



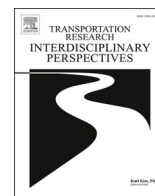
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Transportation Research Interdisciplinary Perspectives

journal homepage: www.sciencedirect.com/journal/transportation-research-interdisciplinary-perspectives



Structural modeling of COVID-19 spread in relation to human mobility

Rezwana Rafiq^{a,*}, Tanjeeb Ahmed^a, Md Yusuf Sarwar Uddin^b

^a Institute of Transportation Studies, University of California, Irvine, CA 92697-3600, USA

^b Department of Computer Science and Electrical Engineering, University of Missouri-Kansas City, MO 64110, USA

ARTICLE INFO

Keywords:

COVID-19 pandemic
Human mobility
Infection rate
SEM
Latent factor
COVID-19 policies
Spatial effect
Big data

ABSTRACT

Human mobility is considered as one of the prominent non-pharmaceutical interventions to control the spread of the pandemic (positive effect from mobility to infection). Conversely, the spread of the pandemic triggered massive changes to people's daily schedules by limiting their movement (negative effect from infection to mobility). The purpose of this study is to investigate this bi-directional relationship between human mobility and COVID-19 spread across U.S. counties during the early phase of the pandemic when infection rates were stabilizing and activity-travel behavior reflected a fairly steady return to normal following the drastic changes observed during the pandemic's initial shock. In particular, we applied Structural Regression (SR) model to investigate a bi-directional relationship between COVID-19 infection rate and the degree of human mobility in a county in association with socio-demographic and location characteristics of that county, and state-wide COVID-19 policies. Combining U.S. county-level cross-sectional data from multiple sources, our model results suggested that during the study period, human mobility and infection rate in a county both influenced each other, but in an opposite direction. Metropolitan counties experienced higher infection and lower mobility than non-metropolitan counties in the early stage of the pandemic. Counties with highly infected neighboring counties and more external trips had a higher infection rate. During the study period, community mitigation strategies, such as stay at home order, emergency declaration, and non-essential business closure significantly reduced mobility whereas public mask mandate significantly reduced infection rates. The findings of this study will provide important insights to policy makers in understanding the two-way relationship between human mobility and COVID-19 spread and to derive mobility-driven policy actions accordingly.

Introduction

The ongoing COVID-19 pandemic has caused immense disruptions to our day-to-day normal life. Admittedly, human mobility plays an important role in spreading the virus (Bhouri et al., 2021; Bryant and Elofsson, 2020; Chang et al., 2021). For this reason, to control the spread of the virus various human mobility restrictions, such as stay at home order, limited public gathering, non-essential business closures had been extensively exercised as a crucial non-pharmaceutical intervention tool all over the world. Despite the wide adoption of these mobility restrictions, the extent of the relationship between human mobility and the spread of the disease in a *quantitative* manner is still unclear. The purpose of this study is to understand this underlying relationship across U.S. counties during the early phase of the pandemic.

There had been several studies that investigated the relationships between travel pattern and COVID-19 spread. While analyzing the virus spread across the cities in China, Kraemer et al. (2020) and Chinazzi

et al. (2020) both concluded that movement restrictions, supplemented with public health policies were most effective in reducing the spread of the disease. Other studies outside China found a significant positive relationship between COVID-19 cases and mobility. Carteni et al. (2020) found mobility as the best predictor of daily COVID-19 cases when conducting a multiple regression analysis using sample data collected in Italy between January to March 2020. In another study in Italy, mobility restriction was estimated to achieve a reduction of 45 percent infections and 200,000 hospitalizations during the same time frame (Gatto et al., 2020). On the other hand, some recent studies explored how infection rate affected mobility and obtained significant negative connections from infection to mobility (Hu et al., 2021; Kim and Kwan, 2021).

The purpose of this study was to address the following three research gaps. First, arguably, the relationship between human mobility and COVID-19 spread can hold in both directions: mobility can affect the spread and the spread can also influence the mobility in reverse. To the best of our knowledge, no prior studies captured this bi-directional

* Corresponding author.

E-mail addresses: rfaq@uci.edu (R. Rafiq), ahmed.t@uci.edu (T. Ahmed), muddin@umkc.edu (M. Yusuf Sarwar Uddin).

<https://doi.org/10.1016/j.trip.2021.100528>

Received 19 June 2021; Received in revised form 13 November 2021; Accepted 24 December 2021

Available online 29 December 2021

2590-1982/© 2021 The Author(s).

