medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

How Timing of Stay-home Orders and Mobility Reductions Impacted First-Wave COVID-19 Deaths in US Counties

Michelle Audirac, Mauricio Tec, Lauren Ancel Meyers, Spencer Fox and Cory Zigler*
The University of Texas at Austin COVID-19 Modeling Consortium

Abstract

SARS-CoV-2 transmission continues to evolve in the United States following the large second wave in the Summer. Understanding how location-specific variations in non-pharmaceutical epidemic control policies and behaviors contributed to disease transmission will be key for designing effective strategies to avoid future resurgences. We offer a statistical analysis of the relative effectiveness of the timing of both official stay-at-home orders and population mobility reductions, offering a distinct (but complementary) dimension of evidence gleaned from more traditional mechanistic models of epidemic dynamics. Specifically, we use a Bayesian hierarchical model fit to county-level mortality data from the first wave of the pandemic from Jan 21 2020 through May 10 2020 to establish how timing of stay-at-home orders and population mobility changes impacted county-specific epidemic growth. We find that population mobility reductions generally preceded stay-at-home orders, and among 356 counties with a pronounced early local epidemic between January 21 and May 10 (representing 195 million people and 32,000 observed deaths), a 10 day delay in population mobility reduction would have added 16,149 (95% credible interval [CI] 9,517 24,381) deaths by Apr 20, whereas shifting mobility reductions 10 days earlier would have saved 13,571 (95% CI 8,449 16,930) lives. Analogous estimates attributable to the timing of explicit stay-at-home policies were less pronounced, suggesting that mobility changes were the clearer drivers of epidemic dynamics. Our results also suggest that the timing of mobility reductions and policies most impacted epidemic dynamics in larger, urban counties compared with smaller, rural ones. Overall, our results suggest that community behavioral changes had greater impact on curve flattening during the Spring wave compared with stay at home orders. Thus, community engagement and buy-in with precautionary policies may be more important for predicting transmission risk than explicit policies.

1 Introduction

SARS-CoV-2, the causative virus of COVID-19, continues to threaten the world with nearly 9.2M reported cases and 231,500 deaths as of November 1st, 2020 [John Hopkins University, 2020]. Retrospective understanding of the effectiveness of previous precautionary efforts remains key for guiding policy decisions to avert current and future resurgences. However, there are many population-specific factors that influence pandemic trajectories including social and demographic characteristics, dynamic behavioral responses, as well as the timing of non-pharmaceutical interventions (NPIs) and precautionary policies [Khan et al., 2020; Desmet & Wacziarg, 2020].

It is clear that both NPIs and behavioral changes have impacted transmission of the virus (Jinjarak et al. [2020]; Lin & Meissner [2020]; Kapoor et al. [2020]; Abouk & Heydari [2020]; Pei et al. [2020]; Du et al. [2020]). However, decoupling the relative impacts of these factors presents important challenges as discussed in (Abouk & Heydari [2020]; Courtemanche et al. [2020]), and despite the existing evidence, there remains no consensus about the relative impact of these measures on epidemic growth. A key problem particularly in the United States, is that policy changes and behav-

*Corresponding author: cory.zigler@austin.utexas.edu

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

ior shifts happened rapidly in late March and April in response to growing epidemics and happened simultaneously across the country, so it has thus far been difficult to disentangle the relative impact of behavioral responses from NPIs.

In this work, we quantify relative impact of heterogeneous control efforts and responses relative to local epidemic conditions. Specifically, we developed a Bayesian hierarchical statistical model of county-level COVID-19 mortality across 440 counties that had a pronounced epidemic between January 21, 2020 and May 10, 2020 and had either: a) an official stay-at-home order in place or b) a reduction in population mobility of at least 50% relative to baseline during this time period. Using the model we evaluate how the timing of stay-at-home orders or population mobility reductions contributed to epidemic control, and focus particularly on differences across the urban-rural divide across the country in terms of local pre- and post-intervention trends. A similar framework has previously been shown to capture important heterogeneous dynamics to reasonably forecast deaths during the first stages of the US epidemic [Woody et al., 2020].

Our model specification offers refinement over traditional statistical approaches used in this realm (e.g. difference-in-differences or cross-sectional regression) for its ability to separate modeling of the natural trajectory of the epidemic curve from changes due to other trends or external forces such as an intervention. Compared with more mechanistic epidemiological models, our statistical approach has two key advantages in that it (1) doesn't necessitate explicit specification or assumptions about the underlying disease dynamics and (2) is better suited to explicitly account for and quantify key sources of demographic heterogeneity around the country. Thus, the approach pursued here can be viewed as a hybrid approach that attempts to capture the advantages of both statistical and mechanistic models to elucidate the relative impact of behavioral changes and NPIs [Feltham et al., 2020].

2 Methods

2.1 County-Level COVID Death and Demographic Data

Our analysis relies on three distinct data sources. The first, which is itself a combination of data sources compiled and reported by Killeen et al. [2020], comprises data on socioeconomic factors that may affect the spread of epidemiological outbreaks, along with confirmed COVID-19 deaths at the county level from the COVID-19 Data Repository compiled by the the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University.

Also comprised in this dataset are the dates in which six different explicit policy interventions were put in place in each county: bans of 500+ and 50+ people gatherings; closure mandates of schools, non-essential businesses and restaurants; and orders to stay at home. Intervention policies were most-often put in place state wide, but there are cases in which counties adopted such measures before their respective states. Generally speaking, orders to stay at home came last in a quick succession of policies that occurred primarily between March 15 and April 1. We consider a single date of policy intervention for each county corresponding to the date of the orders to stay at home. Appendix A.1 describes the timing of all six different policy interventions.

As a summary measure of county characteristics known to relate to residents' contact rates and the reproduction number (population density, modes of travel, distance to major airports) and other factors expected to vary across the spectrum of rural and urban areas, we augment this county-level data set with the National Center for Health Statistics (NCHS) Urban-Rural Classification Scheme for Counties [Ingram & Franco, 2012], which classifies each US county to be in one of the following six categories: 1) large central metro; 2) large fringe metro; 3) medium metro; 4) small metro; 5) micropolitan; 6) non-core. In addition, we extract from the US Census American Community Survey other county-level covariates that have been reported to have a strong relationship with COVID-19 death rate and are not captured by the NCHS county classification. The percent of black residents and percent of hispanic residents are included to account for the apparent disparities between infection and comorbidity and death rates among these populations relative to other races and ethnicities. To account for the age-related risk of death, we include the percentage of residents that are 65 years or older. To capture any particular behaviors specific to students (whose main activity quickly became completely remote amid the early epidemic stages), we also include the percentage of residents attending college.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

NCHS	Number of counties in US	Number of counties in dataset	Population covered in dataset	Days between policy intervention and threshold	Days between mobility intervention and threshold	days between mobility and policy intervention
Large central metro	68	57	93 M	5 (0,8)	0 (-4,5)	-4 (-5,-2)
Large fringe metro	368	120	48 M	0 (-5,6)	-5 (-8,0)	-4 (-9,-2)
Medium metro	373	121	41 M	0 (-5,4)	-5 (-9,-2)	-6 (-11,-2)
Small metro	358	67	9 M	0 (-6,5)	-5 (-10,-1)	-4 (-11,-2)
Micropolitan	641	52	3 M	-3 (-7,2)	-10 (-12,0)	-6 (-10,-1)
None-core	1339	23	0.72 M	-5 (-8,0)	-11 (-16, -5)	-7 (-11,-4)
Total	3147	440	194.72 M	0 (-5,5)	-5 (-9,0)	-5 (-10,-2)

Table 1: Description of the counties coverage in the analysis grouped by NCHS along with the median number of days between interventions and death threshold. In parenthesis, values for the first and quantiles are included (25%tile, 75%tile).

2.2 Human Mobility Data from SafeGraph

The final data source comes from SafeGraph, a company that provides anonymized population mobility datasets representing 45 million smartphone devices. SafeGraph data aggregates visit counts to numerous points of interest (POIs) classified into categories.

