are at normal (pre-covid) social mobility. The thick black curve is the 7-day rolling average of the official reported mortality, while the green disk shows the threshold for "herd immunity"

3 Discussion and Conclusions

By simultaneously considering the time series of mortality incidence in every US county, and controlling for the time-varying effects of local social distancing interventions, we demonstrated for the first time a dependence of the epidemic growth of COVID-19 on population density, as well as other climate, demographic, and population factors. We further constructed a realistic, but simple, first-principles model of infection transmission that allowed us to extend our heuristic linear model of the dataset into a predictive nonlinear model, which provided a better fit to the data (with the same number of parameters), and which also accurately predicted late-time data after training on only an earlier portion of the data set. This suggests that the model is well-calibrated to predict future incidence of COVID-19, given realistic predictions/assumptions of future intervention and climate factors. We summarized some of these predictions in the final section of Results, notably that only a small fraction of US counties (with less than 10% of the population) seem to have reached the level of herd immunity, and that relaxation of mobility restrictions without counter-measures (e.g., universal mask usage) will likely lead to increased daily mortality rates, beyond that seen in the Spring of 2020.

In any epidemiological model, the infection rate of a disease is assumed proportional to population density [41], but, to our knowledge, its explicit effect in a real-world respiratory virus epidemic has not been demonstrated. The universal

reach of the COVID-19 pandemic, and the diversity of communities affected have provided an opportunity to verify this dependence. Indeed, as we show here, it must be accounted for to see the effects of weaker drivers, such as weather and demographics. A recent study of COVID-19 in the United States, working with a similar dataset, saw no significant effect due to population density [42], but our analysis differs in a number of important ways. First, we have taken a dynamic approach, evaluating the time-dependence of the growth rate of mortality incidence, rather than a single static measure for each county, which allowed us to account for the changing effects of weather, mobility, and the density of susceptible individuals. Second, we have included an explicit and real-time measurement of social mobility, i.e., cell phone mobility data provided by Google [22], allowing us to control for the dominant effect of intervention. Finally, and perhaps most importantly, we calculate for each county an estimate of the "lived" population density, called the population-weighted population density (PWD) [23], which is more meaningful than the standard population per political area. As with any population-scale measure, this serves as a proxy—here, for estimating the average rate of encounters between infectious and susceptible people—but we believe that PWD is a better proxy than standard population density, and it is becoming more prevalent, e.g., in census work [43, 44].

We also found a significant dependence of the mortality growth rate on specific humidity (although since temperature and humidity were highly correlated, a replacement with temperature was approximately equivalent), indicating that the disease spread more rapidly in drier (cooler) regions. There is a large body of research on the effects of temperature and humidity on the transmission of other respiratory viruses [27, 45], specifically influenza [46]. Influenza was found to transmit more efficiently between guinea pigs in low relative-humidity and temperature conditions [24], although re-analysis of this work pointed to absolute humidity (e.g., specific humidity) as the ultimate controller of transmission [25]. Although the mechanistic origin of humidity's role has not been completely clarified, theory and experiments have suggested a snowballing effect on small respiratory droplets that cause them to drop more quickly in high-humidity conditions [47–49], along with a role for evaporation and the environmental stability of virus particles [49, 50]. It has also been shown that the onset of the influenza season [26, 51]—which generally occurs between late-Fall and early-Spring, but is usually quite sharply peaked for a given strain (H1N1, H3N2, or Influenza B)—and its mortality [46] are linked to drops in absolute humidity. It is thought that humidity or temperature could be the annual periodic driver in the resonance effect causing these acute seasonal outbreaks of influenza [52, 53], although other influences, such as school openings/closings have also been implicated [54]. While little is yet known about the transmission of SARS-CoV-2 specifically, other coronaviruses are known to be seasonal [45, 55], and there have been some preliminary reports of a dependence on weather factors [56, 57]. We believe that our results represent the most definitive evidence yet for the role of weather, but emphasize that it is a weak, secondary driver, especially in the early stages of this pandemic where the susceptible fraction of the population remains large [58]. Indeed, the current early-summer rebound of COVID-19 in the

relatively dry and hot regions of the Southwest suggests that the disease spread will not soon be controlled by seasonality.