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

relationship. In this study, we investigated this relationship in a quantitative manner within a single framework to clearly understand the complex relationships between mobility and disease spread. Second, the term “mobility” can be characterized by a wide range of factors. Therefore, this term needs to be defined appropriately to capture its combined effect on disease transmission in a geographical area. We assigned a factor scoring to mobility based on three fundamental mobility matrices: fraction of people traveled, frequency of travel, and travel distance. Finally, besides human mobility, the spread of the disease can be directly or indirectly (through mobility) influenced by other factors, such as socio-demographic and location characteristics, community mitigation policies (e.g., mask and quarantine mandate), and human movements from adjacent areas. To understand and explain the relationship between mobility and disease spread, all these direct and indirect effects need to be considered and analyzed in a comprehensive methodological framework. To do so, we developed a Structural Regression (SR) model across U.S. counties by conceptualizing a bi-directional relationship between human mobility and COVID-19 infection along with other influencing factors that could potentially affect both mobility and infection during the early phase of the pandemic. The findings of this study can help policy makers to understand the relationships between human mobility and infection rate on a quantitative scale and thus, can help them to derive mobility-driven policy measures in controlling the spread of the pandemic.

The next sections describe relevant studies regarding the relationships between mobility and infection. After describing the data sources, study timeframe, and data preparation, we provide an overview of the methodology. The methodology section includes the conceptual framework, exogenous and endogenous model variables, and the model estimation technique. Model results are then presented and discussed followed by model validation and scenario testing. Finally, major findings, policy implications, and limitations are presented in the last section.

The severity of COVID-19 and human mobility patterns

This section provides an overview of previous studies relevant to this work with a particular focus on the impact of COVID on human mobility patterns.

Progression and impacts of COVID-19

Coronavirus Disease 2019 or COVID-19 can be regarded as the consequence of the third generation of coronaviruses originated from Wuhan, China in December 2019, and quickly spread to 72 countries within a span of only three months (Li et al., 2020a). While in the beginning, the new cases were traced to the persons who traveled to the infected areas, later, local transmission fueled the exponential rise in cases.

Within the first 11 months of its arrival, COVID-19 has imparted an unprecedented effect on health (physical and mental), lifestyle, and the economy. According to data published by Johns Hopkins University & Medicine (2020), global COVID-19 cases on December 4, 2020, reached over 65 million, of which, 14 million are attributed to the United States. During the same period, COVID-19 claimed over 1.5 million lives globally and around 276 thousand in the United States. The United States has already seen at least two peaks in infection, hospitalization, and death rates. This triggered the enactment of several region-based policies that attempted to limit human movement and interactions. Along with the severe health and mental condition of the general population, COVID-19 has also impacted the normal operation of business, which in turn resulted in a significant rise in the unemployment rates – the starting point of a negative snowball effect on the economy. Yoo and Managi (2020) estimate a 34.6% GDP loss in the United States due to COVID-19, while International Monetary Fund (2021) forecasted that by 2024, the world GDP would fall by 3% after factoring in the losses

imparted by COVID-19. To minimize all these detrimental and overarching effects of COVID-19, policy makers, and researchers from all fields are constantly searching for effective ways to reduce and contain the spread of the disease.

Human mobility and disease outbreak

One of the most important steps in fighting a highly infectious virus like COVID-19 is to limit its spread. Although there are various possibilities of how transmission can occur, the most effective way is by inhaling respiratory droplets containing the virus. Previous occurrences of pandemics, such as the bubonic plague in the 14th century and the cholera pandemics in the 19th century, can be related to the movement of people across the oceans in ships, which usually took several weeks (Markel, 2007). Thanks to the advent of efficient transportation modes, people can make numerous trips within a day and even reach the farthest end of the globe within a day or two. As a result, a highly contagious disease like COVID-19 can spread much faster across a larger region. When looking into the hotspots of COVID-19 in different countries, most of them can be traced to people traveling to or coming from infected areas and possible interactions with infected persons (Li et al., 2020b). But the progression of the infection followed a unique trait in each region. Different countries around the world have experienced the peak number in COVID-19 cases at different time periods and in the United States, not all the regions experienced a surge in the COVID-19 cases at the same time or at the same rate (Johns Hopkins, 2020; The Economic Tracker, 2020a). These regions have diverse socio-economic and travel characteristics that generate a range of mobility patterns, which can be translated to a number of possible human interactions. All this suggest the important role played by human mobility pattern in determining the spread of this disease. And as a result, human mobility pattern drew the attention of many researchers who attempted to understand its relationship with the spread of COVID-19.

