State-level variation for initial COVID-19 dynamics in the United States

Easton R. White[1,2]

[2]Department of Biology, University of Vermont, VT 05405, USA

Abstract

In the development of an epidemic, metrics such as R_0, doubling time, and case fatality rates are important in understanding and predicting the course of an epidemic. However, if collected over country or regional scales, these metrics hide important smaller-scale, local dynamics. We examine how commonly used epidemiological metrics differ for each individual state within the United States during the initial COVID-19 outbreak. We found that the case number, and trajectory of cases, differs considerably between states. Although individual states are clearly not independent, they can serve as mini, natural experiments in how different demographic patterns and government responses can impact the course of an epidemic. Thus, these results should be used to better understand, in near real-time, what actions are working most effectively.

Keywords: SARS-CoV-2, COVID-19, spatial heterogeneity, doubling time

Daily updates to figures in this manuscript are available at: https://github.com/eastonwhite/COVID19_US_States

[1]Easton.White@uvm.edu

Introduction

The global SARS-CoV-2 (COVID-19) pandemic began in Wuhan, China in late 2019 (WHO 2020). As of March 22th, 335,955 cases have been reported across 171 countries and regions. There has been several sets of efforts to track the progression of the outbreak across the world and within countries. For example, John Hopkins University Center for Systems Science and Engineering (CSSE) has compiled data from various sources, including the US Center for Disease Control and the World Health Organization, to present a global picture of COVID-19 cases and deaths (Dong *et al.* 2020). These efforts have allowed for international scientific research and political decision-making. Although data are collected at local scales (e.g. within hospitals), in an emerging pandemic data is typically reported at the country level. This allows for interesting comparisons between countries (Anderson *et al.* 2020) and for information from an earlier affected country to be used to slow the outbreak in other places. For instance, South Korea was able to flatten their outbreak curve through early and widespread testing as well as strict quarantine policies (citation). However, country-level analyses still hide more local dynamics that are important to the overall epidemic progression.

Spatial heterogeneity is important for population dynamics generally (Levin 1992) and in particular for understanding the progression of infectious disease dynamics (Grenfell *et al.* 1995). Spatial heterogeneity can include differences in local population density, movement patterns, suitability of environmental conditions for transmission, among other factors. For instance, Keeling *et al.* (2001) showed how spatial distribution and size of farms affected the 2001 UK Foot and Mouth Epidemic.

Here we examine the progression of COVID-19 for state-level differences within the United States. We examine how commonly-used metrics, including doubling time, can vary state to state and compared to the US as a whole. Although not independent units, we can use state-level data to understand the progression of the outbreak across different replicates within a country. Our results can be used to infer what actions by which states are currently most effective in managing the outbreak. We also provide a link to daily updates of our findings as the exact quantitative results are likely to change over the course of the epidemic.