As a proxy measure for overall mobility behavior in each county, we extracted data on the number of visits per day to POIs; in particular to schools, colleges, restaurants, bars, parks and museums, and obtained time series of daily total visits for each county. Although Safegraph provides visits data for several other types of POIs, as well as data on the number of minutes devices remain at home or at work, we limited our analysis only to those categories that had a comprehensive coverage across all the range of counties. The average total visits per day between Jan 15, 2020 and Feb 15, 2020 was used to establish county-baseline levels of mobility. We define the date of the mobility-based intervention to be the date on which the right-aligned ten-day moving average of total visits decreased 50% relative to baseline.

2.3 Analysis Data Set

County-specific daily counts of COVID19 cases/deaths were extracted from Jan 21, 2020 through May 10, 2020. During that period, 940 counties (out of a total of 3221 US counties) had reached a threshold of 3 deaths per 10 million residents and had at least 5 cumulative deaths.

We restrict attention to those counties that instituted a stay at home order or exhibited a population mobility reduction of at least 50% relative to baseline, yielding two final analysis data sets: one containing the 420 counties that put a stay-at-home order in place regardless of their mobility patterns, and the other with the 356 counties whose mobility data reflect a 50% decrease in total visits regardless of whether a stay-home order was put in place. In total, these two datasets comprised data on 440 distinct counties, with 336 having both policy and mobility interventions, and 84 (20) having only a policy (mobility) intervention. To maintain focus on the short-term impact of interventions, we limited the observation window for each county to 29 days after a stay-at-home intervention or mobility reduction. Table 1 summarizes number of counties and population covered in each NCHS classification category, indicating that this relatively small fraction of US counties covers approximately 195 million residents.

2.4 Statistical Methods

To model county-specific death trajectories, we use a Bayesian hierarchical model for the time series of each county's 7-day centered rolling average of daily deaths. The model consists of a log-linear model for the expected number of deaths in each county using a negative binomial distribution and a quadratic function of "epidemic time," defined as the number of days elapsed since reaching a deaths threshold of 3 per 10 million residents (signalling local arrival of the epidemic). Latent heterogeneity

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

in time trends across counties is captured by county-specific random effects for each degree of the quadratic time function. A related quadratic trend model was shown to be useful in forecasting the early transmission of the disease in US metropolitan areas [Woody et al., 2020].

Full details of the model specification appear in Appendix B, but we note here its salient features. Quadratic time polynomials were used to parsimoniously capture nonlinearity of daily death trajectories that take different forms depending on multiple county-specific features. Specifically, each county's trajectory is modeled to vary according to (i.e., interact with) the county's NCHS county classification and the percent of resident college attendees, residents over age 65, black residents, and Hispanic residents. The county-specific random effects permit additional heterogeneity in county-level death trajectories not captured by these features.

To capture the impact of interventions (either stay-at-home or mobility reductions), the model is specified to entail a change in the quadratic time function after the introduction of a stay-at-home order or mobility reduction, where the change in trajectory occurs at a lag of 12 days after the intervention date[1]. The extent of post-intervention change in the death trajectory is specified to depend on: 1) the number of days between the stay-at-home order or mobility reduction and the introduction of the epidemic in the county (i.e., the intervention timing) and 2) the county's NCHS county classification. The county time series used for model fitting ran from the date of the first death in the county through a maximum 29 day period post-intervention (representing the 12 day intervention lag plus 17 day post-intervention trend).

The analysis consists of three distinct variations of model:

1. *Stay-at-home model*: Here, the "intervention" and relative timing is defined according to the date at which a county instituted a stay-at-home order. This model is fit to the 420 counties in the analysis data set that had stay-at-home orders.
2. *Mobility model*: Here, the "intervention" and relative timing is defined according to the date at which a county reached a 50% reduction from baseline mobility patterns. This model is fit to the 356 counties in the analysis data set that reached such a reduction during the study time period.
3. *Double intervention model*: This model, fit to the 336 counties exhibiting both mobility reductions and stay-at-home orders, specifies two analogous "post-intervention" changes in death trajectories: one corresponding to the date of the stay-at-home order, and another corresponding to the date of the mobility reduction.

We fit all models using the `R` language (`3.6.3`) with the package `rstanarm(2.19.3)` [Goodrich et al., 2018]. Inferences are based on posterior simulations from the models, including simulations from the posterior-predictive distribution of deaths for each county under different hypothesized timing of stay-at-home orders or mobility reductions. Differences in cumulative deaths across counties were calculated by aggregating posterior-predicted daily deaths for all counties in a given NCHS county category on a given date under a given timing scenario then accumulating resulting values across days through April 20 and comparing the resulting predictions under the observed vs. early intervention scenario and the observed vs. late intervention scenario. Full details appear in Appendix B.

3 Results

3.1 Epidemic Timing of Stay-at-Home Orders and Mobility Reductions

Among the 440 counties in the present analysis, death thresholds of 3 per 10 million residents, representing our characterization of local arrival of the epidemic, were reached at different calendar times across the period of March 1 to April 15, with clear differences across levels of NCHS county classification. According to the average date of reaching this deaths threshold, the epidemic arrived

[1] A delay consistent with the first quartile of the distribution of time between infection and death. [Lauer et al., 2020] estimate the first quartile of time between infection to symptoms to be 3.8 days, while [Yang et al., 2020] estimate the first quartile of days between symptoms to death at 10. By adding both we obtain an approximation of the left tail of the distribution of time lag of deaths. A similar calculation is done by [Wilson et al., 2020] using IQR values.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

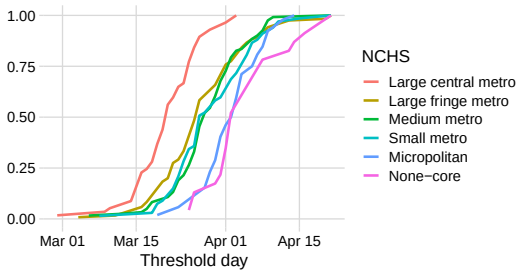


Figure 1: The introduction of COVID-19 varied across NCHS classification. y-axis represents the cumulative percentage of counties in each category that had reached the threshold of 3 deaths per 10million residents by the corresponding calendar date.

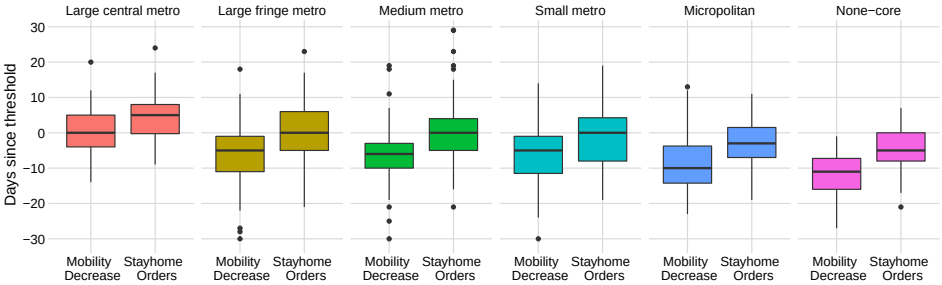


Figure 2: Epidemic timing of mobility reduction and stay-at-home orders across NCHS groups. Days since threshold (y-axis) of 0 represents the day that the threshold of 3 deaths per 10M residents was reached, negative numbers represent days before the threshold.

earliest in large central metro areas (average date March 21), followed by large fringe, medium, and small metro areas (average dates March 27, March 28, and March 29, respectively), with the latest epidemic arrival in micropolitan and non-core counties (average dates April 2 and April 4, respectively). The full distribution of dates of epidemic arrival across county categories appears in Figure 1.

Among the 420 counties that instituted stay-at-home orders, the average date of the order was March 28. In terms of "epidemic timing," the median number of days between a county reaching the deaths threshold of 3/10 million and institution of the stay-at-home order was 0 across all counties, corresponding to a stay at home order on the date the threshold was reached. Since the calendar dates of stay-at-home orders were relatively homogeneous, variability in the epidemic timing of the stay-at-home orders was largely dictated by the timing of the epidemic arrival. This led the epidemic timing of stay-at-home orders across categories of NCHS urban-rural classification to exhibit the opposite temporal ordering of epidemic arrival: stay-at-home orders occurred earliest in epidemic time in non-core and micropolitan counties (median 5 and 3 days before deaths threshold, respectively), followed by small and medium metro areas (median day between days between stay-at-home order and threshold of 0, indicating orders on threshold date), with large fringe and large central metro areas having stay-at-home orders latest in epidemic time (median days between deaths threshold 0 and 5 days after threshold). See Table 1 and Figure 2 for more detail on the epidemic timing of stay-at-home orders across NCHS categories.