Relationship between human mobility and COVID-19 outbreak

The high airborne transmission capacity of COVID-19 might suggest that areas with high population density will have a quicker spread. Although this is theoretically true, there are other factors at play that can slow or accelerate the spread. The findings from the study conducted by Hamidi et al. (2020b) attest to this statement. From their structural equation model, they found no significant effect of population density on the COVID-19 infection rate. Rather, mortality rates are found to be lower in high density, possibly due to the availability of better health-care services. In a follow-up longitudinal study by the same authors, they even found a negative significant relationship between density and infection rate when the size of the metropolitan area is controlled (Hamidi et al., 2020a). This study used a multi-level model, incorporating day and county-level data of the United States, and is also among the first which explained these findings with respect to the high degree of connectivity that is more likely in sizeable metropolitan areas.

There have been several studies in China which attempted to investigate the progression of COVID-19 from the epicenter and within other regions. With the help of a simulation model, Li et al. (2020b) analyzed the spread of COVID-19 through documented and undocumented individuals considering their travel pattern among 375 cities in China. The temporal range of the analysis was set to two weeks, starting from January 10, 2020, which coincides with the 2020 Spring Festival Period when a huge number of people travel across the country. Their analysis shows that 79% of the documented infections were caused by undocumented cases when no movement restrictions were imposed. In contrast to the previous study, Kraemer et al. (2020) and Chinazzi et al. (2020) analyzed the effectiveness of travel restrictions on the spread of COVID-19 across the cities in China. Both studies concluded that travel restrictions are highly effective when the spread is confined to a small area. Otherwise, public health policies, supplemented with movement

Table 1
Description of the variables and summary statistics (N = 3, 140).

Variable	Source	Description	Min	Max	Mean	Std. dev
<i>Socio-demographic Characteristics</i>						
Age 18 – 24 yrs.	U.S. Census Bureau (2018)	Percentage of people aged 18 to 24 years	0.52	46.98	8.69	3.45
Age 25 – 44 yrs.	U.S. Census Bureau (2018)	Percentage of people aged 25 to 44 years	11.43	41.18	23.43	3.29
Age 45 – 64 yrs.	U.S. Census Bureau (2018)	Percentage of people aged 44 to 64 years	10.80	42.99	26.84	2.86
Age 65 or above yrs.	U.S. Census Bureau (2018)	Percentage of people aged 65 or above years	3.20	56.71	18.79	4.66
African-Americans	U.S. Census Bureau (2018)	Percentage of African-Americans	0.00	87.40	8.93	14.47
Male	U.S. Census Bureau (2018)	Percentage of male people	41.39	79.00	50.08	2.37
Labor force	U.S. Census Bureau (2018)	Percentage of population in labor force	14.18	71.27	46.81	6.10
HH with Internet	U.S. Census Bureau (2018)	Percentage of HH having internet connection	35.15	96.27	75.98	8.78
<i>Location Characteristics</i>						
Metropolitan status	Ingram and Franco (2013)	Metropolitan status indicator based on 2013 NCHS urban–rural classification scheme: metro = 1, nonmetro = 0	0	1	0.37	0.48
Presence of airport	Federal Aviation Administration (2021)	Presence of an airport in a county: yes = 1, no = 0	0	1	0.15	0.35
Points of interests	Maryland Transportation Institute (2020)	Number of points of interests in a county for crowd gathering per 1000 people.	8	699	131.47	42.10
<i>Human Mobility</i>						
Stay at home	Maryland Transportation Institute (2020)	Daily percentage of residents staying at home (i.e., no trips with a non-home trip end more than one mile away from home). In the model, we considered the daily percentage of residents "going out of home" for various purposes, which was calculated as the complement of the percentage of residents staying at home.	2	61	20.99	4.94
Non-work trips per person	Maryland Transportation Institute (2020)	Daily average number of nonwork trips per person. Total number of identified and weighted trips on each day in each county divided by the county population (trip lengths lower than 300 m are dropped).	1.20	4.66	2.91	0.34
Person-miles traveled	Maryland Transportation Institute (2020)	Daily average person-miles traveled. Total weighted person-miles traveled on each day in each county across all travel modes (car, train, bus, plane, bike, walk, etc.) divided by the county population.	9.90	124.90	40.37	10.25
COVID-19 infection rate	Johns Hopkins Univ. data repository (Dong et al., 2020)	Number of COVID-19 daily new cases per 100K people (three-day moving average). Natural log form was used in the model.	0.00	117.75	1.70	3.99
<i>COVID-19 Status</i>						
Active cases	Maryland Transportation Institute (2020)	Number of active COVID-19 cases per 1000 people	0.08	7.35	1.93	1.36
Imported cases	Maryland Transportation Institute (2020)	Number of daily external trips by infectious persons from out of state/county	0	247829	4062	13508
<i>COVID-19 Policies</i>						
Stay at home order	Fullman et al. (2020)	Number of days from the first enactment of stay at home order until May 1, 2020	25	73	46.07	16.43
Public mask mandate	Fullman et al. (2020)	Number of days from the first enactment of public mask mandate until May 1, 2020	0	60	28.21	19.10
Quarantine mandate	Fullman et al. (2020)	Number of days from the first enactment of quarantine mandate until May 1, 2020	0	68	45.79	23.11
Gathering restrictions	Fullman et al. (2020)	Number of days from the first enactment of gathering restrictions until May 1, 2020	22	80	64.35	13.86
Emergency declarations	Fullman et al. (2020)	Number of days from the first enactment of emergency declarations until May 1, 2020	76	92	81.09	3.01
Non-essential business closure	Fullman et al. (2020)	Number of days from the first enactment of non-essential business closure until May 1, 2020	49	73	65.28	5.41

