Interplay of global multi-scale human mobility, social distancing, government interventions, and COVID-19 dynamics

Aniruddha Adiga*[a], Lijing Wang*[a,b], Adam Sadilek*[c], Ashish Tendulkar*[c], Srinivasan Venkatramanan[a], Anil Vullikanti[a,b], Gaurav Aggarwal[c], Alok Talekar[c], Xue Ben[c], Jiangzhuo Chen[a], Bryan Lewis[a], Samarth Swarup[a], Milind Tambe[c],Madhav Marathe†[a,b]

[a]Biocomplexity Institute and Initiative, University of Virginia, Charlottesville, VA, USA
[b]Department of Computer Science, University of Virginia, Charlottesville, VA, USA
[c]Google, Mountain View, CA, USA

Abstract

This work quantifies the impact of interventions to curtail mobility and social interactions in order to control the COVID-19 pandemic. We analyze the change in world-wide mobility at multiple spatio-temporal resolutions – county, state, country – using an anonymized aggregate mobility map that captures population flows between geographic cells of size 5 km^2. We show that human mobility underwent an abrupt and significant change, partly in response to the interventions, resulting in 87% reduction of international travel and up to 75% reduction of domestic travel. Taking two very different countries sampled from the global spectrum, we observe a maximum reduction of 42% in mobility across different states of the United States (US), and a 68% reduction across the states of India between late March and late April. Since then, there has been an *uptick* in flows, with the US seeing 53% increase and India up to 38% increase with respect to flows seen during the lockdown. As we overlay this global map with epidemic incidence curves and dates of government interventions, we observe that as case counts rose, mobility fell – often *before* stay-at-home orders were issued. Further, in order to understand mixing within a region, we propose a new metric to quantify the effect of social distancing on the basis of mobility. We find that population mixing has decreased considerably as the pandemic has progressed and are able to *measure* this effect across the world. Finally, we carry out a counterfactual analysis of delaying the lockdown and show that a one week delay would have doubled the reported number of cases in the US and India. To our knowledge, this work is the first to model in near real-time, the interplay of human mobility, epidemic dynamics and public policies across multiple spatial resolutions and at a global scale.

1 Introduction

The COVID-19 pandemic is arguably the most acute public health emergency since the 1918 influenza pandemic. It has already infected over 6.67 million people and resulted in 391K deaths across the globe[1]. The economic impact is expected to be 3-10 trillion dollars[2]. The pandemic has affected almost every country in the world[3] and has resulted in an unprecedented response by governments across the world to control its spread. Pharmaceutical interventions are not generally available at this stage (with the exception of remdesivir under FDA's expanded access [1]) and thus, countries have had to rely exclusively on behavioral

*Equal contribution
†Corresponding author
[1]Source: `https://covid19.who.int/` as of June 04, 2020.
[2]`https://www.un.org/development/desa/en/news/policy/wesp-mid-2020-report.html`
[3]Detailed global cases data available at `https://nssac.bii.virginia.edu/covid-19/dashboard/`

interventions that involve some form of social distancing. The rapid spread of the pandemic has forced countries to institute strict social distancing measures. Each social distancing policy is characterized by: (*i*) when and how gradually it started, (*ii*) the length of time for which it was enforced, (*iii*) the scope and pervasiveness, and (*iv*) the stringency (total lockdown versus stay-at-home advisories). For example, across the US, social distancing policies were instituted at the state level in a progressive manner: a declaration of a state of emergency by many states, followed by school closures, with the final extreme measure being stay-at-home or shelter-at-home orders (SAHO) [2], which saw the closure of recreational centers, parks, restaurant dine-in services, etc. As a contrast, a country-wide lockdown was instituted in India which subsequently led to suspended road, rail, and air transport. Non-essential services and schools were closed and individual-level mobility was severely curtailed.

Evaluating the level of public response to the global, country, and state level restrictions is important to understand COVID-19 dynamics. However, it is important to do this without compromising individual privacy. As pointed out by multiple public health experts and demonstrated in the literature [3, 4], aggregate mobility data, acquired through mobile phone location history, global positioning systems data, direction requests data, etc., indicate a considerable reduction of activity by individuals during the pandemic, and thus act as surrogate data sources to understand compliance with social distancing measures. With appropriate data-sharing policies, these data sources can be used to study social distancing, while also ensuring individual privacy through a combination of anonymization, aggregation and noising techniques that provide the needed privacy guarantees. There have been a number of recent studies along these lines, for example, in China using Baidu data, in the US using mobility data, and at a global scale using airline traffic [5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15]

The Google COVID-19 Aggregated Mobility Research Dataset (cf. Appendix A, henceforth called interchangeably mobility map or flows) provides a global, time-varying anonymized mobility map of flows at a resolution of 5km^2. Figure 1 provides an overview of the volume of local mobility at the level of cell resolution, county, state, and country. The data set has guaranteed differential privacy while capturing mobility flows (MF) at every level of spatio-temporal resolution. The global coverage enables us to undertake multi-scale global analyses of changes in mobility patterns. We combine this data with two additional data layers: data from the University of Virginia COVID-19 Dashboard[4] that provides detailed, global epidemic surveillance data and data on social distancing guidelines in US states and India (at the national level). The integrated layered map provides a new way to assess the impact of changes in global mobility patterns on COVID-19 dynamics. Our analysis can be summarized in the three broad findings described below; more specific findings and additional discussion can be found in Section 2.5.

First, the global analysis reveals that human mobility underwent an abrupt change in response to the COVID-19 pandemic. The abrupt change resulted in a 90% daily reduction in international travel between selected countries[5], about 40-50% daily reduction of aggregate human mobility in the United States (US) and about 60-70% daily reduction of mobility in India during a 30-45 day period starting in late March and ending in late April. Since the first week of May there has been a steady increase in mobility globally. The data provides the first empirical global-scale evidence of mobility reduction and confirms various reports suggesting such a noticeable change.

Second, using the integrated map, we analyze the effects of both a social-distancing advisory and a lockdown using two countries, the US and India. We choose the two countries due to their size, population density, epidemiological context and differences in how policies are being implemented. Also, both countries started implementing social distancing measures at similar times (third and fourth week of March). Thanks to the resolution of the data, our approach captures mobility patterns at a sub-state level (county or administrative level 2) as well. In the US, we observe a high degree of correlation between rising case counts and drop in mobility, with considerable flow reduction occurring before SAHO. In order to quantify mixing within the population, we create a metric called the social distancing index which indicates that across counties of the US, population mixing has reduced with the progression of the pandemic. We also report how inter-state mobility was significantly impacted in both US and India.

Finally, we carry out a counterfactual computational experiment to study the impact of delaying the lockdowns in the US and India. We find that even a one week delay would have caused substantial increase

[4] `https://nssac.bii.virginia.edu/covid-19/dashboard/`

[5] We have selected a few illustrative countries to keep the discussion tractable, which were salient from an epidemiological perspective; additional country data can be reported if requested.

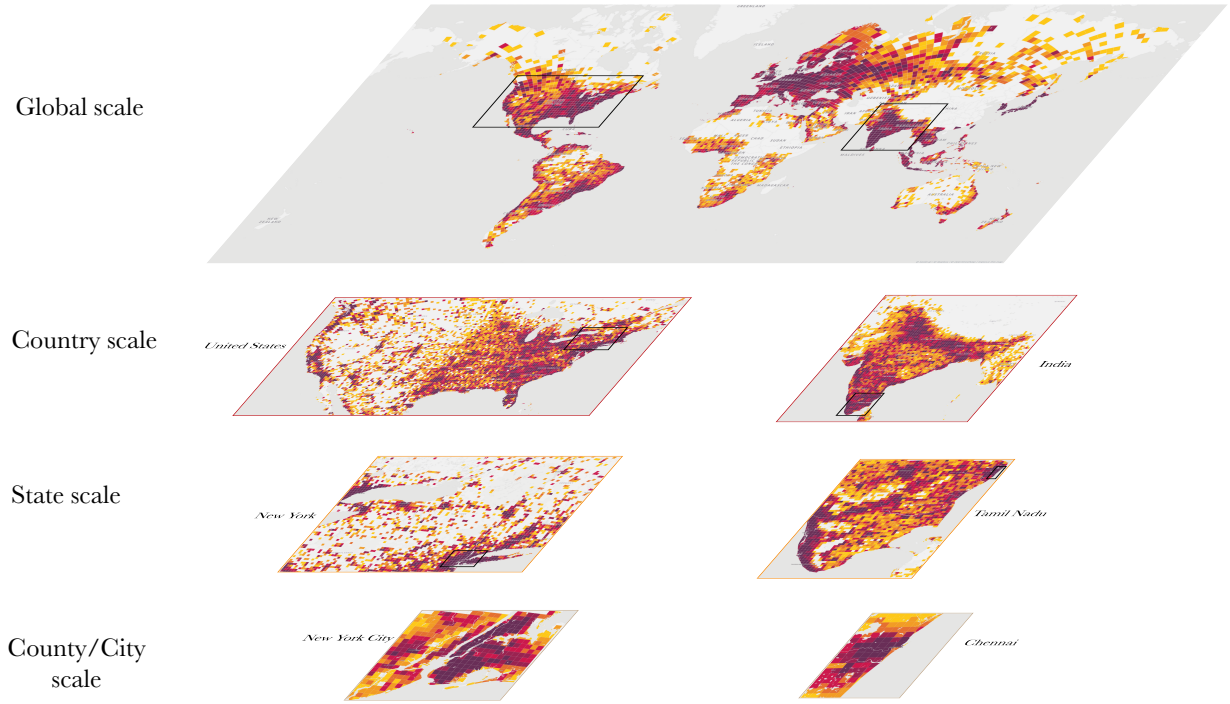


Figure 1: Multiple scales of mobility: mobility flows between 5 km^2 grid cells are aggregated to appropriate spatial resolution for the analyses in this paper. The figure shows progressively the mobility volume at various S2 cell levels (L12 to L6) and geographical scales (county, state, country and globe).

in the number of infections in both the countries. The results confirm the importance of timely lockdown in both the countries.