Among the 356 counties exhibiting at least a 50% reduction in mobility relative to baseline, the average date at which this reduction was reached was March 22, with a median day between mobility and reduction and reaching the deaths threshold of 3/10 million of 5 days before the threshold. The relative epidemic timing at which counties across different NCHS categories reached this mobility reduction showed the same ordering as the dates of stay-at-home orders, but were generally shifted earlier, with median days between deaths threshold and mobility reductions of 11, 10, 5, 5, 5 days before reaching the death thresholds for non-core, micropolitan, small metro, medium metro, and large fringe metro areas, respectively, and 0 for large central metro areas indicating mobility re-

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

ductions on the date the death threshold was reached. See Table 1 and Figure 2 for more detail on the epidemic timing of mobility reductions. Note in particular that, for the 336 counties with both mobility reductions and stay-at-home orders, the mobility reductions occurred earlier, with mobility reductions tending to precede stay-at-home orders by a median of 5 days across all NCHS categories. Appendix A.2 provides additional descriptives of mobility changes, indicating in particular that, among the counties with both stay at home orders and at least a 50% reduction in mobility: a) most exhibited steady drops in mobility in the 10 days preceding the date at which the 50% reduction threshold was met; b) by the date of the stay at home order, mobility reductions had reached, on average, approximately 70%; and c) decreased mobility held relatively constant in the 29 days following implementation of the stay-at-home order. Knowledge that mobility was substantially decreased by the time stay-at-home orders took place and subsequently held fairly constant during the study time frame are important features when interpreting modeled results.

3.2 Modeling Results and Effectiveness of Intervention Timing

Results from fitting the models in Section 2.4 indicate that county-specific death trajectories significantly differed according to counties' NCHS category, percent of college student residents, percent of Hispanic residents, and percent of Black residents, as indicated by significant interactions between these terms and quadratic time parameters. Appendix B.3 provides further details, along with a table summarizing the estimates of all model parameters.

To evaluate the impact of the timing of stay-at-home and mobility reductions, we use the posterior distribution of parameters from the model in Section 2.4 to compare the modeled trajectory of daily deaths in each county and a model-based prediction of what the daily death trajectory would have been under different intervention timing. Specifically, we offer posterior predictions for each county under the hypothetical scenarios where the stay-at-home order or 50% mobility reduction threshold occurred 10 days before or 10 days after the observed date, tabulating estimated deaths under these scenarios through April 20 at least or 30 days since threshold. Appendix B.5 presents estimated curves associated with three specific counties to illustrate the model's flexibility for characterizing epidemic trajectories for different county characteristics and intervention timing configurations. Figure 3 shows the average fitted curves of daily deaths (per capita) under the observed intervention scenarios and hypothetical intervention timings, separated by NCHS category.

Figure 3(a) depicts modeled daily death trajectories for different timing of stay-at-home orders using posterior predictions from the *Stay-at-home model*. While average trajectories are largely overlapping between the observed and late intervention timing, there is evidence that intervening 10 days earlier would have significantly impacted the trajectory of daily deaths, particularly for the more urban counties classified as large central, large fringe, or medium metro areas. Table 2a provides estimates of cumulative deaths differences and differences in deaths per 100,000 residents under earlier and later introduction of stay-at-home orders. Overall, the model predicts that implementing stay-at-home orders 10 days earlier would have led to 9,903 (95% CI: 4,307 13,990) fewer deaths through April 20 across the 420 counties in the *Stay-at-home model*, with effects concentrated in the large central, large fringe, and medium metro counties. This corresponds to a reduction of 5.2 (95% CI: 2.3, 7.4) deaths per 100,000 residents. The evidence that delayed action would have led to extra deaths is weaker; adopting stay-at-home orders 10 days later would have led to an additional 3,995 (95% CI: -244 8,402) deaths and 2.1 (95% CI: -0.1, 4.4) deaths per 100,000 residents, although note that these the uncertainty intervals around these estimates (and those for each individual NCHS category) cannot rule out the possibility of no impact of delayed stay-at-home orders. Appendix B.4 provides results in terms of the average days between epidemic initiation and the daily death peak and the height of the daily death peak.

Figure 3(b) depicts estimated daily death curves for different timing of mobility reduction timing using posterior predictions from the *Mobility model*. While the general shape of trajectories is similar to those for the stay-at-home interventions, important differences emerge. For large central, large fringe, and medium metro areas, there is more pronounced evidence that earlier mobility intervention impacted the daily death trajectories; intervening earlier is predicted to have reduced the time until peak death rate and the height of peak death rate. In addition, and unlike in the stay-at-home model, having intervened later than observed is estimated to impact the daily death trajectories among these counties, with longer times to peak and higher peak death rates. As in the analysis of stay-at-home timing, there is evidence of similar patterns in the small metro and micropolitan coun-

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

	Cumulative Deaths on Apr 20				
NCHS Category	*Observed Deaths*	*Deaths Averted by Earlier Action* (2.5%,97.5%)	*Deaths Added by Later Action* (2.5%,97.5%)	*Deaths Averted* per 100,000 residents (2.5%,97.5%)	*Deaths Added* per 100,000 residents (2.5%,97.5%)
Large central metro	20,553	8,003 (3,255 11,431)	3,725 (-84 7,791)	8.7 (3.5 12.4)	4.1 (-0.1 8.5)
Large fringe metro	7,345	1,213 (-1,035 2,730)	249 (-847 1,496)	2.5 (-2.2 5.7)	0.5 (-1.8 3.1)
Medium metro	2,900	825 (19 1,316)	129 (-428 706)	2.1 (0.0 3.4)	0.3 (-1.1 1.8)
Small metro	662	-21 (-356 215)	-141 (-277 14)	-0.2 (-4.3 2.6)	-1.7 (-3.4 0.2)
Micropolitan	325	48 (-164 188)	-39 (-137 102)	1.6 (-5.3 6.1)	-1.3 (-4.4 3.3)
None-core	75	-64 (-248 25)	-15 (-55 27)	-10.5 (-40.7 4.1)	-2.5 (-9.0 4.4)
total	31,860	9,903 (4,370 13,990)	3,995 (-244 8,402)	5.2 (2.3 7.4)	2.1 (-0.1 4.4)

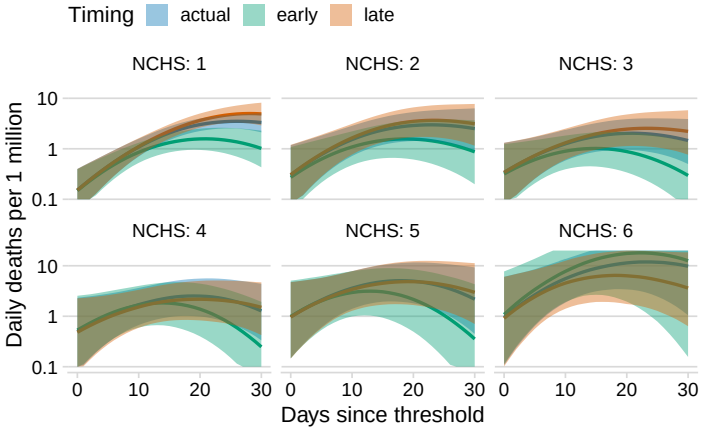
(a) For 420 counties in the *Stay-at-home model*