restriction measures, would be more effective in minimizing the spread. This indicates the need for a well-concerted set of policies tailored to the particular extent of the outbreak (Tian et al., 2020).

Outside China, Badr et al. (2020) conducted the first quantitative analysis of COVID-19 spread in relation to changes in mobility patterns. With the help of cell phone data, they represented mobility patterns using a county-level trip-based measure that indicates the degree of social distancing. The findings suggest a high positive correlation (0.7) between the reduction of infection and social distancing, which is most likely to be noticeable after 9 to 12 days. In a similar investigation, Carteni et al. (2020) estimated a multiple linear regression model using data obtained from sample interviews in Italy between January 2020 and March 2020. Despite having a much simpler measure of mobility, i. e., the number of persons making at least one trip per day, their observation of the 21 days lagged effect of mobility on infection rate supports the finding of the previous study. Mobility was also found to be the best predictor of the daily COVID-19 cases. Another study in Italy, conducted during the same time frame, estimated an average reduction of 45% infections and 200,000 hospitalization due to mobility restriction measures (Gatto et al., 2020).

Infection rate can also act as a determinant of mobility as suggested by Chinazzi et al. (2020) and Engle et al. (2020) when they found a negative correlation between infection rate and mobility after incorporating various factors that can affect the virus spread. Hu et al. (2021) obtained similar findings in their investigation of the factors affecting human mobility pattern during the COVID-19 pandemic with help of big data, which comprises more than 20 million person-level observations. The authors measured mobility with the help of three indicators, namely, percentage of people staying at home, non-work trips per person, and person-miles traveled. By isolating the policy interventions in a time-dependent modeling framework, they were able to conclude that infection rate has a significantly greater effect on the change in mobility compared to policy measures. But the effect of infection rate in reducing mobility diminished even further as the pandemic progressed, possibly due to the increased understanding of the disease (Kim and Kwan, 2021).

This study in the context of prior studies

All the aforementioned studies confirmed the strong effect of human mobility patterns on the extent of the COVID-19 outbreak. While they used different measures limited to one or two variables to represent mobility, none of the studies included multiple measures of mobility to evaluate their *combined* effect. Since a wide range of activity-travel characteristics collectively defines the overall mobility of a certain geographic region, it is important to characterize mobility appropriately to explore its relationships with disease transmission. In our study, we defined a latent “human mobility” factor from three different perspectives: 1) the percentage of people going outside home, which represented the overall out-of-home movement, 2) the number of non-work trips that reflected travel frequency, and 3) person-miles traveled, which described the average travel distance during the pandemic. We, thus, investigated the combined effect of mobility on disease transmission.

The relationship between human mobility and COVID-19 infection spread is bi-directional. For example, higher mobility might induce higher disease spread (positive effect from mobility to infection) whereas the higher spread of disease might restrict people’s mobility (negative effect from infection to mobility). While previous studies investigated either the former relationship or the latter one, we investigated the bi-directional relationships between mobility and infection spread in a comprehensive framework in our study.

A pandemic spread in a wide geographic area is arguably a very complex socio-biological phenomenon that is not only affected by human mobility but also influenced by a set of socio-demographic and location characteristics of an area, infection situation of neighboring areas, and community mitigation policies (policies to slow the spread of

the disease). In this study, we conceptualized a framework to capture all these relationships in a comprehensive manner by using structural equation modeling (SEM).