1.1 Methods

Data sets:

Google COVID-19 Aggregated Mobility Research Dataset, which contains the anonymized relative MF aggregated over users within a 5 km^2 cell. All the flow data is aggregated to the county (administrative level 2) and state (administrative level 1) for the US and India. We refer the reader to Appendix A for a detailed description of the dataset.
COVID-19 surveillance data via the UVA COVID-19 surveillance dashboard [16]. It contains daily confirmed cases and death count worldwide. The data is available at the level of a county in the US and at a state level in India. Daily case counts and death counts are further aggregated to weekly counts.

Metrics:

In order to quantify the effects of social distancing through changes in mobility, we introduce two metrics namely the Flow Reduction Rate (FRR) and Social Distancing Index (SDI). While FRR measures the reduction in *connectivity* of a region to the outside world, SDI measures the change in *mixing* within the region. Additionally we use a case growth rate (CGR) and effective reproductive number (R_{eff}) to quantify the corresponding changes in case incidence.
Flow Reduction Rate (FRR): One way to measure the impact of social distancing is to compare the levels of connectivity before and after the SAHO/lockdowns. Given a set of nodes V, with flows from i to j during week t denoted by $f_{ij}(t)$, we first compute the average outflows during the pre-pandemic period for node $i \in V$ as $\bar{f}_i = \frac{1}{|T_p|} \sum_{t_p \in T_p} \sum_{j \in V} f_{ij}(t_p)$ over first T_p weeks of year 2020 (in our study we considered

the first 6 weeks of 2020). For a given node i, $\mathrm{FRR}_i(t)$ is then defined as

$$\mathrm{FRR}_i(t) = \frac{\sum_{j \in V} f_{ij}(t) - \bar{f}_i}{\bar{f}_i}. \tag{1}$$

This defines a unit-less relative change in outflows from node i for any given week t with respect to $\bar{f}_i$. Henceforth, we omit i and t in the notation. Since MF is mostly symmetric (counting trips in both directions), without loss of generality we work with outgoing flows. Note that FRR is scale-agnostic and can be computed for a county, state or country.

Social Distancing Index (SDI): In order to quantify the *mixing* or movement within a county, we consider the flows between the 5 km^2 cells in it. This is motivated by the fact that, under extreme case of social distancing (i.e., stay-at-home), the inter-cell flows will be significantly reduced. Let V denote the set of cells within a county. Let $\mathbf{F}(t)$ denote the normalized flow matrix of the county at week t with normalized flow from cell i to cell j defined as $F_{ij}(t) = \frac{f_{ij}(t)}{\sum_{j \in V} f_{ij}(t)}$. We compare $\mathbf{F}(t)$ to two matrices, the uniform matrix $\mathbf{U}$ and the identity matrix $\mathbf{I}$. The uniform matrix with entries $U_{ij} = \frac{1}{|V|}$, where $|V|$ is the cardinality or the number of cells in V denotes equiprobable movement between cells and indicates a scenario of high level of mixing or movement between cells. On the other hand, the identity matrix I, with entries $I_{ii} = 1$ and zero otherwise, indicates a scenario where all flows are within a cell and no mixing or flow happens between cells.

The SDI quantifies the closeness of $\mathbf{F}(t)$ to $\mathbf{U}$ and $\mathbf{I}$ and is defined as

$$SDI(t) = \frac{||\mathbf{F}(t) - \mathbf{U}||_2}{||\mathbf{F}(t) - \mathbf{U}||_2 + ||\mathbf{F}(t) - \mathbf{I}||_2}. \tag{2}$$

$SDI(t)$ value close to one indicates the closeness of $\mathbf{F}(t)$ to the identity matrix while a value close to zeros indicates its closeness to the uniform matrix.

Case count growth rate (CGR): Denoting the new confirmed case count at week t as n_t, the CGR of week $t+1$ is computed as $log(n_{t+1} + 1) - log(n_t + 1)$, where we add 1 to smooth zero counts.

Effective reproductive number ($\mathbf{R_{eff}}$) [17]: This is the average number of secondary cases per infectious case in a population made up of both susceptible and non-susceptible hosts. In our study, we use an SEIR model fitted to the normalized case incidence to estimate $\mathrm{R_{eff}}$ as described below.

Simulations:

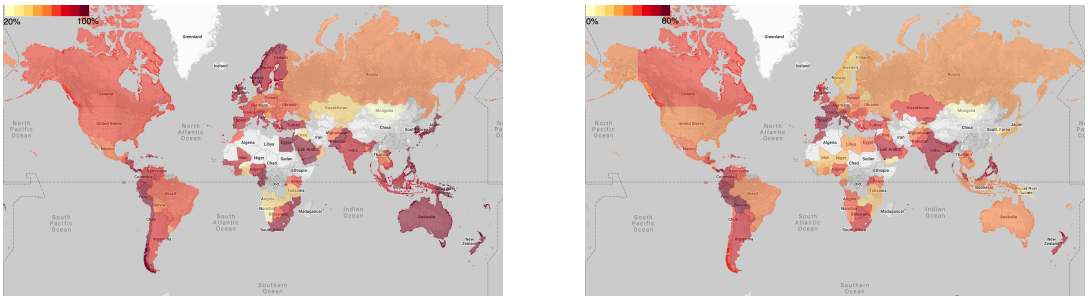
We evaluate the impact of early measures (SAHO in the US and a curfew-like lockdown in India) along similar lines as presented in [18]. As for the disease simulation, we employ a compartmental SEIR model [19, 20] . We set the disease parameters as follows: mean incubation period 5 days, mean infectious period 5 days, delay from onset to confirmation 7 days and case ascertainment rate of 15% [21].

We calibrate a weekly $\mathrm{R_{eff}}$ using simulation optimization to match the new confirmed cases per 100k (smoothed using Savitzky-Golay filter with filter window size of 7) at the state or country level, which is referred as the *normal scenario*. We search for $\mathrm{R_{eff}}$ in the plausible range. To simulate and compare the effect of a delayed intervention, we consider two counterfactual scenarios which involve *one-week delay* and a *two-week delay* in imposing said interventions. Accordingly, we extend the $\mathrm{R_{eff}}$ schedule by persisting a $\mathrm{R_{eff}}$ value in week t for one (or two) more weeks, and mirror the normal schedule with a one (or two) week delay. We then simulate using the modified $\mathrm{R_{eff}}$ schedule to produce the counterfactual simulated epicurve. Monte Carlo simulation is applied to quantify the uncertainty by adding noise to $\mathrm{R_{eff}}$ with noise level of AWGN with $\sigma = 0.025$. We then report the number of cumulative cases avoided as of May 30 for the counterfactual scenarios. We also compare the peak time and peak values of the delay scenarios with respect to the normal scenario for United States.

2 Results

2.1 Global mobility flow analysis

Different countries of the world experienced the effects and reacted differently to the evolving pandemic. Just considering the mobility and comparing the total outflows in January 2020 (as a baseline) with the



(a) (b)

Figure 2: (**Global Data**) Global flow analysis: (a) Inter-country and (b) intra-county MF reduction across the world with flows from January and April in 2020. The MF are aggregated to country level. FRR for inter-country flow is computed with flows satisfying $i \neq j$, while intra-country flow reduction FRR is computed with flows satisfying $i = j$. The darker the color, the stronger the flow reduction is. We observe 74% and 87% inter-country flow reduction for the US and India, respectively. The average intra-country flow reductions for the USA and India are 39% and 72%, respectively.

total outflow in April, we observe a 74% reduction in inter-country mobility in the US, and 87% in India. Figure 2 shows other countries. The domestic mobility flow reduction for the US and India are 39% and 72%, respectively.

In order to study the impact of social distancing policies on human mobility and COVID-19 dynamics in parts of the world, we consider at least one country from each continent (except Antartica) and present their case numbers and respective MF in Figure 3. It is interesting to note that as the pandemic set in, people from these countries started reducing their mobility. The MF reduction in most of the countries started in the week of 2020/03/08-2020/03/14. Further, we observe that most countries show a reduction in flow a week or two *prior* to the formal lockdown announcements[6].

Analysis of the impact of MF on the infection spread in terms of weekly new confirmed cases and cumulative growth rate (CGR) enables us to understand the interplay between mobility and cases. Towards this goal, we present MF and the weekly new confirmed case count (in Figure 3b), and MF reduction with growth of new confirmed cases (in Figure 3c). In all these countries we observe a substantial reduction in MF. The dynamics of the disease spread are not apparent from the weekly number of new cases but a consistent drop in CGR over the subsequent weeks indicates a slowing down of the spread of COVID-19 across all countries.

We now turn to an in-depth analysis of mobility and disease dynamics for the US and India, which in many ways span the spectrum of human mobility, social distancing, government interventions, and COVID-19 dynamics.

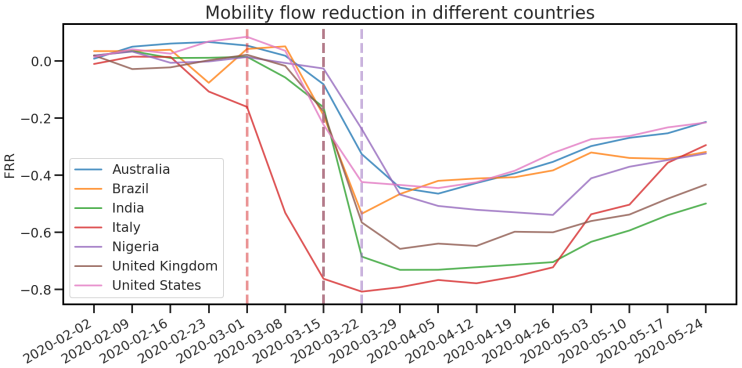
2.2 Research findings for the US

Temporal analysis of MF changes with the COVID-19 progression and government social distancing orders

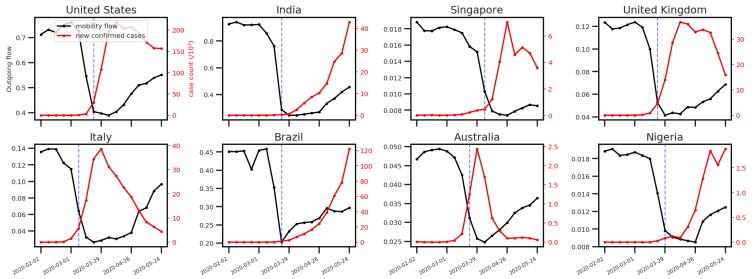
The temporal range is from week 2020/01/19-2020/01/25 (the first confirmed case appeared in the US) to week 2020/05/24-2020/05/30. The analysis is conducted at both state and county levels.