	Cumulative Deaths on Apr 20				
NCHS Category	*Observed Deaths*	*Deaths Averted by Earlier Action* (2.5%,97.5%)	*Deaths Added by Later Action* (2.5%,97.5%)	*Deaths Averted* per 100,000 residents (2.5%,97.5%)	*Deaths Added* per 100,000 residents (2.5%,97.5%)
Large central metro	20,577	8,465 (4,364 11,665)	9,709 (4,221 16,164)	9.1 (4.7 12.6)	10.5 (4.6 17.4)
Large fringe metro	7,323	3,985 (2,514 5,007)	5,101 (2,480 9,093)	9.4 (6.0 11.9)	12.1 (5.9 21.5)
Medim metro	2,975	1,508 (877 1,982)	1,480 (433 2,962)	4.1 (2.4 5.4)	4.0 (1.2 8.1)
Small metro	663	-118 (-847 229)	-153 (-336 113)	-1.7 (-12.0 3.2)	-2.2 (-4.8 1.6)
Micropoitan	311	97 (-94 204)	-56 (-164 92)	4.1 (-4.0 8.7)	-2.4 (-7.0 3.9)
None-core	63	-194 (-2,412 7)	-47 (-75 11)	-49.2 (-611.7 1.9)	-11.9 (-18.9 2.9)
Total	31,912	13,571 (8,449 16,930)	16,149 (9,517 24,381)	7.5 (4.7 9.3)	8.9 (5.2 13.4)

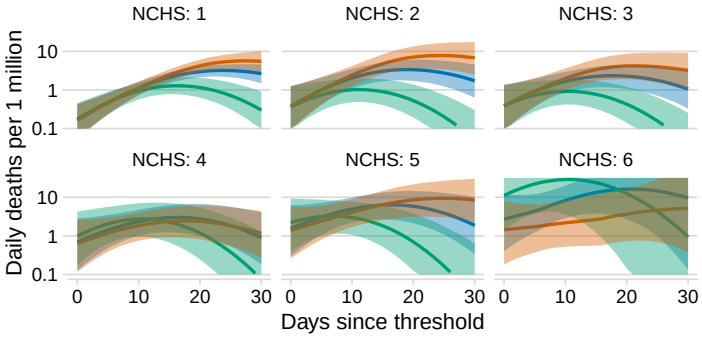
(b) For 356 counties in the *Mobility decrease model*

Table 2: Estimated (median and 95% credible intervals) COVID-19 deaths averted (added) by 10 day earlier (later) stay-at-home order or mobility reduction, through April 20, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



(a) **Stay-at-home order model**



(b) **50% mobility decrease model**

Figure 3: Estimated death trajectories for observed timing, 10-day earlier timing, and 10-day delayed timing of stay-at-home orders and mobility reductions. Solid lines are posterior median estimates of deaths per 1 million residents at t days after epidemic arrival. Shaded areas are point wise 90% posterior intervals.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

ties, but the associated uncertainty renders the evidence inconclusive. The level of uncertainty for the non-core counties makes it difficult to glean anything. In fact, as evident from Figure 2, mobility interventions in these least urban counties occurred before the arrival of the epidemic (median value of 11 days before, c.f. Table 1), so it is expected that a 10 day earlier or later intervention would have little or no effect because the epidemic conditions in such counties were practically the same as when the observed mobility change occurred. Table 2b shows estimates of the differences in cumulative deaths by April 20 for earlier or later mobility reductions in the 356 counties used for the *Mobility model*. Overall, shifting mobility reductions 10 days earlier in these counties would have averted an estimated 13,571 (95%CI: 8,449 16,930) deaths, while delaying mobility reductions by 10 days would have led to an additional 16,149 (95%CI: 9,517 24,381) deaths over what was observed. These estimates correspond, respectively, to 7.5 (95% CI 4.7, 9.3) and 8.9 (95% CI 5.2, 13.4) deaths per 100,000 residents, with most pronounced impacts accruing in the large central, large fringe, and medium metro areas. Thus, unlike in the analysis of stay-at-home orders, both earlier and later mobility reductions are estimated to significantly impact deaths, and these impacts are generally more pronounced (in terms of deaths per 100,000 residents) for the mobility reductions than for the stay-at-home orders. Overall across all NCHS categories and within the large central, large fringe, and medium metro areas, the estimates of deaths per 100,000 residents are uniformly higher for the impact of earlier or delayed timing of mobility reductions[2]. Additional numeric summaries in terms of peak characteristics appear in Table 4.

In total, the results from the *Mobility model* relative to the *Stay-at-home model* match expectations, since stay-at-home orders typically happened after a significant decrease in mobility had already taken place (*c.f.* Table 1), with mobility drops persisting beyond the first date reaching a 50% reduction from baseline. This result is corroborated in the *Double intervention model* that simultaneously considers bends in death trajectories based on both the stay-at-home order date and the mobility reduction date. In this model, we see that after accounting for changes in deaths attributable to the timing of the mobility reductions, the impact of stay-at-home orders was negligible, with trend-altering parameters not significantly different from zero. Note that interpretation of the *Double intervention model* is complicated by the possible interplay between mobility patterns and adoption of stay-at-home orders, a point to which we return in Section 4.

4 Discussion

We have offered a rigorous statistical evaluation of the relative effectiveness of the timing of both stay-at-home orders and population mobility reductions. The statistical analyses here expresses county-specific curves of daily COVID-19 deaths in terms of a pre- and a post intervention trajectories, dictated by both county-level features and the epidemic timing of policy and behavioral efforts for epidemic control. This formulation permitted inferences regarding relative effectiveness, in terms of deaths avoided, of different epidemic timings of both an explicit policy intervention and one measure of population mobility reductions.

The descriptive analysis of the timing of stay-at-home orders and total visits to points of interest clarified both the heterogeneity in intervention timing relative to local epidemic conditions and that mobility reductions often occurred prior to stay-home-orders. By-and-large, more urban counties classified as large central, large fringe, or medium metro areas tended to have the epidemic arrive earlier in calendar time, with mobility and policy interventions in these areas tending to be later in epidemic time relative to less urban counties. Also, counties in our data set reveal that the total visits to POI did not reduce much more after stay-home orders were adopted across all the rural-urban spectrum, as the total visits during the study time period had often reached close to its minimum observed level at the time a stay-at-home order was instituted.

With model specifications specifically tailored to various dimensions of county-level heterogeneity, the suite of statistical models fit indicated that the timing of mobility reductions was more important for dictating changes in the daily deaths trajectories than the timing of official stay-at-home orders for more urban counties, while the uncertainty of the results for rural counties are rendered inconclusive. The analysis of mobility interventions indicated that hypothetically shifting mobility decreases

[2] An alternative analysis evaluating impacts of action at different numbers of days earlier or later (not shown) produced conceptually similar results; the earlier the action, the more reduction in cumulative deaths in urban counties, with more pronounced impacts from the *Mobility model* than from the *Stay-at-home model*.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

ten days earlier would have reduced deaths, with hypothetically later mobility intervention resulting in more deaths. In contrast, the analysis of stay-at-home orders showed less pronounced impact of intervention timing, indicating some benefit of early action but more limited impact of delayed action relative to the observed timing of the stay-at-home order. These results were corroborated by the *Double intervention model* that estimated little or no impact of stay-at-home timing after accounting for the timing of mobility decrease. Taken together, these results point towards stay-home orders having little benefit above-and-beyond that induced by the persistent drop in mobility. However, interpretation of the relative effectiveness of mobility versus policy interventions generally relies on an interplay between these two types of "interventions" that is not fully resolved in the present work. The relative timing of mobility decreases relative to timing of stay-at-home orders makes clear that there were other drivers of behavior change. We point out that since our mobility measure is determined by visits to all points of interest, other polices such as school and restaurant closures inevitably impacted this measure to some degree that is separate from the stay-at-home order. Parsing mobility decreases attributable to specific policies other than stay-at-home order is important future work. Additional influence of non-policy drivers of behavior change such as awareness-driven voluntary actions on visits and other mobility metrics is likely to have also played a role in a manner not captured by the present analysis. This is not to say that the stay-at-home orders had no effect, even though total visits had often decreased more than 60% and close to its minimum in advance of official stay-at-home orders, these orders likely served to maintain low mobility after its initial decrease. Using the *Double intervention* model to attribute the entirety of impact on deaths to the timing mobility reductions over that of stay-at-home orders would rely on the tenuous assumptions that mobility reductions did not impact the timing of stay-at-home orders and that stay-at-home orders played no role in the apparent maintenance of reduced mobility beyond the date of reaching a 50% reduction. For example, work in [Abouk & Heydari, 2020] describes how, at the state level, stay-home orders increased the actual presence at home, a metric of mobility that did not seem as sound for use at the county level with the data sources contained herein. What's more, the timing of official policy interventions may well have been influenced by observed mobility if, for example, policy makers were prompted to adopt an official order to continue behavior changes that were already occurring. An analysis that fully resolves the potential confounding and mediating effects that may result from the influence of mobility reductions on the timing of stay-at-home orders (or *vice versa*) is beyond the scope of this analysis.