We developed a new model of infection in the framework of a renewal equation (see, e.g., [32] and references therein), which we could formally solve for the exponential growth rate. The incubation period in the model was determined by a random walk through the stages of infection, yielding a non-exponential distribution of the generation interval, thus imposing more realistic delays to infectiousness than, e.g., the standard SEIR model. In this formulation, we did not make the standard compartmental model assumption that the infection of an individual induces an autonomous, sequential passage from exposure, to infectiousness, to recovery or death; indeed, the model does not explicitly account for recovered or dead individuals. This freedom allows for, e.g., a back passage from infectious to noninfectious (via the underlying random walk) and a variable rate of recovery or death. We assumed only that the exponential growth in mortality incidence matched (with delay) that of the infected incidence—the primary dynamical quantity in the renewal approach—and we let the cumulative dead count predict susceptible density—the second dynamical variable in the renewal approach—under the assumption that deaths arise from a distinct subset of the population, with lower mobility behavior than those that drive infection (see [21]). Therefore, we fitted the model to the (rolling 2-week estimates of the) COVID-19 mortality incidence growth rate values, λ_{14}, for all counties and all times, and used the per capita mortality averaged over that period, f_D, to determine susceptible density. Regression to this nonlinear model was much improved over linear regression, and, once calibrated on an early portion of the county mortality incidence time series, the model accurately predicted the remaining incidence.

Because we accounted for the precise effects of social mobility in fitting our model to the actual epidemic growth and decline, we were then able to, on a county-by-county basis, "turn off" mobility restrictions and estimate the level of cumulative mortality at which SARS-CoV-2 would fail to spread even without social distancing measures, i.e., we estimated the threshold for "herd immunity." Meeting this threshold prior to the distribution of a vaccine should not be a goal of any community, because it implies substantial mortality, but the threshold is a useful benchmark to evaluate the potential for local outbreaks following the first wave of COVID-19 in Spring 2020. We found that a few counties in the United States have indeed reached herd immunity in this estimation—i.e., their predicted mortality growth rate, assuming baseline mobility, was negative—including counties in the immediate vicinity of New York City, Detroit, New Orleans, and Albany, Georgia. A number of other counties were found to be at or close to the threshold, including much of the greater New York City and Boston areas, and the Four Corners, Navajo Nation, region in the Southwest. All other regions were found to be far from the threshold for herd immunity, and therefore are susceptible to ongoing or restarted outbreaks. These determinations should be taken with caution, however. In this analysis, we estimated that the remaining fraction of susceptible individuals in the counties at or near the herd immunity threshold was in the range of 0.001% to 5% (see [21]). This is in strong tension with initial seroprevalence studies [59, 60]

which placed the fraction of immune individuals in New York City at 7% in late March and 20% in late April, implying that perhaps 75% of that population remains susceptible today. We hypothesize that the pool of susceptible individuals driving the epidemic in our model is a subset of the total population—likely those with the highest mobility and geographic reach—while a different subset, with very low baseline mobility, contributes most of the mortality (see [21]). Thus, the near total depletion of the susceptible pool we see associated with herd immunity corresponds to the highly-mobile subset, while the low-mobility subset could remain largely susceptible. One could explicitly consider such factors of population heterogeneity in a model—e.g., implementing a saturation of infectivity as a proxy for a clustering effect [61–64]—but we found (in results not shown) that the introduction of additional of parameters left portions of the model unidentifiable. Despite these cautions, it is interesting to note that the epidemic curves (mortality incidence over time) for those counties that we have predicted an approach to herd immunity are qualitatively different than those we have not. Specifically, the exponential rise in these counties is followed by a peak and a sharp decline—rather than the flattening seen in most regions—which is a typical feature of epidemic resolution by susceptible depletion.

At the time of this writing, in early Summer 2020, confirmed cases are again rising sharply in many locations across the United States—particularly in areas of the South and West that were spared significant mortality in the Spring wave. The horizon for an effective and fully-deployed vaccine still appears to be at least a year away. Initial studies of neutralizing antibodies in recovered COVID-19 patients, however, suggest a waning immune response after only 2–3 months, with 40% of those that were asymptomatic becoming seronegative in that time period [65]. Although the antiviral remdesivir [66–68] and the steroid Dexamethasone [69] have shown some promise in treating COVID-19 patients, the action of remdesivir is quite weak, and high-dose steroids can only be utilized for the most critical cases. Therefore, the management of this pandemic will likely require non-pharmaceutical intervention—including universal social distancing and mask-wearing, along with targeted closures of businesses and community gathering places—for years in the future. The analysis and prescriptive guidance we have presented here should help to target these approaches to local communities, based on their particular demographic, geographic, and climate characteristics, and can be facilitated through our http://mylocalcovid.uwaterloo.ca/ online simulator dashboard. Finally, although we have focused our analysis on the United States, due to the convenience of a diverse and voluminous data set, the method and results should be applicable to any community worldwide, and we intend to extend our analysis in forthcoming work.