Results and Discussion

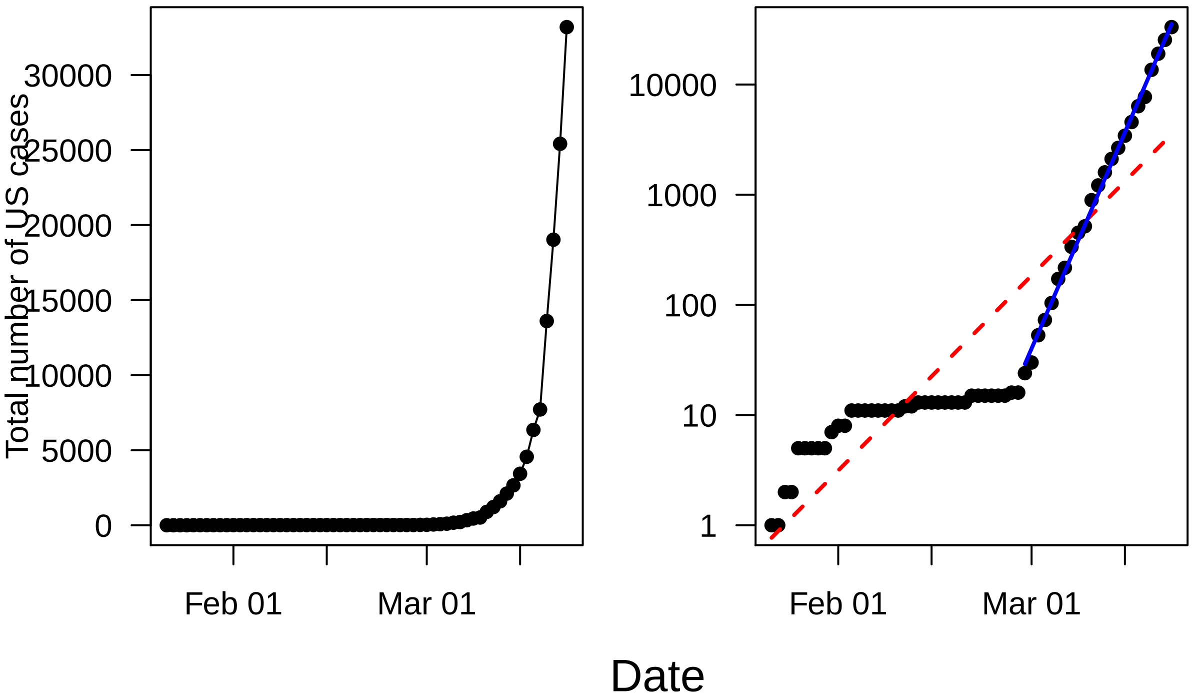

Figure 1: (Left panel) Cases versus time for the whole United States. (Right panel) Log number of cases versus time for the whole United States. The red, dashed line is the line of best fit for all the data and the blue, solid line is the line of best fit since Feburary 29th.

We used data compiled by John Hopkins University Center for Systems Science and Engineering (Dong *et al.* 2020). The United States has seen exponential growth in the number of cases, especially since February 29th (Fig. 1). The exponential growth rate, and corresponding doubling time, has yet to arrive at an equilibrium for the US as a whole (Fig. 2). This may also be due to the need for more data to achieve high enough statistical power (White 2019). Country-level results, however, hide underlying dynamics within each state.

Therefore, we examined how the number of cases changed over time within each state. To properly compare the progression of the epidemic across states, we looked at the log number