In order to analyze the mobility patterns and the disease progression during the initial phase of the pandemic, we perform a correlation analysis on data from 2020/01/19-2020/01/25 to 2020/05/03-2020/05/09. Figure 4a shows Pearson correlations between weekly county level outgoing flow and new confirmed case count by states. We observe that the correlation varies across counties and shows low (median -0.2 in Utah) to high (median -0.82 in New Jersey) negative correlation between MF and new confirmed case count. We also determine the lags of a state's confirmed case that yield the highest correlation with respective MF and

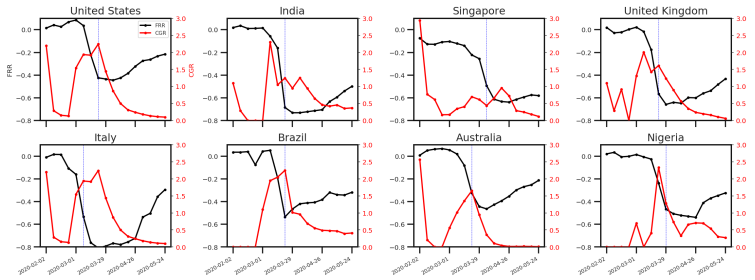
[6]The starting dates of lockdowns in different countries were obtained from the `https://en.wikipedia.org/wiki/COVID-19_pandemic_lockdowns`



(a) Weekly MF reduction rate (FRR).



(b) Weekly MF and new confirmed cases in 8 selected countries around the world.



(c) Weekly FRR and CGR in 8 selected countries around the world.

Figure 3: (**Global Data: A select few countries**) Impact of country wide social distancing policies on human mobility and COVID-19 dynamics. The time range is from 2020/02/02-2020/02/08 to 2020/05/24-2020/05/30. (a) Outgoing flow reduction from eight selected countries across the world: Australia, Brazil, India, Italy, Nigeria, Singapore, United Kingdom, and United States. To obtain these flow reduction rates (FRR), we used outgoing flows in January of each of these countries as their respective baseline and calculated the change with respect to the respective baseline from week 2020/02/02-2020/02/08 to 2020/05/24-2020/05/30. Each line in the plot corresponds to a country and it shows weekly FRR for that country with respect to the baseline. In addition, we use a vertical dashed line to show the week when lockdown or SAHO were imposed by each of these countries. The time of SAHO or lockdown mandates in different countries may overlap with each other. Note that FRR and lockdown lines share the same color for a given country. It is interesting to note that most countries show a decreasing trend in flow a week or two prior to the formal lockdown announcement. Moreover, most of these selected countries have initiated lockdown measures in the later part of March 2020. (b)Weekly outgoing flows (black) vs. new confirmed cases (red) in the selected eight countries. The blue dashed lines depict the start of the lockdown week in each country. We observe that the two lines show opposite trends. (c) Weekly FRR (black) vs. CGR of the new confirmed cases (red) in the eight selected countries. The blue dashed lines depict the start of the lockdown week in each country. We observe that post-lockdown CGR is decreasing for most of these countries.

present the results in Figure 4b and 4c. We observe that most county-level and state-level mobility show highest correlation with new confirmed cases lagged by -2. Figure 4d presents the weekly outgoing flow and new confirmed case count which are aggregated to state level in top 5 states with the largest total confirmed cases.

We further explore the change in MF using the flow reduction rate (FRR) (1). Figure 5a shows a 41% ($Q1 : -50\%$ and $Q3 : -30\%$ in FRR) mobility reduction compared to baseline flows across the states during the week 2020/03/22-2020/03/28 where most states had declared SAHO. There is a 17% reduction in flows during the week of the first school closure order, and 19% flow reduction in the week of the first SAHO. In the subsequent three weeks after the declaration of SAH orders the flows remain nearly constant, but since then there has been a rebound in the flows with flow reduction at around -19%. The relative timing of the reduction in flows indicates that the population complied to the social distancing guidelines and reduced mobility before the SAH orders.

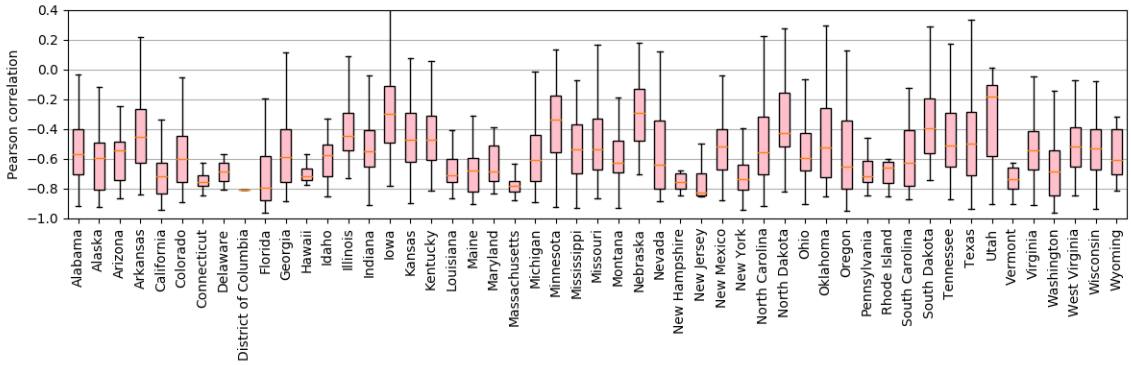
Figure 5b presents the timeline of FRR, the growth rate of the new confirmed cases (CGR) and social distancing orders in the five states with the highest number of confirmed cases. Although, we observe that the number of new confirmed cases increase despite the MF dropping rapidly during the week from 2020/03/01-2020/03/07 to week 2020/03/29-2020/04/04, we observe that the rate of growth of cases drops significantly.

Spatial distribution of mobility patterns and COVID-19 cases

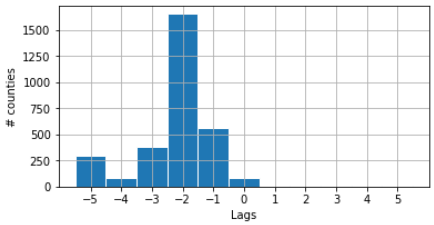
Figure 6 provides a spatio-temporal view of how the FRR coincides with the state-level new cases. During the week of 2020/01/26-2020/02/01, when the US recorded its first few cases, the mobility patterns across all the states show normal behaviour. As of 2020/03/02, 100 confirmed cases were recorded across the US; by March 2020/03/17, all the 50 states had the incidence of COVID-19 and during the week of 2020/03/15-2020/03/21, New York state had recorded nearly 10,000 new cases. We observe that with the progression of the pandemic, predominantly, the most populous states have the highest number of new cases. During the week of 2020/03/15-2020/03/21 a national emergency was declared and in addition several states had already closed schools and we begin to observe an overall reduction in the mobility. We also observe that generally the states with high number of cases also have higher reduction in flow. This becomes more evident in the week of 2020/03/29-2020/04/04 where we observe flows reducing with the case counts. In the interstate mobility matrix, the self-loop flows are suppressed. In order to achieve a sense of adjacency, we group the states according to their respective HHS region designations. During normal times, the inter-state mobility matrix shows considerable flows across all the states. During the week of 2020/03/15-2020/03/21 a few interstate connections start to drop and in the week 2020/03/29-2020/04/04 we observe nearly 70% of the state-pairs that are present during normal times, dropping.

An analysis of mixing within counties

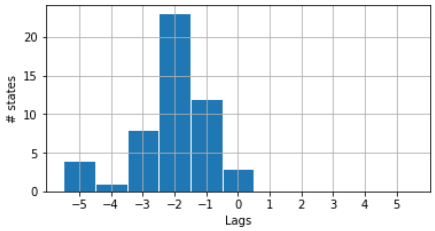
In this analysis, we attempt to capture the mixing within a county by employing the inter-cell flows. Ideally, with social distancing orders in place, one would expect inter-cell flows within a county to drop. Due to reduced inter-cell flows, normalized flows should tend towards an identity matrix and hence $SDI(t)$ defined in Equation (2) should move closer to 1. In Figure 7a, the boxplot represents the variation in $SDI(t)$ across the various weeks of 2020. We observe $SDI(t)$ to be nearly constant until the implementation of national emergency and state-level orders after which we start to observe an increase in SDI (10%). However, by the third week of April we start to observe SDI dropping and staying nearly constant over the weeks of May, which could be attributed to social distancing fatigue, a desire to return to daily routines, and other factors. Although, the SDI has reduced, the values are higher than the baseline values indicating a lower than normal mixing. As a general observation of variations in SDI across counties, we consider five states which have experienced the highest number of cases. The variations in $SDI(t)$ across different weeks of the pandemic can be observed through the choropleth plots in Figure 7b, 7c, and 7d. We observe the overall shading moving towards yellow indicating reduction in mixing within counties. Since the last week of April, we observe a decrease in SDI indicating *social distancing fatigue*. Importantly, the SDI has remained nearly constant over the month of May and median SDI nearly 5% higher than the median baseline values.



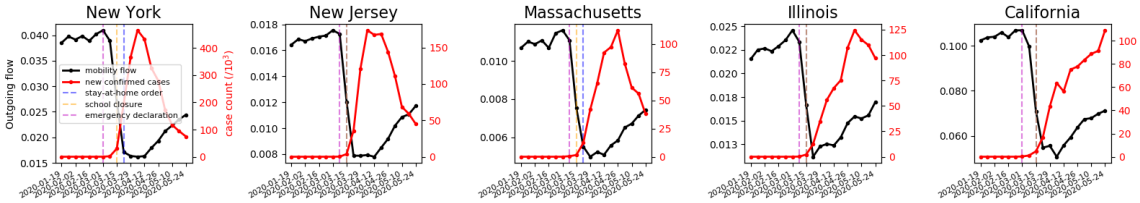
(a) Pearson correlation between county level outgoing flow and new confirmed case count by states.