Acknowledgments We are indebted to helpful comments and discussions by our colleagues, in particular Bruce Bassett, Ghazal Geshnizjani, David Spergel, and Lee Smolin. NA is partially supported by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Colleges and Universities. BPH acknowledges sabbatical support from Grand Valley State University, and is grateful to the hospitality of the University of Waterloo during his stay.

References

1. N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Y. Qiu, J. Wang, Y. Liu, Y. Wei, et al., The Lancet **395**(10223), 507 (2020)
2. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KS, Lau EH, Wong JY, et al. 2020 Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. New England Journal of Medicine 382, 1199–1207, doi: https://doi.org/10.1056/NEJMoa2001316.
3. Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, Spitters C, Ericson K, Wilkerson S, Tural A, et al. 2020 First case of 2019 novel coronavirus in the United States. New England Journal of Medicine 382, 929–936, doi: https://doi.org/10.1056/NEJMoa2001191.
4. Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, Eggo RM, Sun F, Jit M, Munday JD, Davies N, Gimma A, van Zandvoort K, Gibbs H, Hellewell J, Jarvis CI, Cliord S, Quilty BJ, Bosse NI, Abbott S, Klepac P, Flasche S. 2020 Early dynamics of transmission and control of COVID-19: A mathematical modelling study. The Lancet Infectious Diseases 20, 553–558, doi: https://doi.org/10.1016/S1473-3099(20)30144-4.
5. Chu HY, Englund JA, Starita LM, Famulare M, Brandstetter E, Nickerson DA, Rieder MJ, Adler A, Lacombe K, Kim AE, et al. 2020 Early detection of COVID-19 through a citywide pandemic surveillance platform. New England Journal of Medicine 383, 185–187.
6. Ladner JT, Larsen BB, Bowers JR, Hepp CM, Bolyen E, Folkerts M, Sheridan K, Pfeier A, Yaglom H, Lemmer D, Sahl JW, Kaelin EA, Maqsood R, Bokulich NA, Quirk G, Watts TD, Komatsu KK, Waddell V, Lim ES, Caporaso JG, Engelthaler DM, Worobey M, Keim P, Fraser CM. 2020 An early pandemic analysis of SARS-CoV-2 population structure and dynamics in Arizona. mBio 11, e02107-20, doi: https://doi.org/10.1128/mBio.02107-20.
7. Gonzalez-Reiche AS, Hernandez MM, Sullivan MJ, Ciferri B, Alshammary H, Obla A, Fabre S, Kleiner G, Polanco J, Khan Z, et al. 2020 Introductions and early spread of SARS-CoV-2 in the New York City area. Science 369, 297–301.
8. Worobey M, Pekar J, Larsen BB, Nelson MI, Hill V, Joy JB, Rambaut A, Suchard MA, Wertheim JO, Lemey P. 2020 The emergence of SARS-CoV-2 in Europe and North America. Science 370, 564–570.
9. X. Deng, W. Gu, S. Federman, L. Du Plessis, O. Pybus, N. Faria, C. Wang, G. Yu, C.Y. Pan, H. Guevara, et al., Science (2020). https://doi.org/10.1126/science.abb9263
10. Fauver JR, Petrone ME, Hodcroft EB, Shioda K, Ehrlich HY, Watts AG, Vogels CB, Brito AF, Alpert T, Muyombwe A, et al. 2020 Coast-to-coast spread of SARS-CoV-2 during the early epidemic in the United States. Cell 181, 990–996.
11. Park SW, Bolker BM, Champredon D, Earn DJ, Li M, Weitz JS, Grenfell BT, Dusho_J. 2020 Reconciling early-outbreak estimates of the basic reproductive number and its uncertainty: Framework and applications to the novel coronavirus (SARS-CoV-2) outbreak. Journal of the Royal Society Interface 17, 20200144.
12. H. Yu, S. Cauchemez, C.A. Donnelly, L. Zhou, L. Feng, N. Xiang, J. Zheng, M. Ye, Y. Huai, Q. Liao, et al., Emerging infectious diseases **18**(5), 758 (2012)
13. A.D. Storms, M.D. Van Kerkhove, E. Azziz-Baumgartner, W.K. Lee, M.A. Widdowson, N.M. Ferguson, A.W. Mounts, Influenza and other respiratory viruses **7**(6), 1328 (2013)
14. S. Sanche, Y. Lin, C. Xu, E. Romero-Severson, N. Hengartner, R. Ke, Emerging infectious diseases **26**(7) (2020)
15. B. Oliveiros, L. Caramelo, N.C. Ferreira, F. Caramelo, medRxiv (2020)
16. J. Wallinga, M. Lipsitch, Proceedings of the Royal Society B: Biological Sciences **274**(1609), 599 (2007)
17. M. Roberts, J. Heesterbeek, Journal of mathematical biology **55**(5–6), 803 (2007)
18. Dusho_J, Park SW. 2021 Speed and strength of an epidemic intervention. Proceedings of the Royal Society B 288, 20201556.
19. D. Champredon, J. Dushoff, Proceedings of the Royal Society B: Biological Sciences **282**(1821), 20152026 (2015)
20. H. Nishiura, Mathematical Biosciences & Engineering **7**(4), 851 (2010)