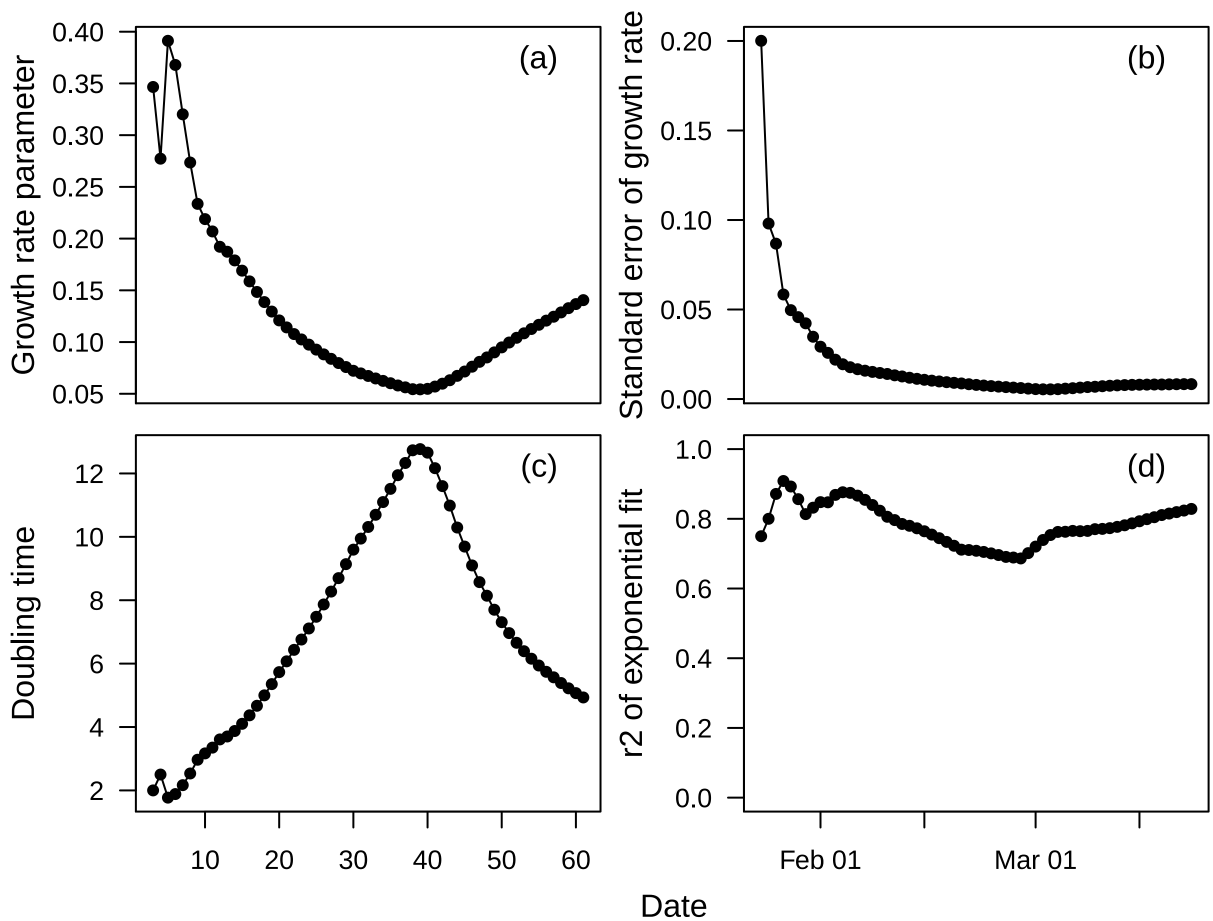

Figure 2: The effect of each additional day of data on (a) exponential growth rate parameter, (b) standard error of the exponential growth rate parameter, (c) doubling time, and (d) the R^2 value for a fitted exponential curve.

of cases since the first day a state reported 25 cases (Fig. 3). On a log scale, a straight line of the cases over time indicates exponential growth where the slope of the line is the exponential growth parameter. We found considerable differences between states in how the outbreak has progressed (Fig. 3). States like New York, New Jersey, and Michigan have experienced a doubling of cases approximately every two days. Conversely, Massachusetts has experienced a doubling time closer to five days. We mapped doubling time across the US and found regional differences where the West Coast and Northeast have seen large doubling times, i.e. slower outbreak dynamics (Fig. 4).