(b) Frequency of lags that produce maximum absolute Pearson's correlation value between county level outgoing flow and lagged new confirmed case count by states.



(c) Frequency of lags that produce maximum absolute Pearson's correlation value between state level outgoing flow and lagged new confirmed case count by states.



(d) Weekly MF vs. new confirmed cases in top 5 states with the largest number of confirmed cases.

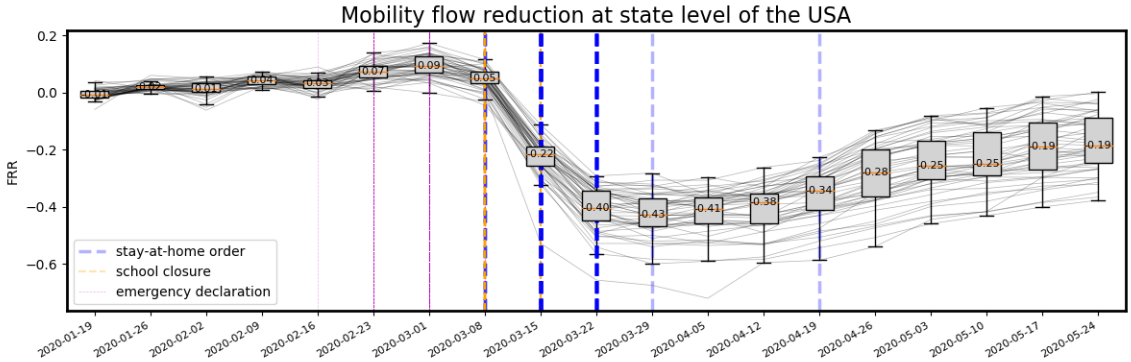
Figure 4: (**US Data**). Correlation analysis of human mobility and COVID-19 dynamics. (a) In order to analyze the human mobility and COVID-19 dynamics during the initial phase of the pandemic, we restrict our time range from 2020/01/19-2020/01/25 to 2020/05/03-2020/05/09 for the correlation analysis. This shows moderate to high negative correlation between MFand new confirmed case count. Analyses in (b) and (c) to identify the lags at which the correlation is maximum (absolute value) reveal similar patterns at both the county-level and state-level. (d) Weekly outgoing flow (black) and new confirmed cases (red) in the top 5 states with the largest number of confirmed cases. The two lines show opposite trends and cross around week 2020/03/22-2020/03/28. The mobility starts to drop before the SAHO (dashed blue vertical line) in 5 states.

2.3 Research findings for India

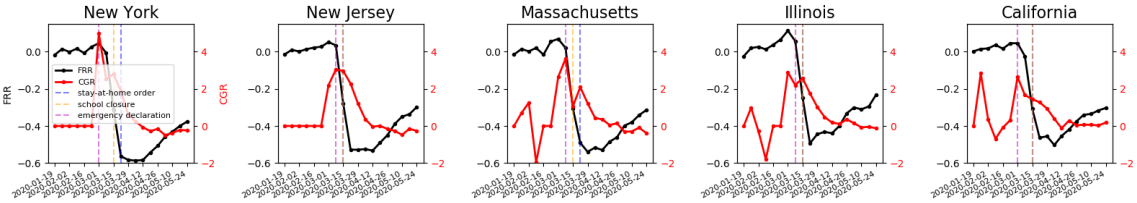
2.3.1 Analysis of mobility flows and its correlation with cases

The temporal range is from week 2020/03/01-2020/03/07 (the first confirmed case appeared in India) to week 2020/05/24-2020/05/30. The analysis is mainly conducted at the state level in India due to lack of authoritative infection data at district level.

Figure 8b shows Pearson correlations between weekly outgoing flow and new confirmed case count at the state level (again the date range restricted from 2020/03/01-2020/03/07 to 2020/05/03-2020/05/09 to



(a) Weekly MF reduction rate (FRR).

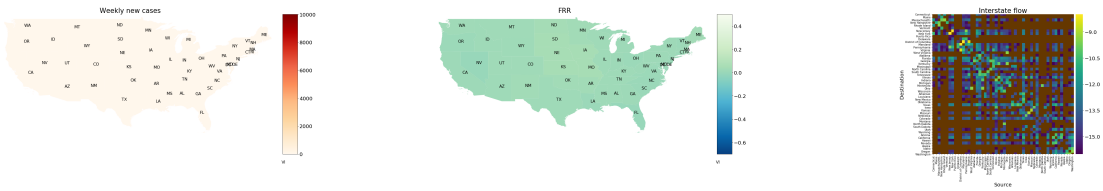


(b) Weekly FRR vs. CGR in top 5 states with the largest number of confirmed cases.

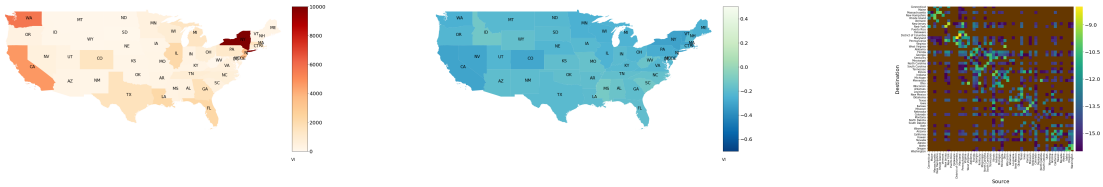
Figure 5: (**US Data**). Impact of state level social distancing policies on human mobility and COVID-19 dynamics. The time range is from 2020/01/19-2020/01/25 to 2020/05/24-2020/05/30. Vertical dashed lines mark the time of state level social distancing mandates including emergency declaration (purple), school closure (orange), and SAHO (blue). The time of mandates at different states may overlap with each other. (a) A spaghetti plot shows the weekly FRR time series of 50 states. A boxplot is used to display variation in samples of 50 states each week. The median value is shown along with the median line (orange, also outliers are disregarded). A positive value indicates an increase in flow and a negative value means reduction in flows when compared with pre-pandemic flows. (b) Weekly FRR (black) vs. CGR of the new confirmed cases (red) in 5 states with the highest number of confirmed cases. Although, in Figure 4d, we see that the number of new cases are increasing, (b) indicates that the growth rate of the new cases drops considerably across states.

capture the initial phase of the pandemic). We observe that the correlation varies across states and shows low (-0.75 in Uttarakhand) to high (-0.27 in Punjab) negative correlation between MF and new confirmed case count. We have ignored smaller states with less than 10 COVID-19 cases from this analysis.

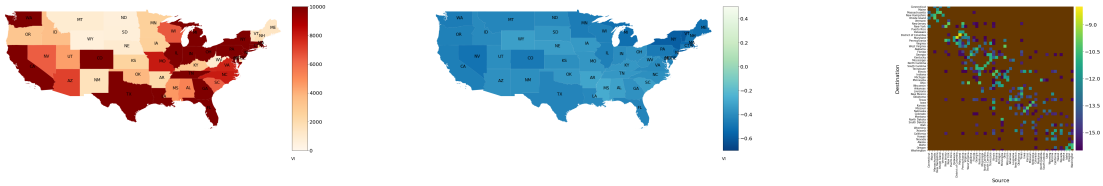
We further explore the MF change by applying flow reduction rate (FRR) (1). Figure 8a shows a 66% ($Q1 : -69\%$ and $Q3 : -57\%$ in FRR) mobility reduction compared to normal flows across the states during the week 2020/03/22-2020/03/28 where the central government had declared the first lockdown. Many states across India, large and small, like Maharashtra, Karnataka, Uttar Pradesh, Kerala, Uttarakhand, Odisha, Bihar, Chattisgarh, Punjab and Manipur declared school closure in the week 2020/03/08/-2020/03/14. Delhi announced school closure a week after these states. We observed 10% flow reduction in the next week after the school closure orders in the week 2020/03/08-2020/03/14. In the following week 2020/03/15-2020/03/21, many corporations issued work from home advisories and government offices decided to function with reduced staff strength with rotation. This seems to have resulted in a significant drop in the flow by 51%. The people's curfew on March 22, 2020 followed by nationwide lockdown for 21 days starting March 25, 2020 further reduced the flows by 2%. These flow levels were maintained for the next couple of weeks till the end of the first lockdown. The second lockdown in India was taken for 19 days between April 15, 2020 to May 3, 2020. We observe a slight increase in flows in the first week of the second lockdown and an increase of 10% in the third week. This increase can be attributed to allowance of certain economic activities in less affected areas post April 20. Over the month of May, the flows have steadily increased by nearly 40%



(a) 2020/01/26-2020/02/01: 4 cases recorded across the country



(b) 2020/03/15-2020/03/21: Week where multiple states implemented social distancing orders.



(c) 2020/03/29-2020/04/04: The week with the highest reduction in flows across the US.

Figure 6: (**US Data**) A comparison of variation in number of new cases to variation in flow reduction and inter-state mobility across different weeks. Column 1 represents the spatial distribution of new cases across the US, column 2 shows the spatial distribution of FRR, and column 3 depicts the heatmap of natural logarithm of the flow between states (states are ordered as per the health and human services grouping to obtain a sense of adjacency, self-loops are suppressed and brown indicates state-pairs not recorded in the data). Row 1 represents baseline times (only a few cases recorded). Row 2 represents the week when New York and California recorded considerable number of weekly new cases and we observe that there is considerable reduction in flows across the US and also interstate mobility. By the week of 2020/03/29-2020/04/04, the weekly number of new cases has increased substantially and almost all states have brought down mobility significantly, with the number of state pairs with non-zero flows dropping by 70%.

compared to flows during the imposition of lockdown. Thus the lockdown orders and graded resumption of economic activities seems to have large impact on mobility in India.

The effect of lockdown orders on mobility can be observed by comparing inter-state flows prior (Figure 9b) and post lockdown (Figure 9c). The usual flow of inter-state mobility has either reduced significantly or dried up completely (e. g. Uttar Pradesh-Madhya Pradesh).