21. N. Afshordi, B. P. Holder, M. Bahrami, D. Lichtblau, arXiv preprint arXiv:2007.00159 (2020)
22. J. Fitzpatrick, D. Karen, COVID-19 community mobility reports. Tech. rep., Google (2020). Available at: https://www.google.com/covid19/mobility/
23. J. Craig, Population Trends **39**, 16 (1985)
24. A. Lowen, S. Mubareka, J. Steel, P. Palese, **3**(10), e151 (2007). https://doi.org/10.1371/journal.ppat.0030151
25. J. Shaman, M. Kohn, Proceedings of the National Academy of Sciences **106**(9), 3243 (2009)
26. J. Shaman, E. Goldstein, M. Lipsitch, American journal of epidemiology **173**(2), 127 (2011)
27. E. Kudo, E. Song, L.J. Yockey, T. Rakib, P.W. Wong, R.J. Homer, A. Iwasaki, Proceedings of the National Academy of Sciences **116**(22), 10905 (2019)
28. Pearce N, Vandenbroucke JP, VanderWeele TJ, Greenland S. 2020 Accurate statistics on COVID-19 are essential for policy guidance and decisions. American Journal of Public Health 110, 949–951, doi: https://doi.org/10.2105/AJPH.2020.305708.
29. D.A. Leon, V.M. Shkolnikov, L. Smeeth, P. Magnus, M. Pechholdová, C.I. Jarvis, The Lancet **395**(10234), e81 (2020)
30. C. Modi, V. Boehm, S. Ferraro, G. Stein, U. Seljak, medRxiv (2020). https://doi.org/10.1101/2020.04.15.20067074. URL https://www.medrxiv.org/content/early/2020/05/14/2020.04.15.20067074
31. J. Heesterbeek, K. Dietz, Statistica Neerlandica **50**(1), 89 (1996)
32. D. Champredon, J. Dushoff, D.J. Earn, SIAM Journal on Applied Mathematics **78**(6), 3258 (2018)
33. T.J. John, R. Samuel, European Journal of Epidemiology **16**(7), 601 (2000). URL https://doi.org/10.1023/A:1007626510002
34. P. Fine, K. Eames, D.L. Heymann, Clinical Infectious Diseases **52**(7), 911 (2011). URL https://doi.org/10.1093/cid/cir007
35. A. Handel, I.M. Longini Jr, R. Antia, Proceedings of the Royal Society B: Biological Sciences **274**(1611), 833 (2007)
36. I.C.H. Fung, R. Antia, A. Handel, PLoS One **7**(6), e36573 (2012)
37. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schunemann HJ, El-harakeh A, Bognanni A, Lot T, Loeb M, et al. 2020 Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: A systematic review and meta-analysis. The Lancet 395, 1973–1987.
38. Wang J, Tang K, Feng K, Lin X, Lv W, Chen K, Wang F. 2021 Impact of temperature and relative humidity on the transmission of COVID-19: A modelling study in China and the United States. BMJ Open 11, e043863.
39. A. Notari, arXiv e-prints arXiv:2003.12417 (2020)
40. R. Xu, H. Rahmandad, M. Gupta, C. DiGennaro, N. Ghaffarzadegan, H. Amini, M.S. Jalali, medRxiv (2020). https://doi.org/10.1101/2020.05.05.20092627. URL https://www.medrxiv.org/content/early/2020/05/24/2020.05.05.20092627
41. M.C. de Jong, O. Diekmann, H. Heesterbeek, in *Epidemic Models: Their structure and relation to data*, ed. by D. Mollison (Cambridge University Press, Cambridge, 1995), pp. 84–94
42. S. Hamidi, S. Sabouri, R. Ewing, Journal of the American Planning Association **0**(0), 1 (2020). URL https://doi.org/10.1080/01944363.2020.1777891
43. Dorling D, Atkins D. 1995 Population density, change and concentration in Great Britain 1971, 1981 and 1991, number 58 in Studies on Medical and Population Subjects, London: HMSO.
44. S.G. Wilson, *Patterns of metropolitan and micropolitan population change: 2000 to 2010* (US Department of Commerce, Economics and Statistics Administration, 2012)
45. M. Moriyama, W.J. Hugentobler, A. Iwasaki, Annual Review of Virology **7**(1), null (2020). URL https://doi.org/10.1146/annurev-virology-012420-022445. PMID: 32196426
46. A.I. Barreca, J.P. Shimshack, American Journal of Epidemiology **176**(suppl_7), S114 (2012). URL https://doi.org/10.1093/aje/kws259
47. R. Tellier, Journal of the Royal Society Interface **6**(suppl_6), S783 (2009)
48. J.D. Noti, F.M. Blachere, C.M. McMillen, W.G. Lindsley, M.L. Kashon, D.R. Slaughter, D.H. Beezhold, PloS one **8**(2), e57485 (2013)