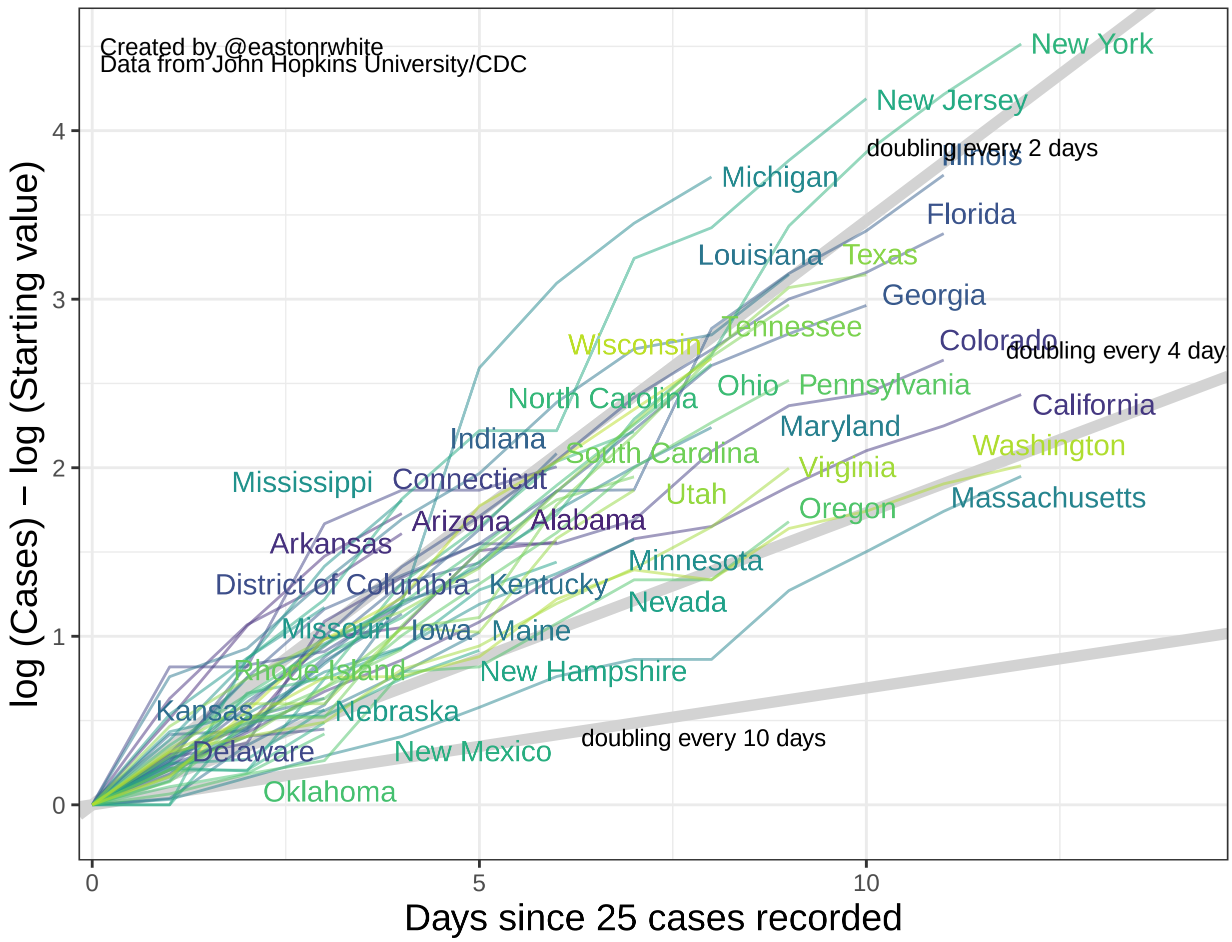

Figure 3: The log number of cases over time for each individual state that recorded more than 25 cases over at least three days. The light grey diagonal lines represent the growth trajectory for doubling times of 2, 4, and 10 days. The log number of the starting value (intial number of cases on first day when at least 25 cases were recorded) had to be subtracted on the y-axis to standarize the graph across states.