2.3.2 Migration pattern analysis for India

Across the world, social distancing measures and the associated economic impact has changed human behavioral patterns. In India, news reports indicated that the lockdown triggered the movement of large groups of people from cities to their distant home towns. In order to detect and analyze the migratory patterns, we consider the time-varying mobility network of India with edges representing connection between district and edge weights as MF between them. An absence of an edge during a time period denotes relatively small amount of flow on that edge and thus is dropped for privacy reasons. Denoting $\mathcal{E}(t)$ as the set of edges at time t, we compute the Jaccard index $J(t)$[7] between the consecutive weeks of 2020. In addition, we also determine the number of edges added and removed across the weeks and show the respective plots in Figure 10.

[7]Given the set of edges $\mathcal{E}(t)$ and $\mathcal{E}(t-1)$, Jaccard index $J(t) = \frac{\mathcal{E}(t)\cap\mathcal{E}(t-1)}{\mathcal{E}(t)\cup\mathcal{E}(t-1)}$, where, $\cap$ and $\cup$ represent the standard intersection and union of sets operations, respectively.

We observe a sharp dip in the Jaccard index during the week of lockdown (March 25, 2020) indicating a large change in edge configuration compared to the previous week. This drop in Jaccard index can be attributed to the large number of edges being dropped during the lockdown week. Also, the number of new edges added from the week of lockdown is minimal. The interstate mobility patterns (cf. 9c) do not indicate large flows across states.

Analysis of the mobility data does not reveal mass migration of population at and during lockdown time and/or the subsequent weeks (see Section 3 for limitations). But we do observe a rise in mobility and the number of new connections being established between the districts in the recent weeks and one could attribute it to the various schemes introduced by the government to transport people to their destination.

2.4 Evaluating the epidemic impact of timely lockdown

Figure 11 and Figure 12 below show the posterior fitting to the smoothed daily new confirmed cases with normal and delay scenarios for multiple states in the US and India, where the blue line is the smoothed number of new confirmed cases per 100k. The red lines and pink shadows are the mean estimate and 95% CIs. The grey lines are fitted R_{eff} estimates. Solid lines denote normal schedule, dotted lines and dashed lines denote the scenarios of one-week and two-week delay. The numerical comparison between normal scenario and delay scenarios is presented in Table 1.

Figure 11 shows delay scenarios for the US and five states in the US. The searching range of R_{eff} is set to [0.5,3.2] for New York and New Jersey [18, 22] and [0.5,3] for the US, Massachusetts, Illinois, and California [22], the upper bound partially to counter the effects of testing ramp-up. We postpone R_{eff} at different weeks for different states according to their SAHO. For the US, we use the week 2020/03/15-2020/03/21 when the first state level SAHO is claimed. The simulations indicate that a delay in implementation of interventions by one week in the US would have led to 2.2M [95% CI: 1.5M - 2.8M] additional cases as of May 30, 2020. If the interventions had been postponed by two weeks, the number of additional cases would have been 6.2M [95% CI: 5.2M - 7.3M]. The ratios indicate nearly 2.2 times the ground truth number of cases for the one week delay scenario and nearly 4.4 times the ground truth number of cases for the two week delay scenario. (numbers are tabulated in Table 1).

In the context of India, the simulation results show similar trends as the US scenario. We set the searching range of R_{eff} to be [0.5,3.0] [23, 24] for the counterfactual simulation. The lockdown was implemented on the 2020/03/25 and hence, we use the R_{eff} of that week to create the one-week and two-week delay scenarios for India and all five states. The scenarios are demonstrated at both the country-level and the state-level and the results are shown in Fig. 12. As indicated by the case count curves, we see an increasing trend and the peak is yet to be appear. Overall, the factor of increase in total number of cases as observed around 2020/05/30 would be nearly 2 times and 4 times compared to actual confirmed cases for the one-week and two-week delay scenarios, respectively. Also, the simulations project that a one week delay in imposition of lockdown would have lead to 150K [95% CI: 120K - 174K] additional cases while a two-week delay would have resulted in 410K [95% CI: 350K - 480K] additional cases (numbers are tabulated in Table 1).

Although, the US has far more reported cases than India at the current time, it should be noted that the testing rate in US (55,000 per million) is significantly higher than that in India (2800 per million) [25]. Also, the two countries are in different phases of the epidemic with the number of new cases slowly declining in the US, while steadily increasing in India.

2.5 Discussion

Key findings:

1. Globally, we observe a considerable reduction in flows across most countries. With international airline traffic suspended by most countries, the inter-country flows have seen a significant drop. The US and India have had a 74% and 87% drop in international flows, respectively.

2. Overall, the state-level social distancing strategies across the US and the lock-down enforcement across India have paid dividends in reducing mobility with the US experiencing a maximum drop of 42% ($Q1 : -48\%$ and $Q3 : -40\%$) compared to usual flows across the states during the week where most

Table 1: Comparison in peak size, peak time, cumulative cases between normal scenario and delay scenarios. A ratio is given by v_{delay}/v_{normal} for peak size and cumulative cases, while a difference is given by $v_{delay} - v_{normal}$ for peak time and cumulative cases. Mean values with 95% CIs are shown in the table. The peak analysis for states where peak does not appear until May 30th is omitted.

	US					
one-week delay	US	NY	NJ	MA	IL	CA
peak size (ratio)	2.45	2.51	2.39	2.19	2.28	-
	[1.93,2.98]	[2.08,2.94]	[1.94,2.83]	[1.66,2.72]	[1.75,2.80]	-
peak time (diff. in days)	3	5	2	6	3	-
	[-20,25]	[-1,12]	[-11,16]	[-1,12]	[-12,18]	-
cumulative cases (ratio)	2.20	2.18	2.11	2.07	2.09	2.13
	[1.77,2.62]	[1.92,2.45]	[1.82,2.40]	[1.72,2.42]	[1.72,2.46]	[1.73,2.53]
diff. in cumulative cases	2161	403	178	104	131	127
(K)	[1541,2781]	[338,468]	[144,212]	[79,129]	[96,166]	[92,162]
two-week delay	US	NY	NJ	MA	IL	CA
peak size (ratio)	5.68	5.23	5.16	4.06	4.33	-
	[4.54,6.82]	[4.57,5.88]	[4.38,5.93]	[3.24,4.88]	[3.54,5.12]	-
peak time (diff. in days)	5	10	5	5	4	-
	[-11,22]	[4,16]	[-5,16]	[-5,15]	[-10,19]	-
cumulative cases (ratio)	4.44	3.92	3.75	3.69	3.83	4.37
	[3.59,5.28]	[3.53,4.32]	[3.32,4.17]	[3.12,4.26]	[3.29,4.36]	[3.54,5.20]
diff. in cumulative cases	6227	997	443	264	341	381
(K)	[5185,7269]	[921,1073]	[408,478]	[234,294]	[307,375]	[315,447]
	India					
one-week delay	IN	DL	GJ	MH	RJ	TN
cumulative cases (ratio)	2.08	1.58	2.74	2.37	2.48	2.02
	[1.45,2.70]	[1.18,1.98]	[1.65,3.83]	[1.57,3.17]	[1.64,3.33]	[1.42,2.63]
diff. in cumulative cases	151	9.8	17.0	50.6	8.1	18.1
(k)	[124.5,177.1]	[8.2,11.4]	[13.9,19.9]	[43.5,57.6]	[6556,9699]	[15.1,21.1]
two-week delay	IN	DL	GJ	MH	RJ	TN
cumulative cases (ratio)	3.92	2.83	7.43	4.72	5.30	4.46
	[2.71,5.15]	[2.14,3.53]	[4.58,4.85]	[4.53734 10.316]	[3.50,7.11]	[3.12,5.81]
diff. in cumulative cases	414.6	25.6	48.3	123.9	22.0	47.5
(K)	[346.6,482.6]	[22.2,29.0]	[40.6,55.9]	[106.3,141.6]	[18.1,25.9]	[41.2 ,53.9]

Table 2: Summary of key observations based on our analysis in the two countries

Observation	**US**	**India**
Maximum reduction in overall mobility	42% (Q1:-48%, Q3:-40%)	68% (Q1:-71%, Q3:-59%)
Reduction in mobility before SAH orders	22% (Q1:-30%, Q2:-10%)	15% (Q1:-26%, Q2:-5%)
Reduction in international outflows	74%	87%
Reduction in inter-state flow pairs	70%	70%
Change in intra-county mixing or SDI	10%	2%
Rebound in mobility since reduced mobility	53.5%	38.2%

states had declared SAHO. In India, on the other hand, there was a maximum drop of 68% ($Q1 : -72\%$ and $Q3 : -66\%$) across states compared to the baseline due the lock-down declaration.

3. US Specific findings:

   (a) The relative timing of the reduction in flows indicates that substantial amount of the population adhered to social distancing guidelines and reduced mobility before the state's imposition of the SAHO. There is a 17% reduction in flows during the week of the first school closure order, and 19% flow reduction in the week of the first SAHO. In the subsequent three weeks after the declaration of SAH orders the flows remain nearly a constant, but since then there has been a rebound in the flows with flow reduction at 19% as per the latest data.

   (b) At the county level, the flow pattern over the weeks of the progression of the pandemic show high (median -0.82 in New Jersey) negative correlation to low (median -0.2 in Utah) negative correlation with the number of new weekly confirmed cases. Importantly, we observe a steady reduction in growth rate with the reduction in mobility.

   (c) We observe a considerable reduction in interstate flows with nearly 70% drop in the intensity of travel between them when compared to pre-pandemic baseline.

(d) SDI (defined in Section 1.1) is a metric to quantify the level of mixing with a region. A higher value of SDI indicates lower level of mixing. We observe a 10% increase in SDI values post SAHO when compared to SDI computed during the pre-pandemic period.

(e) The counterfactual simulations project that the US would have seen up to 2.2M [95% CI: 1.5M - 2.8M] additional cases nationwide as of May 30, 2020. If the social distancing measures were postponed by one week, the number of cases would be 2.20 times the current number of confirmed cases. If the interventions postponed by two weeks, the number of additional cases would be 6.2M (95% CI: 5.2M - 7.3M) in the US and the number of cases would have been 4.44 times the current number of confirmed cases.