49. L.C. Marr, J.W. Tang, J. Van Mullekom, S.S. Lakdawala, Journal of The Royal Society Interface **16**(150), 20180298 (2019). https://doi.org/10.1098/rsif.2018.0298. URL https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2018.0298
50. L. Morawska, in *Proceedings of Indoor Air 2005: the 10th International Conference on Indoor Air Quality and Climate* (Springer, 2005), pp. 9–23
51. J. Shaman, V.E. Pitzer, C. Viboud, B.T. Grenfell, M. Lipsitch, PLoS Biology **8**(2), e1000316 (2010)
52. J. Dushoff, J.B. Plotkin, S.A. Levin, D.J.D. Earn, **101**(48), 16915 (2004). https://doi.org/10.1073/pnas.0407293101
53. J. Tamerius, M.I. Nelson, S.Z. Zhou, C. Viboud, M.A. Miller, W.J. Alonso, Environmental health perspectives **119**(4), 439 (2011)
54. D.J. Earn, D. He, M.B. Loeb, K. Fonseca, B.E. Lee, J. Dushoff, Annals of internal medicine **156**(3), 173 (2012)
55. R.A. Neher, R. Dyrdak, V. Druelle, E.B. Hodcroft, J. Albert, Swiss medical weekly **150**(1112) (2020)
56. R. Xu, H. Rahmandad, M. Gupta, C. DiGennaro, N. Ghaffarzadegan, H. Amini, M.S. Jalali, medRxiv (2020). https://doi.org/10.1101/2020.05.05.20092627. URL https://www.medrxiv.org/content/early/2020/05/24/2020.05.05.20092627
57. M. Schell, B.D. Gonzalez, J. Greene, A. Giuliano, Available at SSRN 3579744 (2020). URL https://doi.org/10.2139/ssrn.3579744
58. Baker RE, Yang W, Vecchi GA, Metcalf CJE, Grenfell BT. 2020 Susceptible supply limits the role of climate in the early SARS-CoV-2 pandemic. Science 369, 315–319.
59. E.S. Rosenberg, J.M. Tesoriero, E.M. Rosenthal, R. Chung, M.A. Barranco, L.M. Styer, M.M. Parker, S.Y.J. Leung, J. Morne, D. Greene, D.R. Holtgrave, D. Hoefer, J. Kumar, T. Udo, B. Hutton, H.A. Zucker, medRxiv (2020). https://doi.org/10.1101/2020.05.25.20113050. URL https://www.medrxiv.org/content/early/2020/05/29/2020.05.25.20113050
60. F.P. Havers, C. Reed, T.W. Lim, J.M. Montgomery, J.D. Klena, A.J. Hall, A.M. Fry, D.L. Cannon, C.F. Chiang, A. Gibbons, I. Krapiunaya, M. Morales-Betoulle, K. Roguski, M. Rasheed, B. Freeman, S. Lester, L. Mills, D.S. Carroll, S.M. Owen, J.A. Johnson, V.A. Semenova, J. Schiffer, N.P. Thornburg, medRxiv (2020). https://doi.org/10.1101/2020.06.25.20140384. URL https://www.medrxiv.org/content/early/2020/06/26/2020.06.25.20140384
61. V. Capasso, G. Serio, Mathematical Biosciences **42**(1–2), 43 (1978)