An important feature of the analysis is that its focus on the relative effectiveness of intervention timing does not characterize the relative effectiveness of implementing vs. not implementing the policies or behavior changes. The method pursued here intentionally focuses posterior predictions that perturb the observed trajectory around only a relatively short window spanning the observed intervention dates, specifically avoiding the type of extrapolation that would be required, for example, to learn what might have happened absent a policy or during time frames that expand substantially outside the range of intervention dates. Furthermore, the analysis is unable to explicitly account for differences in the execution of the official orders. In truth, execution of policies or implications of mobility changes may well vary across the US, with the present analysis unable to capture such variation beyond that which may be captured by NCHS urban-rural classification. Finally, the inferences here are valid insofar as the statistical model with quadratic time trends represents a reasonable approximation to the shape of the death trajectories after the initial introduction of cases and for a limited time frame afterward. Reproducing the full epidemiological cycle is not the intention of the model; we only attempt to quantify intervention impacts in the time frame surrounding observed stay-at-home and mobility reduction dates and during the early epidemic stages.

Intentionally exchanging explicit epidemic dynamics for a statistical model focusing on a relatively reduced time frame is an important feature of the analysis. While mechanistic approaches are specifically tailored to echo detailed nuances of epidemic growth, we intentionally simplify this nuance with a quadratic time function in exchange for the ability to use the mechanics of hierarchical regression modeling to ingest county-level features and links to county-specific external factors, all while continuing to borrow information across all counties. The statistical evidence generated from a model such as this is designed to complement that obtained from more traditional mechanistic epidemic models and, in fact, our estimates of averted deaths align with the evidence from the metapopulation SEIR model used in Pei et al. [2020] in that both analyses estimate approximately 50% fewer deaths from shifting epidemic control policies earlier by 7-10 days.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Our work was specifically designed to surmount the challenges of disentangling the intertwined local characteristics and events that unfolded around the time of the first wave of COVID-19 interventions. The limitations of this analysis notwithstanding, we move a step closer into parsing these events by providing evidence for the timing of mobility-related behavior changes as an important determinant of local daily COVID-19 deaths. These results point towards the need to investigate how official reopening policies and other policies that varied across counties interplay with changes in mobility beyond the time frame considered here and into the later phases of the US COVID-19 epidemic.

References

Rahi Abouk and Babak Heydari. The immediate effect of covid-19 policies on social distancing behavior in the united states. *Available at SSRN*, 2020.

Charles Courtemanche, Joseph Garuccio, Anh Le, Joshua Pinkston, and Aaron Yelowitz. Strong social distancing measures in the united states reduced the covid-19 growth rate: Study evaluates the impact of social distancing measures on the growth rate of confirmed covid-19 cases across the united states. *Health Affairs*, pp. 10–1377, 2020.

Klaus Desmet and Romain Wacziarg. Understanding spatial variation in covid-19 across the united states. Technical report, National Bureau of Economic Research, 2020.

Zhanwei Du, Xiaoke Xu, Lin Wang, Spencer J Fox, Benjamin J Cowling, Alison P Galvani, and Lauren Ancel Meyers. Effects of proactive social distancing on covid-19 outbreaks in 58 cities, china. *Emerging infectious diseases*, 26(9):2267, 2020.

Eric M Feltham, Laura Forastiere, Marcus Alexander, and Nicholas A Christakis. No increase in covid-19 mortality after the 2020 primary elections in the usa. *arXiv preprint arXiv:2010.02896*, 2020.

Ben Goodrich, Jonah Gabry, Imad Ali, and Sam Brilleman. rstanarm: Bayesian applied regression modeling via stan. *R package version*, 2(4):1758, 2018.

D Ingram and S Franco. NCHS urban-rural classification scheme for counties. *Vital and Health statistics. Series 2, Data Evaluation and Methods Research*, (154):1–65, January 2012. ISSN 0083-2057, 2333-0872. URL `https://europepmc.org/article/med/22783637`.

Yothin Jinjarak, Rashad Ahmed, Sameer Nair-Desai, Weining Xin, and Joshua Aizenman. Accounting for global covid-19 diffusion patterns, january-april 2020. Technical report, National Bureau of Economic Research, 2020.

John Hopkins University. Coronavirus resource center, 2020. URL `https://coronavirus.jhu.edu/us-map`.

Rolly Kapoor, Haedong Rho, Kinpritma Sangha, Bhavyaa Sharma, Ajay Shenoy, and Guanghong Xu. God is in the rain: The impact of rainfall-induced early social distancing on covid-19 outbreaks. *Available at SSRN 3605549*, 2020.

Sadiya Khan, Megan McCabe, Amy Krefman, Lucia C Petito, Xiaoyun Yang, Kiarri Kershaw, Lindsay Pool, and Norrina B Allen. A county-level susceptibility index and coronavirus disease 2019 mortality in the united states: A socioecological study. *medRxiv*, 2020.

Benjamin D. Killeen, Jie Ying Wu, Kinjal Shah, Anna Zapaishchykova, Philipp Nikutta, Aniruddha Tamhane, Shreya Chakraborty, Jinchi Wei, Tiger Gao, Mareike Thies, and Mathias Unberath. A County-level Dataset for Informing the United States' Response to COVID-19. April 2020.

Nan M Laird and James H Ware. Random-effects models for longitudinal data. *Biometrics*, pp. 963–974, 1982.

Stephen A Lauer, Kyra H Grantz, Qifang Bi, Forrest K Jones, Qulu Zheng, Hannah R Meredith, Andrew S Azman, Nicholas G Reich, and Justin Lessler. The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine*, 172(9):577–582, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Zhixian Lin and Christopher M Meissner. Health vs. wealth? public health policies and the economy during covid-19. Technical report, National Bureau of Economic Research, 2020.

Sen Pei, Sasikiran Kandula, and Jeffrey Shaman. Differential effects of intervention timing on covid-19 spread in the united states. *Science Advances*, 2020.

Nick Wilson, Amanda Kvalsvig, Lucy Telfar Barnard, and Michael G Baker. Case-fatality risk estimates for covid-19 calculated by using a lag time for fatality. *Emerging infectious diseases*, 26(6):1339, 2020.

Spencer Woody, Mauricio Garcia Tec, Maytal Dahan, Kelly Gaither, Michael Lachmann, Spencer Fox, Lauren Ancel Meyers, and James G Scott. Projections for first-wave covid-19 deaths across the us using social-distancing measures derived from mobile phones. *medRxiv*, 2020.

Xiaobo Yang, Yuan Yu, Jiqian Xu, Huaqing Shu, Hong Liu, Yongran Wu, Lu Zhang, Zhui Yu, Minghao Fang, Ting Yu, et al. Clinical course and outcomes of critically ill patients with sars-cov-2 pneumonia in wuhan, china: a single-centered, retrospective, observational study. *The Lancet Respiratory Medicine*, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

A Additional Details on Timing of Policies and Mobility Reductions

A.1 Other Policy Interventions

Amid the initial spread of COVID-19, an array of non-pharmaceutical interventions including bans on gatherings, closures of schools and restaurants, and stay-at-home orders were implemented from mid-March to early April across the US. In most cases, these interventions applied to entire states, which limited the variability in calendar timing across policies across metro, micropolitan and non-core counties.