4. India specific findings:

(a) The mobility reduced by 15% in India a week prior to the first of the multiple series of lockdowns (this week saw a few states closing down schools). The MF dropped to 62% during the week of formal lockdown announcement following the "People's curfew". The reduction in mobility continued until the end of the first lockdown, which shows by and large compliance with policy intervention for social distancing. A steady decline in FRR over the weeks of May indicate rebound in the flow and can be attributed to either fatigue or graded relaxation is less affected zones in the country.

(b) Pearson's correlation coefficient between the MF and the number of confirmed cases is mostly negative with median of -0.43. The moderate correlation indicates that concurrent with the rising number of newly found cases there was a drop in MF.

(c) The interstate mobility among the Indian states reduced significantly. The flows have stopped completely or have registered a very large drop. This implies that the SDI for Indian states approached the values close to one very quickly after the lockdown was put in place indicating reduced mixing within districts.

(d) The counterfactual simulations project that a week's delay in lockdown implementation would have lead to 151K [95% CI: 124.5K-177.1K] additional cases while a two week delay would have resulted in 414.6K [95% CI: 346.6K-482.6K] additional cases as of May 30, 2020. These numbers imply that the total number of cases would have been nearly 2 times and 4 times the actual confirmed cases for one-week and two-week delay in interventions, respectively.

3 Limitations

An important aspect that is not captured in our analysis is COVID-19 testing. Some of the increase in cases is likely due to the increased testing in US and India. Thus, confirmed cases as reported are likely an under count of the prevalence; but the difference is hard to quantify exactly. Testing in India as well as the US was largely confined to individuals reporting to clinics with symptoms. Nevertheless, we believe our results show significant correlation patterns between human mobility and COVID-19 dynamics. It also provides a clear evidence of the abrupt changes in mobility patterns across the world in relatively small time window. The mobility map, like other such data sets, have a bias based on device ownership, when the devices are used, etc. For instance, Google mobility data is limited to smartphone users who have opted in to Google's consumer Location History feature, which is off by default. These data may not be representative of the population as a whole, and furthermore their representativeness may vary by location. Importantly, these limited data are only viewed through the lens of differential privacy algorithms, specifically designed to protect user anonymity and obscure fine detail. Moreover, comparisons across rather than within locations are only descriptive since these regions can differ in substantial ways.

We did not observe large MF between states due to migratory workers at the start of the lock down period (as well as during lock down period) in India. This can be potentially due to relative flow volumes needed to maintain differential privacy and ownership bias. Finally, a decrease in aggregate mobility does not provide precise estimates in the decrease in the social interaction. Estimating social interactions precisely will require different kinds of digital traces.

Research-in-context

Evidence before the study

We have searched Google scholar, Nature, Science, Lancet, medRxiv, for journal articles and news reports that have studied the mobility patterns during the COVID-19 pandemic. We used keywords such as "mobility reduction COVID-19", "mobility patterns COVID-19" "effect of social distancing on mobility" to filter out articles relevant to our study. Further, we used references within to obtain other related studies on the topic of mobility changes during COVID-19 crisis.

We found a few articles that have assessed the human mobility during the pandemic. Out of these articles, only a handful address the interplay between mobility, public policies, and disease dynamics.

Added value to study

As per the literature review, this work is the first to model the complex interplay between COVID-19 spread, human mobility, and government policies in near real-time at multiple spatial resolutions, such as, global, country, state, and sub-state level. In addition, this is the first study to employ the Google COVID-19 Aggregated Mobility Research Dataset (described in the main text) for this study. Although, the manuscript focuses on two countries, the analysis can be extended to any specific country. Counterfactual analysis to study the effects of delayed interventions is one-of-a-kind.

Interpretations

The study reveals a huge drop in mobility globally. The effect of this drop in mobility shows a concurrent drop in growth rate of COVID-19 cases. Counterfactual analysis reveal that the government interventions, at least in the US and India, were timely and averted a significant number of cases.

3.1 Data sharing statement

The Google COVID-19 Aggregated Mobility Research Dataset used for this study is available upon request to the corresponding authors and with permission of Google, LLC. University of Virginia COVID-19 Dashboard that provides detailed, global epidemic surveillance data and used in this study is available for public use on `https://nssac.bii.virginia.edu/covid-19/dashboard/`. The data on social distancing guidelines in US states and India COVID-19 surveillance data would be shared upon request and would be uploaded on `https://dataverse.lib.virginia.edu/`

Acknowledgements

We thank Aaron Schneider, Aaron Stein, Ahmed Aktay, Alvin Raj, Amy Chung-Yu Chou, Andrew Oplinger, Ashley Zlatinov , Blaise Aguera y Arcas, Bryant Gipson, Charina Chou, Christopher Pluntke, Damien Desfontaines, Eric Tholome, Ewa Dominowska, Gregor Rothfuss, Iz Conroy, Janel Thamkul, Janet Whiteman, Jason Freidenfelds, Jeff Dean, Karen Lee Smith, Katherine Chou, Leeron Morad, Lizzie Dorfman, Marlo McGriff , Mia Vu, Michael Howell, Paul Eastham, Rif Saurous, Rishi Bal, Royce Wilson, Ruth Alcantara, Shawn O'Banion, Stephanie Cason, Thomas Roessler, Vivien Hoang, Yanning Zhang, Manish Gupta, Pankaj Gupta, Ashwani Sharma and Divy Thakkar for their support and guidance. We also thank members of the Biocomplexity Institute and Initiative, University of Virginia for useful discussion and suggestions. This work was partially supported by National Institutes of Health (NIH) Grant 1R01GM109718, NSF BIG DATA Grant IIS-1633028, NSF DIBBS Grant ACI-1443054, NSF Grant No.: OAC-1916805, NSF Expeditions in Computing Grant CCF-1918656, CCF-1917819, US Centers for Disease Control and Prevention 75D30119C05935, DTRA subcontract/ARA S-D00189-15-TO-01-UVA, and a collaborative seed grant from the UVA Global Infectious Disease Institute.

References

[1] Benjamin N. Rome and Jerry Avorn. Drug evaluation during the Covid-19 pandemic. *New England Journal of Medicine*, Apr 14, 2020.

[2] Gregory A Wellenius, Swapnil Vispute, Valeria Espinosa, Alex Fabrikant, Thomas C Tsai, Jonathan Hennessy, Brian Williams, Krishna Gadepalli, Adam Boulange, Adam Pearce, et al. Impacts of state-level policies on social distancing in the United States using aggregated mobility data during the COVID-19 pandemic. *arXiv preprint arXiv:2004.10172*, 2020.

[3] Caroline O Buckee, Satchit Balsari, Jennifer Chan, Mercè Crosas, Francesca Dominici, Urs Gasser, Yonatan H Grad, Bryan Grenfell, M Elizabeth Halloran, Moritz UG Kraemer, et al. Aggregated mobility data could help fight COVID-19. *Science (New York, NY)*, 368(6487):145, 2020.

[4] Nuria Oliver, Emmanuel Letouzé, Harald Sterly, Sébastien Delataille, Marco De Nadai, Bruno Lepri, Renaud Lambiotte, Richard Benjamins, Ciro Cattuto, Vittoria Colizza, et al. Mobile phone data and COVID-19: Missing an opportunity? *arXiv preprint arXiv:2003.12347*, 2020.

[5] Isaac I Bogoch, Alexander Watts, Andrea Thomas-Bachli, Carmen Huber, Moritz UG Kraemer, and Kamran Khan. Potential for global spread of a novel coronavirus from China. *Journal of Travel Medicine*, 27(2):taaa011, 2020.

[6] Matteo Chinazzi, Jessica T Davis, Marco Ajelli, Corrado Gioannini, Maria Litvinova, Stefano Merler, Ana Pastore y Piontti, Kunpeng Mu, Luca Rossi, Kaiyuan Sun, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, 368(6489):395–400, 2020.

[7] Aniruddha Adiga, Srinivasan Venkatramanan, James Schlitt, Akhil Peddireddy, Allan Dickerman, Andrei Bura, Andrew Warren, Brian D Klahn, Chunhong Mao, Dawen Xie, Dustin Machi, Erin Raymond, Fanchao Meng, Golda Barrow, Henning Mortveit, Jiangzhuo Chen, Jim Walke, Joshua Goldstein, Mandy L Wilson, Mark Orr, Przemyslaw Porebski, Pyrros A Telionis, Richard Beckman, Stefan Hoops, Stephen Eubank, Young Yun Baek, Bryan Lewis, Madhav Marathe, and Chris Barrett. Evaluating the impact of international airline suspensions on the early global spread of COVID-19. *medRxiv*, 2020.

[8] Moritz U. G. Kraemer, Chia-Hung Yang, Bernardo Gutierrez, Chieh-Hsi Wu, Brennan Klein, David M. Pigott, , Louis du Plessis, Nuno R. Faria, Ruoran Li, William P. Hanage, John S. Brownstein, Maylis Layan, Alessandro Vespignani, Huaiyu Tian, Christopher Dye, Oliver G. Pybus, and Samuel V. Scarpino. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, 368(6490):493–497, 2020.

[9] Dianbo Liu, Leonardo Clemente, Canelle Poirier, Xiyu Ding, Matteo Chinazzi, Jessica T Davis, Alessandro Vespignani, and Mauricio Santillana. A machine learning methodology for real-time forecasting of the 2019-2020 COVID-19 outbreak using internet searches, news alerts, and estimates from mechanistic models, 2020.

[10] Shengjie Lai, Nick W Ruktanonchai, Liangcai Zhou, Olivia Prosper, Wei Luo, Jessica R Floyd, Amy Wesolowski, Chi Zhang, Xiangjun Du, Hongjie Yu, et al. Effect of non-pharmaceutical interventions for containing the COVID-19 outbreak: An observational and modelling study. *medRxiv*, 2020.

[11] UMD. *COVID-19 impact research findings*, 2020.

[12] Arielle Lasry, Daniel Kidder, Marisa Hast, and et al. Timing of community mitigation and changes in reported COVID-19 and community mobility — four U.S. metropolitan areas. *MMWR Morb Mortal Wkly Rep 2020*, 69:451–457, 2020.

[13] Apple. *Mobility Trends Reports*, 2020 (accessed May 14, 2020).