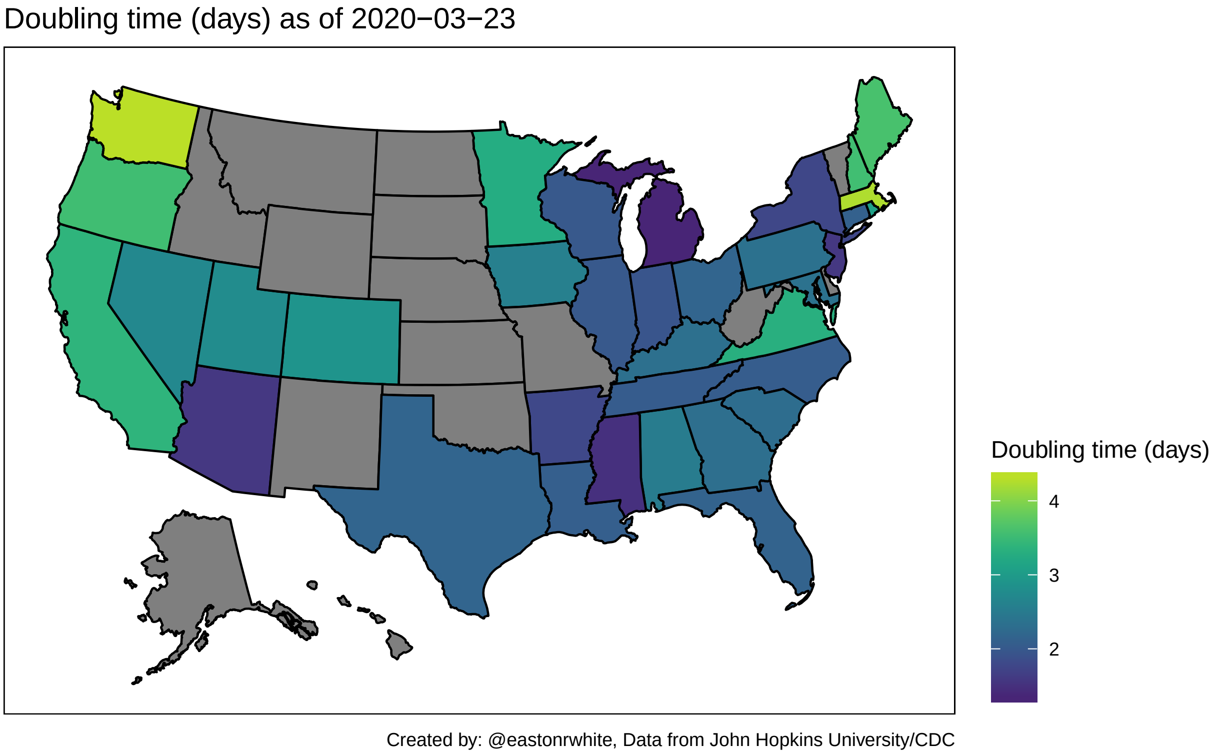

Figure 4: Doubling time (in number of days) across the US states that recorded more than 25 cases over at least three days.

Conclusions and Future Work

We found a large degree of heterogeneity in the number of cases over time across US states. These differences might be for statistical and data collection issues or demographic and governance patterns. The US is still early in the progression of the epidemic. Therefore, the number of detected cases might not have settled into a long-term trajectory. In addition, because of limited testing, there is potentially a large number of unobserved cases (Perkins *et al.* 2020).

Each US state varies considerably across a number of important axes: wealth, access to healthcare, number of international travelers, age distribution, population density, among other factors. In addition, much of the response to COVID-19 has been done at the state, as opposed to federal, government level in the US (citations). These responses have varied a lot with different laws and such… (citations). Future work will focus on explaining the distinct dynamics observed across the different US states.

Code availability and acknowledgements

All code and corresponding data is freely available at https://github.com/eastonwhite/COVID19_US_States. The original raw data has been compiled by the Johns Hopkins University Center for Systems Science and Engineering at (https://github.com/CSSEGISandData/COVID-19). The authors received no specific funding for this work.

References

Anderson, R.M., Heesterbeek, H., Klinkenberg, D. & Hollingsworth, T.D. (2020). How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet*, 395, 931–934.

Dong, E., Du, H. & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*, S1473309920301201.

Grenfell, B.T., Bolker, B.M. & Kleczkowski, a. (1995). Seasonality and Extinction in Chaotic Metapopulations. *Proceedings of the Royal Society B: Biological Sciences*, 259, 97–103.

Jombart, T., van Zandvoort, K., Russell, T., Jarvis, C., Gimma, A. & Abbott, S. *et al.* (2020). Inferring the number of COVID-19 cases from recently reported deaths. *medRxiv*.

Keeling, M.J., Woolhouse, M.E.J., Shaw, D.J., Matthews, L., Chase-Topping, M. & Haydon, D.T. *et al.* (2001). Dynamics of the 2001 UK foot and mouth epidemic: Stochastic dispersal in a heterogeneous landscape. *Science*, 294, 813–817.

Levin, S.A. (1992). The Problem of Pattern and Scale in Ecology. *Ecology*, 73, 1943–1967.

Park, A.W. (2012). Infectious disease in animal metapopulations : The importance of environmental transmission. *Ecology and Evolution*, 2, 1398–1407.

Perkins, A., Cavany, S.M., Moore, S.M., Oidtman, R.J., Lerch, A. & Poterek, M. (2020). Estimating unobserved SARS-CoV-2 infections in the United States. *medRxiv*.

Schreiber, S.J. (2010). Interactive effects of temporal correlations, spatial heterogeneity and dispersal on population persistence. *Proceedings of the Royal Society B: Biological Sciences*, 277, 1907–1914.

White, E.R. (2019). Minimum time required to detect population trends: The need for long-term monitoring programs. *BioScience*, 69, 40–46.

WHO. (2020). *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)*.