62. D. Mollison, in *Population Dynamics of Rabies in Wildlife* (1985), pp. 223–234
63. R.J. De Boer, Journal of virology **81**(6), 2838 (2007)
64. A. Farrell, C. Brooke, K. Koelle, R. Ke, BioRxiv p. 547349 (2019)
65. Q.X. Long, X.J. Tang, Q.L. Shi, Q. Li, H.J. Deng, J. Yuan, J.L. Hu, W. Xu, Y. Zhang, F.J. Lv, et al., Nature Medicine pp. 1–5 (2020)
66. J.H. Beigel, K.M. Tomashek, L.E. Dodd, A.K. Mehta, B.S. Zingman, A.C. Kalil, E. Hohmann, H.Y. Chu, A. Luetkemeyer, S. Kline, D. Lopez de Castilla, R.W. Finberg, K. Dierberg, V. Tapson, L. Hsieh, T.F. Patterson, R. Paredes, D.A. Sweeney, W.R. Short, G. Touloumi, D.C. Lye, N. Ohmagari, M.d. Oh, G.M. Ruiz-Palacios, T. Benfield, G. Fätkenheuer, M.G. Kortepeter, R.L. Atmar, C.B. Creech, J. Lundgren, A.G. Babiker, S. Pett, J.D. Neaton, T.H. Burgess, T. Bonnett, M. Green, M. Makowski, A. Osinusi, S. Nayak, H.C. Lane, New England Journal of Medicine **0**(0), null (2020). URL https://doi.org/10.1056/NEJMoa2007764
67. J. Grein, N. Ohmagari, D. Shin, G. Diaz, E. Asperges, A. Castagna, T. Feldt, G. Green, M.L. Green, F.X. Lescure, E. Nicastri, R. Oda, K. Yo, E. Quiros-Roldan, A. Studemeister, J. Redinski, S. Ahmed, J. Bernett, D. Chelliah, D. Chen, S. Chihara, S.H. Cohen, J. Cunningham, A. D'Arminio Monforte, S. Ismail, H. Kato, G. Lapadula, E. L'Her, T. Maeno, S. Majumder, M. Massari, M. Mora-Rillo, Y. Mutoh, D. Nguyen, E. Verweij, A. Zoufaly, A.O. Osinusi, A. DeZure, Y. Zhao, L. Zhong, A. Chokkalingam, E. Elboudwarej, L. Telep, L. Timbs, I. Henne, S. Sellers, H. Cao, S.K. Tan, L. Winterbourne, P. Desai, R. Mera, A. Gaggar, R.P. Myers, D.M. Brainard, R. Childs, T. Flanigan, New England Journal of Medicine **382**(24), 2327 (2020). URL https://doi.org/10.1056/NEJMoa2007016

68. Y. Wang, D. Zhang, G. Du, R. Du, J. Zhao, Y. Jin, S. Fu, L. Gao, Z. Cheng, Q. Lu, et al., The Lancet (2020)
69. P. Horby, W.S. Lim, J. Emberson, M. Mafham, J. Bell, L. Linsell, N. Staplin, C. Brightling, A. Ustianowski, E. Elmahi, B. Prudon, C. Green, T. Felton, D. Chadwick, K. Rege, C. Fegan, L.C. Chappell, S.N. Faust, T. Jaki, K. Jeffery, A. Montgomery, K. Rowan, E. Juszczak, J.K. Baillie, R. Haynes, M.J. Landray, medRxiv (2020). https://doi.org/10.1101/2020.06.22.20137273. URL https://www.medrxiv.org/content/early/2020/06/22/2020.06.22.20137273