[14] Brennan Klein et al. *Reshaping a nation: Mobility, commuting, and contact patterns during the COVID-19 outbreak*, 2020-05-11 (version 3).

[15] Song Gao, Jinmeng Rao, Yuhao Kang, Yunlei Liang, Jake Kruse, Doerte Doepfer, Ajay K Sethi, Juan Francisco Mandujano Reyes, Jonathan Patz, and Brian S Yandell. Mobile phone location data reveal the effect and geographic variation of social distancing on the spread of the COVID-19 epidemic. *arXiv preprint arXiv:2004.11430*, 2020.

[16] UVA. *UVA COVID-19 Surveillance Dashboard*, 2020 (accessed May 14, 2020).

[17] Gaston De Serres, Nigel J Gay, and C Paddy Farrington. Epidemiology of transmissible diseases after elimination. *American Journal of Epidemiology*, 151(11):1039–1048, 2000.

[18] Sen Pei, Sasikiran Kandula, and Jeffrey Shaman. Differential effects of intervention timing on COVID-19 spread in the United States. *medRxiv*, 2020.

[19] Herbert W Hethcote. The mathematics of infectious diseases. *SIAM review*, 42(4):599–653, 2000.

[20] Srinivasan Venkatramanan, Jiangzhuo Chen, Arindam Fadikar, Sandeep Gupta, Dave Higdon, Bryan Lewis, Madhav Marathe, Henning Mortveit, and Anil Vullikanti. Optimizing spatial allocation of seasonal influenza vaccine under temporal constraints. *PLoS computational biology*, 15(9):e1007111, 2019.

[21] 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.

[22] Kevin Systrom. The metric we need to manage COVID-19 Rt: The effective reproduction number. `http://systrom.com/blog/the-metric-we-need-to-manage-covid-19/`, April 12, 2020.

[23] Sourish Das. Prediction of COVID-19 disease progression in India: Under the effect of national lockdown. *arXiv preprint arXiv:2004.03147*, 2020.

[24] Manisha Mandal and Shyamapada Mandal. COVID-19 pandemic scenario in India compared to China and rest of the world: A data driven and model analysis. *medRxiv*, 2020.

[25] Worldometer. `https://www.worldometers.info/coronavirus/`, 2020.

[26] Aleix Bassolas, Hugo Barbosa-Filho, Brian Dickinson, Xerxes Dotiwalla, Paul Eastham, Riccardo Gallotti, Gourab Ghoshal, Bryant Gipson, Surendra A Hazarie, Henry Kautz, et al. Hierarchical organization of urban mobility and its connection with city livability. *Nature communications*, 10(1):1–10, 2019.

[27] Royce J Wilson, Celia Yuxin Zhang, William Lam, Damien Desfontaines, Daniel Simmons-Marengo, and Bryant Gipson. Differentially private SQL with bounded user contribution. *arXiv preprint arXiv:1909.01917*, 2019.

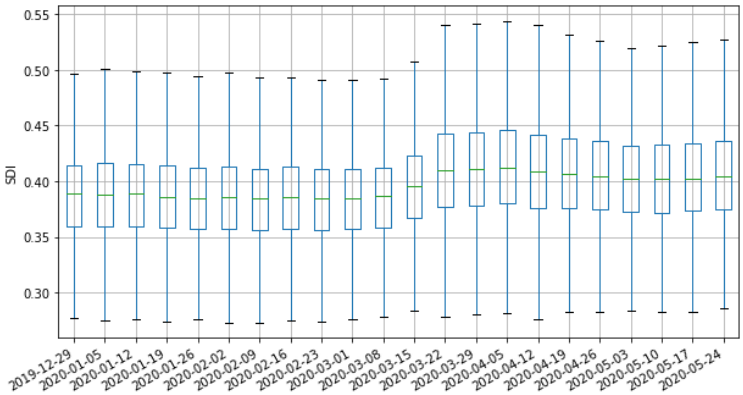
Appendices

A Google COVID-19 Aggregated Mobility Research Dataset

The dataset contains anonymized mobility flows aggregated over users who have turned on the Location History setting, which is off by default. This is similar to the data used to show how busy certain types of places are in Google Maps — helping identify when a local business tends to be the most crowded. The dataset aggregates flows of people from region to region, which is here further aggregated at multiple geographical resolutions weekly.

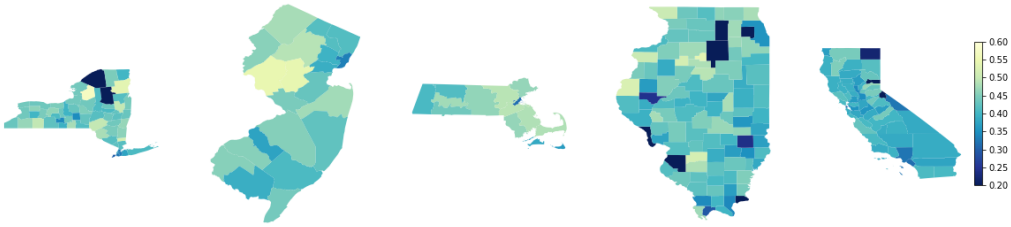
To produce this dataset, machine learning is applied to logs data to automatically segment it into semantic trips [26]. To provide strong privacy guarantees, all trips were anonymized and aggregated using a differentially private mechanism [27] to aggregate flows over time. This research is done on the resulting heavily aggregated and differential private data. No individual user data was ever manually inspected, only heavily aggregated flows of large populations were handled.

All anonymized trips are processed in aggregate to extract their origin and destination location and time. For example, if users traveled from location a to location b within time interval tIn assessing public levels of compliance to social distancing, several analyses have revealed reduction in overall mobility.

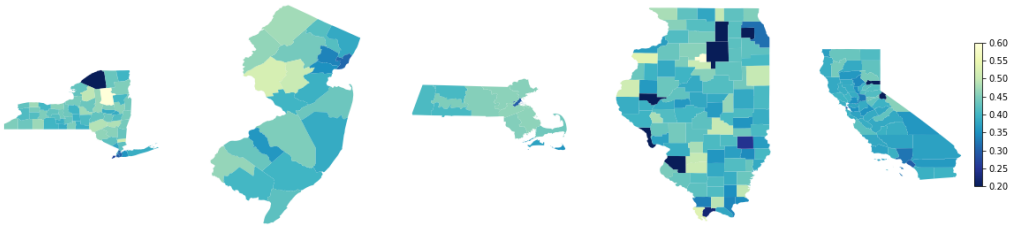


(a)

(b) SDI in the counties of New York, New Jersey, Massachusetts, Illinois, California pre-pandemic.

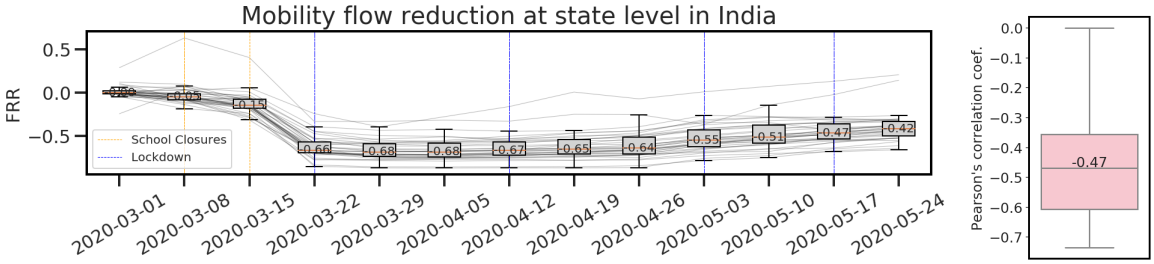


(c) SDI in the counties of New York, New Jersey, Massachusetts, Illinois, California in the week of 2020/03/29-2020/04/04



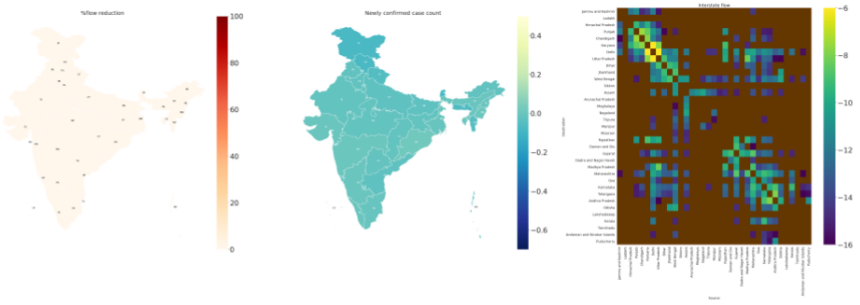
(d) SDI in the counties of New York, New Jersey, Massachusetts, Illinois, California in the week of 2020/05/24-2020/05/30

Figure 7: SDI at a county level. (a) Variation in SDI across all the counties of the US (obtained from the cell-to-cell flows within a county) for various weeks in the 2020. We observe that from week of 2020/03/15 (when most orders where implemented) the overall SDI starts to increase, implying that the mixing reduced. But since the last week of April we observe the overall SDI dropping and staying nearly constant over the weeks of May. (b), (c), and (d) shows the choropleth plots with counties shaded according to their SDI values for five states with highest number of confirmed cases pre- and post-state-orders. (d) Indicates the date when states have relaxed the social distancing order. Although, the mixing has increased, the SDI values are higher than the baseline values.