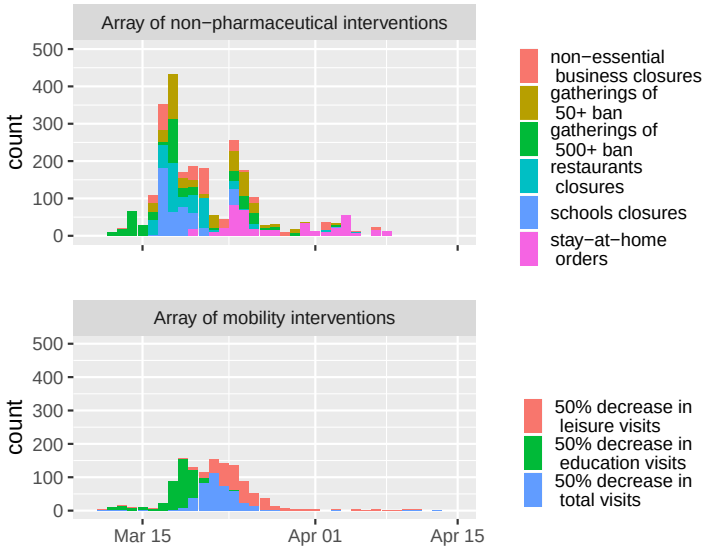


Figure 4: Adoption timeline of an array of both policy and mobility interventions. On the top, counts of dates for six different NPIs, such as stay-home orders, school closures and other more lenient measures, are stacked. On the bottom, counts of days when total visits reached a 50% decrease are included. To illustrate the components that make up total visits, the days "education" and "leisure" visits decreased 50% are also shown; with the education classification covering schools and colleges, and leisure covering restaurants, bars, parks and museums.

Figure 4 shows the number of counties that adopted each of the six policy interventions on a given day. Most bans and closures occurred in the same small window of time and were preceded by stay-at-home orders, which lagged other policies by 8 days in average. No statewide official orders to shelter-in-place were issued in North Dakota, Nebraska and Arkansas.

A.2 Additional Mobility Details

Most of the 356 counties with mobility interventions exhibited a steady decrease in mobility within a window of ten days, with March 22 being the average date at which counties reached the 50% decrease from baseline visits to all POIs, corresponding to the definition of a mobility-based intervention. Figure 4, shows the distribution of dates on which counties mobility-intervention, depicted alongside the dates that counties achieved a 50% decrease in visits to two sub-categories of POIs: 1) schools and colleges and 2) leisure destinations such as restaurants, bars, parks and museums.

To further illustrate how the timing of mobility interventions compared to that of policy interventions among the 336 counties with both, Figure 5 depicts the decrease in mobility (relative to baseline) that was observed on the day a stay-at-home order took effect. All counties had already shown a marked reduction, with micropolitan and non-core counties showing a reduction of 66% on the day of the policy, and other more urban counties having even more reduced mobility, with an average reduction of 70% mobility when the a stay-at-home order took effect. Thus, all counties indicate

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

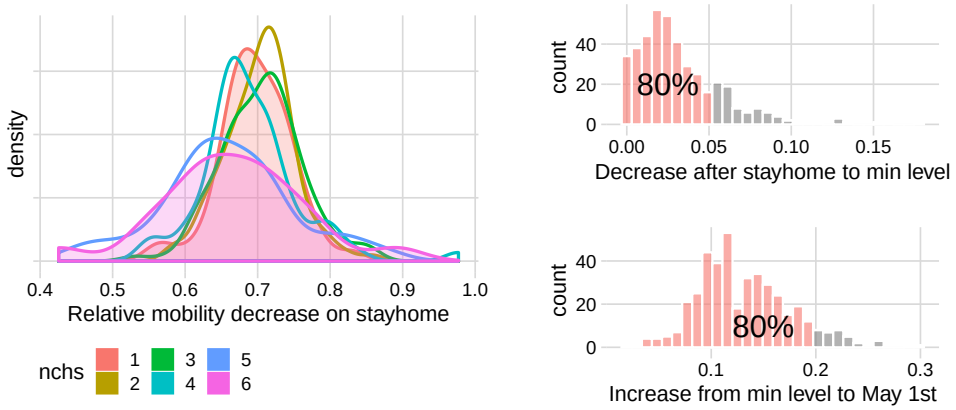


Figure 5: Left: Shows the relative decrease in mobility with respect to the baseline on the day the stay-at-home order was enacted. Right top: Shows the additional decrease in mobility since stay-home up to each county's minimum mobility level. Right bottom: Contains the distribution of the mobility increase beginning at each county's minimum levet up to May 1st.

the potential for behavior and mobility changes to have impacted epidemic dynamics in advance of official stay-home.

Along with this, reductions in mobility were close to a level of saturation when stay-home went into effect with total visits to POI not reducing much more. The distribution on the right panel in Figure 5 shows that for 80% of the counties in our dataset, the mobility levels observed the day stay-home stay-home orders were adopted, were at most 5% above their absolute minimum levels. After the minimum levels were reached, mobility slowly started to move upwards with 80% of counties lingering no more than 20% above those levels on May 1st.

B Details of the Statistical Modeling Approach

B.1 Bayesian Hierarchical Negative Binomial Model Specification

The statistical models for county-specific death trajectories rely on two important features. First, we define local "epidemic time" in terms of the number of days elapsed since local arrival of the epidemic, defined as the date at which a county reached a deaths threshold of 3 deaths per 10 million residents. Second, because reports of daily deaths often exhibit erratic day-to-day variation across counties (e.g., one county may report most cases on Monday while another county on Wednesday), we fit the statistical models described below to the 7-day centered rolling average of daily deaths, effectively removing day-of-week reporting effects which are not important for the purpose of this analysis.

More precisely, let y_{it} be the 7-day moving average of daily deaths observed in county i at time t, where t represents the number of days since the deaths threshold was reached. First, we assume

$$y_{it} \sim \text{NegBin}(\lambda_{it}, r), \tag{1}$$

where the parameterization is such that $\mathbb{E}[y_{it}] = \lambda_{it}$ and $\mathbb{V}[y_{it}] = \lambda_{it}(1 + \lambda_{it}/r)$; the unknown parameter r indicates over-dispersion: the smaller it is, the higher the variance. For λ_{it} we use a log link and a per capita normalization. Let N_i be the population of county i, then

$$\log\left(\frac{\lambda_{it}}{N_i}\right) = f(\boldsymbol{x}_i, d_i, t) + r_i(t)$$

where t is the number of days since the threshold date, $\boldsymbol{x}_i$ are the county-level features, and d_i represents intervention timing as the number of days between the threshold date and the date the intervention was enacted. r_i is a county-specific residual time function that captures the portion of the curve not explainable by the county-level features and intervention contained in f.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

To specify the form of f we use a quadratic polynomial in t that changes shape after introduction of the intervention and across different levels of county-level features. Specifically,

$$\begin{aligned} f(\boldsymbol{x}_i, d_i, t) := &\sum_{k=0}^{2} \left[\alpha_k + \boldsymbol{x}_{i,pre}^\top \boldsymbol{\beta}_k\right] t^k \\ &+ \sum_{k=1}^{2} \left[\eta_k^I + d_i \eta_k^d + \boldsymbol{x}_{i,post}^\top \boldsymbol{\eta}_k^c\right] \mathbb{1}(t \geq d_i)(t - d_i + 1)^k. \end{aligned} \tag{2}$$

The first summation in (2) specifies the polynomial time trend during the pre-intervention period: $(\alpha_0, \alpha_1, \alpha_2)$ can be thought of as "baseline" intercept, linear, and quadratic polynomial parameters, which are each shifted by $(\beta_0, \beta_1, \beta_2)$ according to county-level features in $\boldsymbol{x}_{i,pre}$. That is, each county's pre-intervention deaths trajectory is dictated in part by the features of that county, where the features contained in $\boldsymbol{x}_{i,pre}$ are: NCHS classification, the percent of college attendees, the percent aged 65 and older, the percent of black, and the percent hispanic of the county.

The second summation in (2) specifies the bend in the quadratic death trajectory after the introduction of the intervention. Here, d_i represents the date that an intervention occurred in county i plus a lag of 12 days. The term $\mathbb{1}(t \geq d_i)$ represents the indicator equal to 1 when epidemic time surpasses the time after the intervention lag, and $(t - d_i + 1)$ is the days elapsed since the intervention lag. The parameters η_k^I for $k = 1, 2$ can be thought of (respectively) as the "baseline" shift in the linear, quadratic terms of the polynomial after the introduction of the intervention. Analogously η_k^d can be thought as shifts in the linear and quadratic terms according to the timing of the intervention, implying different post-intervention trajectory shapes for counties that intervened at different points in their local epidemic time. Similarly, η_k^c dictate shifts in the linear and quadratic terms according to the covariates in $\boldsymbol{x}_{i,post}$, where $\boldsymbol{x}_{i,post}$ in the present analysis includes only the NCHS county classification.

Note the omission of the zero-degree polynomial terms in the post-intervention terms of the second summation in (2), intentionally omitting mean shifts in the polynomial function upon the introduction of the intervention. This is done to specify continuity in the polynomial trend across the introduction of the intervention, as we would not expect the intervention to initiate an immediate jump or drop in daily deaths.

Finally, the idiosyncratic factor r_i is specified using county-specific random effects [Laird & Ware, 1982] for each polynomial coefficient

$$r_i(t) := \sum_{k=0}^{2} \xi_k^{(i)} t^k, \quad (\xi_0^{(i)}, \xi_1^{(i)}, \xi_2^{(i)})^\top \sim N(\mathbf{0}, \Sigma),$$

implying that each county is modeled to have it's own pre-intervention polynomial time trend.

The above model specification is used for the *Stay-at-home model* and *Mobility model* described in the main text. The *Double intervention model* has a three-fold specification, where equation (2) is augmented with a third summation that has the same construction as the second summation. The second term specifies the bend after the mobility intervention, and the third term captures the additional or residual effect of the stay-at-home order that is not captured by the change in mobility.

B.2 Posterior Inference

We base inference on simulations from the posterior distributions of the parameters in models of the form (2) simulations from the models. To evaluate impacts of intervention timing, we simulate from the posterior predictive distribution of deaths from each county's time polynomial, but where the intervention timing, d_i, is replaced with alternative hypothetical timing corresponding to earlier or later intervention. Specifically, under different values of d_i, we take a sample from the posterior predictive distribution for each county and take group averages by NCHS category for each time t. Group-average curves are calculated by averaging point-wise posterior predictive quantities among counties within the same NCHS category. In addition, we compute the number of days since the threshold at the peak and the number of per capita deaths at the peak for each one of the average curves per NCHS. We repeat this process for 1,000 posterior samples to produce a set of 1,000 aggregate statistics and curves for each NCHS.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

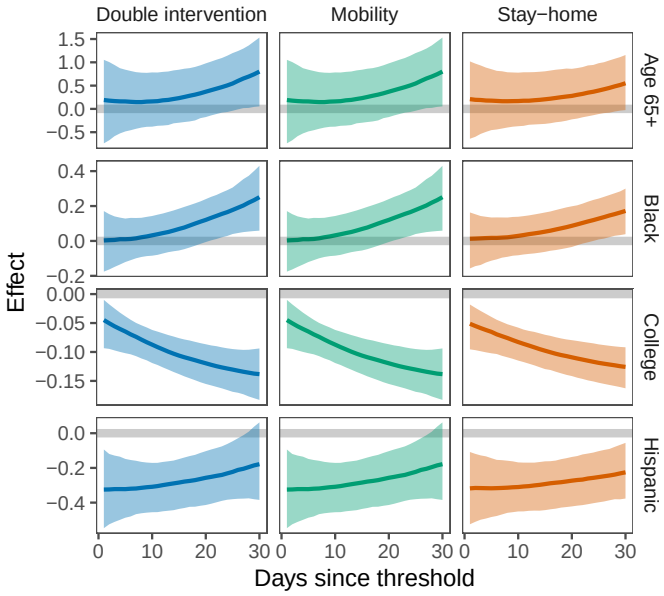


Figure 6: Shows the time-varying of the confounding variables. For each covariate the figure shows median and 90% posterior credible intervals. The zero-th line is shown in gray in each subplot. Effects that do not consider contain the zero-th line are statistically significant.

Note that, since model (2) is specified to borrow information across all counties, results for a given county may be different in all three of the above models, even if the mobility and stay-at-home intervention dates coincide.

B.3 Drivers of Pre-intervention Trajectory Heterogeneity

With the rationale that different county characteristics are associated with differently-shaped local epidemic death curves, the model from Section 2.4 permits individual counties to have pre-intervention daily death trajectories that explicitly depend on the county-level covariates: NCHS Urban-Rural Classification Scheme for Counties and percent of residents that are Black, Hispanic, aged 65 years and older, and attending college. Note that given the specification of the model, the differences that are due to the timing of the intervention influence are captured separately.

For both the stay-at-home and mobility models, we find that NCHS category, percent of college student residents, and percent of Hispanic residents had significant impact on death trajectories. In the double intervention model, these same characteristics plus the percent of Black residents had significant impact. A table presenting the coefficient estimates associated with each county characteristic and its interaction with the linear and quadratic time polynomial terms appears in Table 3 in the Appendix. Note that the coefficient estimates in the table are not easily interpreted nor directly comparable in magnitude across the three models since the time polynomial in each model is orthogonalized using the `poly` function in `R`, but judgments of statistical significance remain valid. Figure 6 summarizes the covariate impacts across the different models. The figure shows posterior estimates and uncertainty intervals of the interaction between time polynomials and coefficients effects. More precisely, for each covariate j with corresponding polynomial coefficients $\boldsymbol{\beta}^j = (\beta_0^j, \beta_1^j, \beta_2^j)$ in the regression model (2), we take 1,000 samples from their posterior distribution and for each sample compute an estimate of the total contribution of covariate j through time given by $c^j(t) = \sum_{k=0} t^k \beta_k^j$. Figure 6 is showing the pointwise median and quantiles at each time t of the posterior samples of the total contribution $c^j(t)$. Observe that all models yield similar estimates for the confounding variables effects, but the percent of Hispanic residents and college students are the most statistically significant variables, as evidenced by the fact that the curves do not contain the zero line. Note that these results should be interpreted as descriptive summaries of how these daily death trajectories varied across levels of these characteristics. For example, the estimates for

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

	Stay-home only		Mobility only		Double intervention	
Parameter	*Mean*	*[5%, 95%]*	*Mean*	*[5%, 95%]*	*Mean*	*[5%, 95%]*
Degree=0						
(Intercept)	-11.3*	[-15.8, -6.79]	-9.71*	[-14.4, -5.42]	-10*	[-14.6, -5.38]
NCHS-2	0.27	[-0.11, 0.67]	0.87*	[0.44, 1.26]	1*	[0.54, 1.43]
NCHS-3	0.01	[-0.42, 0.5]	0.55*	[0.11, 0.99]	0.65*	[0.2, 1.09]
NCHS-4	-0.17	[-0.74, 0.4]	0.03	[-0.63, 0.65]	0.25	[-0.37, 0.9]
NCHS-5	0.4	[-0.34, 1.08]	0.9*	[0.29, 1.54]	1.04*	[0.28, 1.74]
NCHS-6	0.12	[-0.96, 1.18]	-0.64	[-2.48, 1.09]	0.22	[-1.77, 2.2]
% in college	-0.09*	[-0.12, -0.07]	-0.08*	[-0.11, -0.06]	-0.08*	[-0.11, -0.06]
% black	0.07	[-0.03, 0.16]	0.07	[-0.02, 0.18]	0.11*	[0.01, 0.21]
% hispanic	-0.29*	[-0.45, -0.15]	-0.36*	[-0.5, -0.22]	-0.34*	[-0.47, -0.2]
% +65 age	0.28	[-0.26, 0.82]	0.07	[-0.47, 0.57]	0.04	[-0.54, 0.62]
Degree=1						
(Linear)	-19.7	[-313, 274]	20.7	[-217, 266]	-46.8	[-287, 168]
NCHS-2	-37.6*	[-70.6, -3.77]	8.29	[-19.6, 35.1]	11.5	[-18.7, 42.7]
NCHS-3	-50*	[-84.3, -18]	-15.5	[-44.3, 14]	-14.8	[-49.5, 15.3]
NCHS-4	-102*	[-148, -61.4]	-91.2*	[-134, -49]	-86.3*	[-130, -42.5]
NCHS-5	-98.5*	[-154, -41.7]	-42.7*	[-81.4, -2.42]	-63.3*	[-116, -13.1]
NCHS-6	-140*	[-229, -52.1]	-151*	[-252, -59.9]	-62.2	[-223, 95.7]
% in college	-2.56*	[-4.2, -0.87]	-1.89*	[-3.24, -0.43]	-1.74*	[-3.09, -0.43]
% black	7.39	[-0.14, 14.6]	1.85	[-3.64, 7.67]	5.43	[-0.38, 11.1]
% hispanic	4.62	[-4.96, 14.2]	5.4	[-1.18, 12.6]	4.6	[-2.4, 11.8]
% +65 age	15.9	[-19, 48.7]	13	[-16.3, 40.3]	19.3	[-6.67, 46.6]
Degree=2						
(Quadratic)	-172	[-407, 52.9]	-20.7	[-176, 130]	-4.05	[-176, 154]
NCHS-2	-7.46	[-28.2, 12]	0.43	[-14.5, 14.7]	0.38	[-14.3, 14.3]
NCHS-3	4.2	[-17.2, 25.5]	-2.64	[-17.3, 11.2]	0.51	[-14.7, 16]
NCHS-4	-10.4	[-38.2, 16.7]	-15.7	[-35.3, 3.34]	-15.6	[-36.5, 4.3]
NCHS-5	-13	[-45.3, 20.9]	9.42	[-9.01, 28]	-4.78	[-30, 18.8]
NCHS-6	-47.2*	[-94.3, -0.89]	-2.15	[-52, 46.9]	36.6	[-42.5, 116]
% in college	0.61	[-0.63, 1.9]	-0.04	[-0.99, 0.83]	-0.07	[-0.95, 0.85]
% black	2.65	[-3.14, 8.56]	-2.77	[-6.7, 1.12]	-1.08	[-5.09, 3.28]
% hispanic	2.32	[-4.39, 9.12]	1.37	[-3.42, 5.89]	0.93	[-3.68, 5.96]
% +65 age	9.06	[-17.2, 36.8]	1.14	[-16.6, 19.7]	-2.31	[-21.7, 17.7]