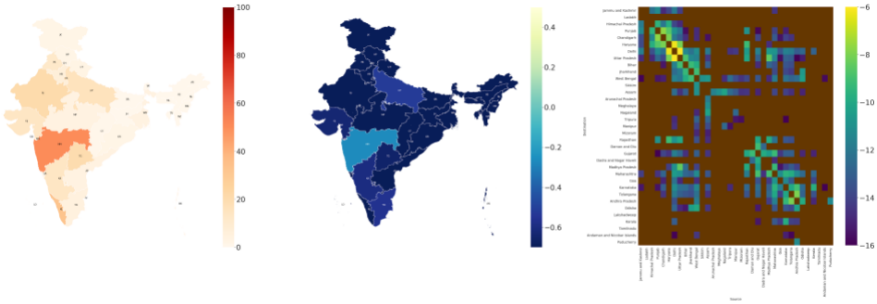


(a) (b)

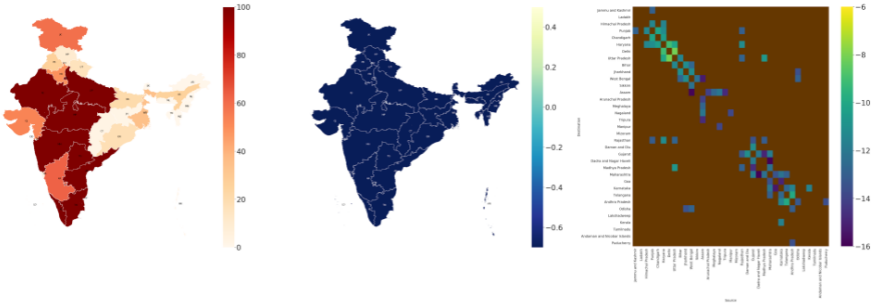
Figure 8: (**India Data**) Impact of state level social distancing policies on human mobility and COVID-19 dynamics. The time range is from 2020/03/01-2020/03/07 to 2020/05/24-2020/05/30. Vertical dashed lines mark the time of state level social distancing mandates including school closure (orange), and lockdown order (blue). The time of different mandates at different states may overlap with each other. (a) A spaghetti plot showing the weekly FRR time series for the 28 states and 8 union territories. A boxplot is used to display variation in samples of each of these entities every week. The median value is shown along with the median line (orange). A positive value means flow increase compared with pre-pandemic flows, while a negative value means a reduction in flows. We observed significant reduction in MF after the on-set of COVID-19 pandemic in India. The lockdown measures further reduced the flow. Towards the end of second lockdown (May 3, 2020), we observe slight increase in the MF and since then there has been a steady increase in flows. (b) Distribution of Pearson's correlation coefficient between state/union territory level outgoing flow and new confirmed case count at weekly levels during the initial phase of the pandemic. Note that we have only considered states with total cases more than 10 for this analysis. We observed Pearson's correlation coefficient of -0.47 for India, which indicates moderate negative correlation between the outgoing MF and newly confirmed cases per week.



(a) 2020/01/26-2020/02/01: 1 case recorded across the country



(b) 2020/03/15-2020/03/21: Week where cases detected in multiple states in India and school closures were ordered.



(c) 2020/03/29-2020/04/04: The week with the highest reduction in flows across India and this is the first full week of nation-wide lockdown.

Figure 9: (**India Data**) A comparison of variation in number of new cases to variation in flow reduction and inter-state mobility across different weeks. Column 1 represents the spatial distribution of new cases across India, column2 shows the spatial distribution of FRR, and column 3 depicts the heatmap of natural logarithm of the flow between states (states are ordered in order to preserve a sense of adjacency as far as possible with self-loops suppressed and brown shade indicates state-pairs not recorded in the data.). Row 1 represents normal times (only a few cases recorded). Row 2 represents the week when Maharashtra and Delhi started recording number of weekly new cases and we observe that there is considerable reduction in flows across India and also interstate mobility. By the week of 2020/03/29-2020/04/04, the weekly new cases has increased substantially and the nation-wide lockdown has brought down mobility substantially, with the numbers of state pairs with non-zero flows dropping by nearly 70% like in the US.

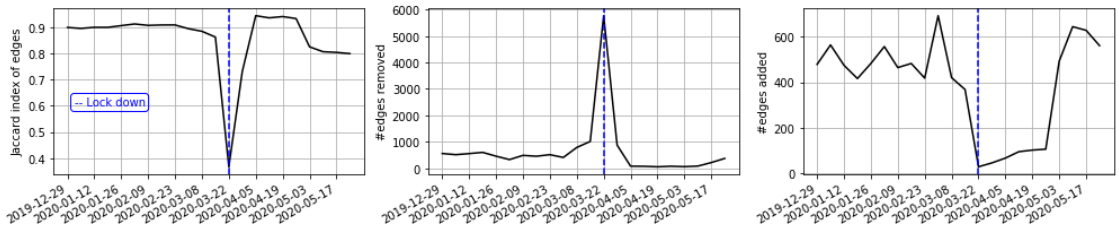


Figure 10: Migration analysis: Change in Jaccard index, number of edges (or connections between districts) removed, and number of new edges added across multiple weeks of 2020. Across the plots, the pattern at the lockdown announcement week indicates a significant drop in connections and flow between districts.

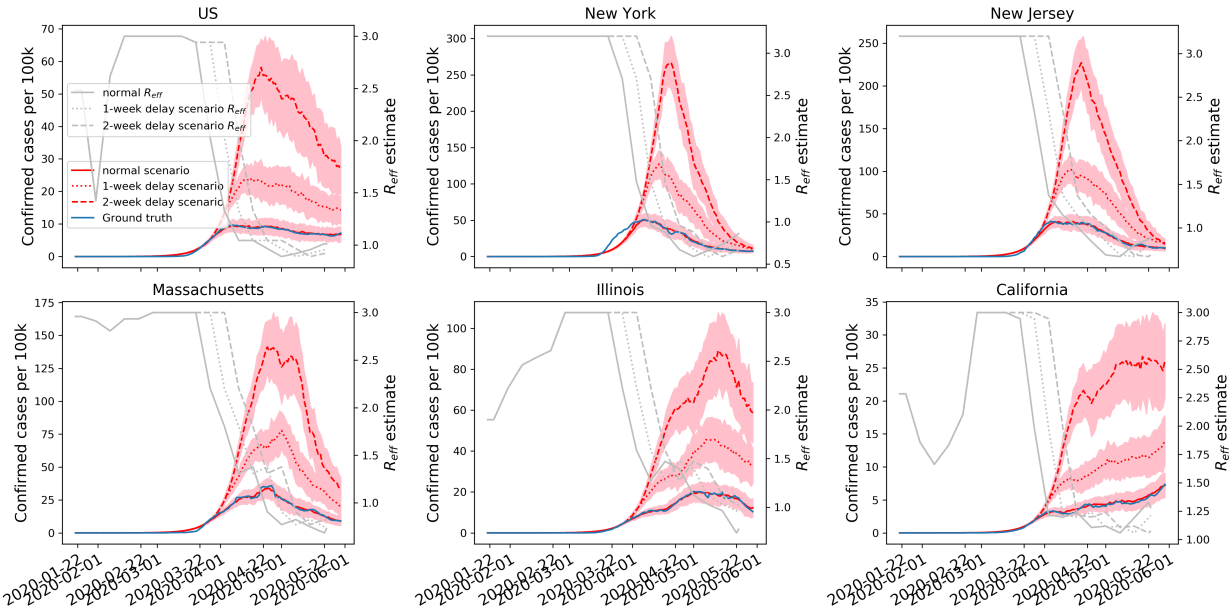


Figure 11: Posterior fitting to daily new confirmed cases with delayed intervention scenarios in the US and top five states with the highest number of confirmed cases. The blue line is the smoothed number of new confirmed cases per 100k. The red lines and pink shadow are the mean estimate and 95% CIs. The grey lines are fitted R_{eff} estimates. Solid lines denote normal scenario, dotted lines and dashed lines denote the scenario of one-week and two-week delay.

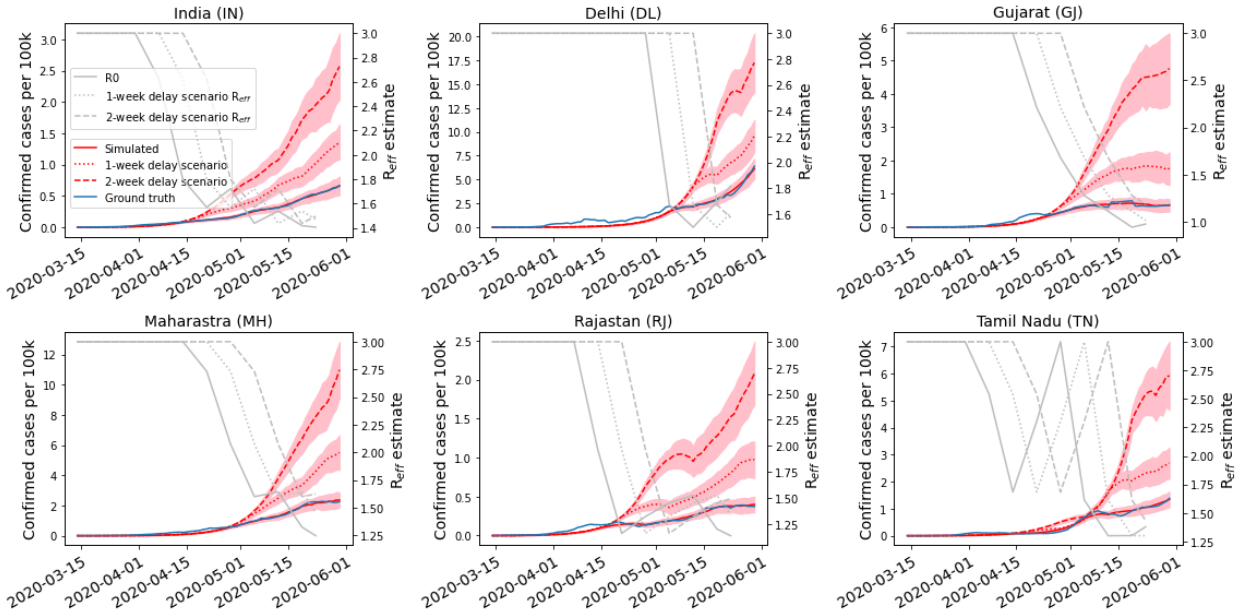


Figure 12: Delayed intervention scenarios for India and a few states with highest number of confirmed cases and India. The blue line indicates the smoothed number of daily cases, red lines correspond to simulation results with scenarios appropriately indicated, and the grey lines show the estimated R_{eff} and scenarios created by prolonging the R_{eff} for one week and two weeks. In general, we observe that with a one-week delay in implementation of lockdown, the total number of cases nearly doubles, while for the two-week scenario, we observe the cases increasing four folds in India. Mostly, the number of new cases is steadily increasing across India.