Table 3

Hispanic population are first negative but gradually increase towards zero. One could hypothesize that this relationship could be attributed to the fact that counties with higher Hispanic population, particularly southern counties, also had more extant COVID safety awareness since the epidemic took hold later in calendar time.

Importantly, while we only evaluated explicit pre-intervention heterogeneity with respect to this relatively small subset of covariates that were included in the model, the inclusion of county-specific random effects is designed to capture additional heterogeneity in death curves attributable to other factors.

B.4 Additional Summaries of Trajectory Characteristics under Hypothetical Intervention Timing

Table 4 provides numerical estimates (median and inter-quartile range (IQR) intervals) of the average number of days since local epidemic initiation (3/10 million death threshold) to the peak in daily deaths and the height of the daily death peak under each intervention timing. The table suggests statistically significant differences in NCHS 1-3 between the actual and early estimates as well as between actual and late estimates. For example, for both models early intervention in NCHS 1 is associated with a change in deaths at the peak of approximately 2 less deaths per 1 million, and an earlier peak of 5 days for the stay-home model and 8 days for the mobility model. Differences in NCHS 4-6 are not statistically significant for either model since the IQR intervals strongly overlap.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

	Days since threshold at peak			Deaths per 1 million at peak		
NCHS	*Early*	*Actual*	*Late*	*Early*	*Actual*	*Late*
1	21 (2)	26 (1)	28 (2)	1.59 (0.42)	3.50 (0.16)	4.94 (0.89)
2	19 (4)	23 (2)	23 (1)	1.56 (0.49)	2.95 (0.29)	3.59 (0.52)
3	15 (2)	21 (2)	23 (2)	1.02 (0.25)	1.99 (0.16)	2.47 (0.41)
4	14 (2)	19 (2)	20 (2)	1.77 (0.58)	2.46 (0.33)	2.19 (0.49)
5	13 (4)	18 (2)	20 (3)	3.08 (1.40)	4.97 (0.79)	4.84 (1.57)
6	22 (16.25)	23 (14)	18 (4)	18.47 (54.90)	12.12 (9.74)	6.65 (4.17)

(a) Stay-at-home order model

	Days since threshold at peak			Deaths per 1 million at peak		
NCHS	*Early*	*Actual*	*Late*	*Early*	*Actual*	*Late*
1	16 (2)	24 (1)	27 (1)	1.30 (0.37)	3.20 (0.19)	5.66 (1.13)
2	11 (2)	19 (2)	24 (1)	1.02 (0.28)	3.33 (0.29)	7.74 (1.50)
3	10 (2)	18 (1)	21 (2)	0.92 (0.25)	2.31 (0.18)	4.15 (0.88)
4	11 (3)	16 (1)	18 (2)	2.75 (1.14)	2.96 (0.41)	2.41 (0.80)
5	7 (2)	16 (3)	25 (5)	3.27 (1.20)	6.28 (1.03)	9.65 (3.44)
6	11 (12)	21 (25.25)	33 (25)	35.13 (72.47)	16.97 (49.00)	8.34 (71.33)

(b) Mobility decrease model

Table 4: Summary statistics for fitted daily deaths curves per NCHS. The table shows the median and interquartile range (IQR) in parenthesis for the peak of the average curve per NCHS from 1000 posterior samples. Significant effects can be seen in NCHS 1-3.

B.5 County Level Model Posterior Predictions

Figure 7 shows the observed and hypothetical trajectories for three selected counties counties: King County, Washington; Kings County, New York; and Jefferson County, Louisiana. In King County, Washington, a hypothetical late adoption of a stay-home order does not appear to have a significant impact on the post-intervention death curve trajectory, suggesting that its curve had already flattened by the time the policy was adopted, likely as a result of awareness brought by national attention as the first US cases emerged in this county. In Kings County, New York, the 50% mobility decrease was reached three days after the stay-home was adopted and yet the models' estimation show that a late mobility-intervention would have a significantly higher peak than a late policy-intervention and that it would more dramatically bend the post-intervention trend. Finally, in the case of Louisiana's Jefferson County, even if both the policy and mobility interventions happened exactly on the same day, the late policy-intervention counterfactual has a considerable overlap with the original fit contrasting with the late mobility-intervention counterfactual.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.24.20238055; this version posted November 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

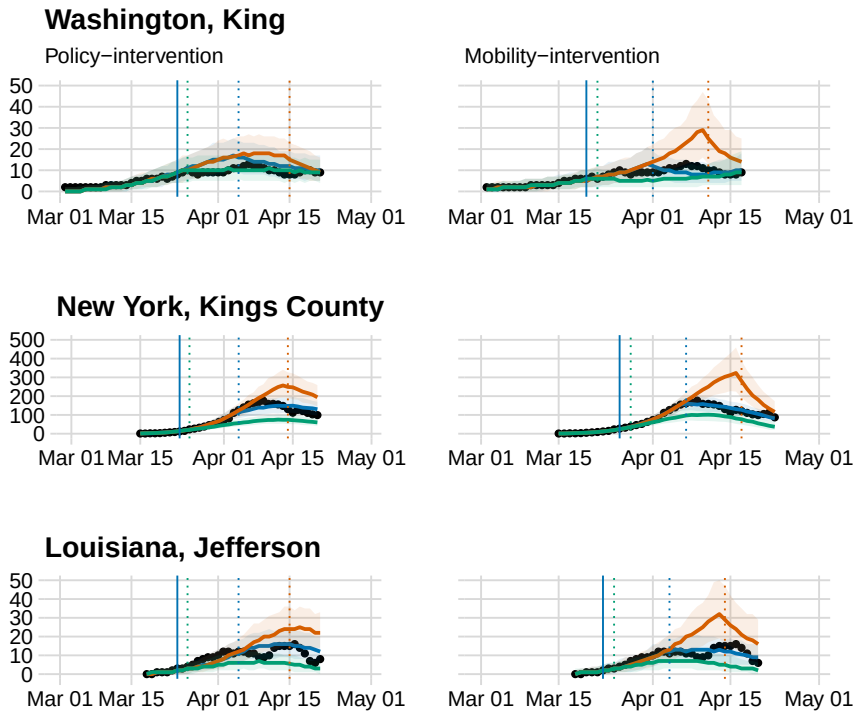


Figure 7: The early and late intervention timing counterfactuals for three counties. We simulate 1000 draws from the Bayesian posterior predictive distribution of the (log) expected counts for each county, and take point-wise medians and 5%/95% quantiles for each t: in blue the fitted credible intervals; in green (red) the early (late) credible bands around the counterfactual medians. The three plots on the left show estimates using the stay-at-home model, and on the right estimates from the mobility model. The vertical dotted lines depict the actual and hypothetical date of the intervention plus the intervention lag.