Data and sample

This section describes the data used in this study, their sources, and the rationale for choosing the study timeframe.

Data sources

This study used data from multiple sources which are listed in Table 1. It also shows the description of the variables, sources, and summary statistics. Following is the list of data sources used in this study:

- MTI COVID-19 Impact Analysis Platform (Maryland Transportation Institute, 2020)
- Johns Hopkins University COVID-19 infection data repository (Dong et al., 2020)
- Statewide COVID-related policy enactment dates (Fullman et al., 2020)
- U.S. Census Bureau (2018)

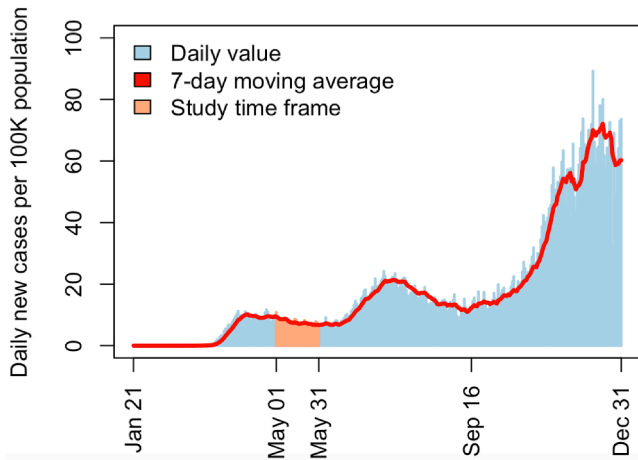
The MTI data is collected from the public web portal for our study timeframe (discussed later in this section). The MTI portal provides data on four major categories at the state- and county-level for each day from January 2020 to April 2021. These four categories are mobility and social distancing (e.g., staying at home, person-miles traveled, number of trips), COVID and health (e.g., new daily cases, active cases, tests done), economic impact (e.g., unemployment rate, working from home), and vulnerable population (e.g., people over 60, African-Americans). MTI collected privacy-protected mobile device location data and then applied a rigorous imputation algorithm to infer information from data on various socio-demographic, location, and travel characteristics. A rigorous multi-level weighing and validation of the imputed data were then integrated with the COVID data, and finally, data are made available to the online platform. For details, readers are referred to MTI’s web portal (<https://data.covid.umd.edu/>). The list of variables collected from the MTI portal, their definitions, and summary statistics are listed in Table 1.

We primarily used the MTI dataset for getting county-level mobility data, such as the proportion of people staying at home, average distance traveled, and average non-work trips made during the pandemic. On the other hand, to get data on daily COVID-19 infected cases, we used Johns Hopkins University (JHU) COVID-19 infection dataset. This dataset provides U.S. state- and county-level COVID data (cumulative confirmed cases, deaths, and recovered) since the first case in the U.S. in January 2020.

The state-level COVID-related policy enactment timelines were extracted from Fullman et al. (2020). The dataset lists various policies such as stay at home orders, non-essential business closures, bar and restaurant closures, mask mandates, gathering restrictions, and quarantine mandates with the announcement, enactment, and expiration dates. The socio-economic and location attributes of counties were obtained from the U.S. Census Bureau (2018) and the 2013 National Center for Health Statistics (NCHS) urban–rural classification data (Ingram and Franco, 2013). The full list of variables is provided in Table 1. This processing from multiple heterogeneous sources falls into the “variety” category of Big data analysis (IBM, 2020).

Data preparation

The county-level cross-sectional data for the whole month of May 2020 was considered in this study (the rationale for this choice is described in the next section). Daily data for each variable collected



Data Source: JHU COVID-19 data repository (Dong et al., 2020)

Fig. 1. Daily new COVID cases from January to November 2020 in the U.S. and study time frame.

from MTI and JHU COVID-19 data portal were collapsed into a single day value for the study month, one value for each county, to generate an average value per day, thus forming the cross-sectional data for our study time window. COVID-19 policy variables corresponded to the number of days from the first enactment date of a particular policy until the first day of the study window (May 1, 2020). This captures the time-lagged effects of state-level policies on human mobility and COVID-19 infection cases. Unlike the other variables, socio-demographic and location variables are not specific to the study time window because these data are neither available at a monthly time interval nor they are available for the year 2020 when the study was conducted. Considering that aggregate population data are not supposed to vary considerably within a short span of a few years, we used the most recent available data for these variables (2018 for U.S. Census Bureau and 2013 for NCHS

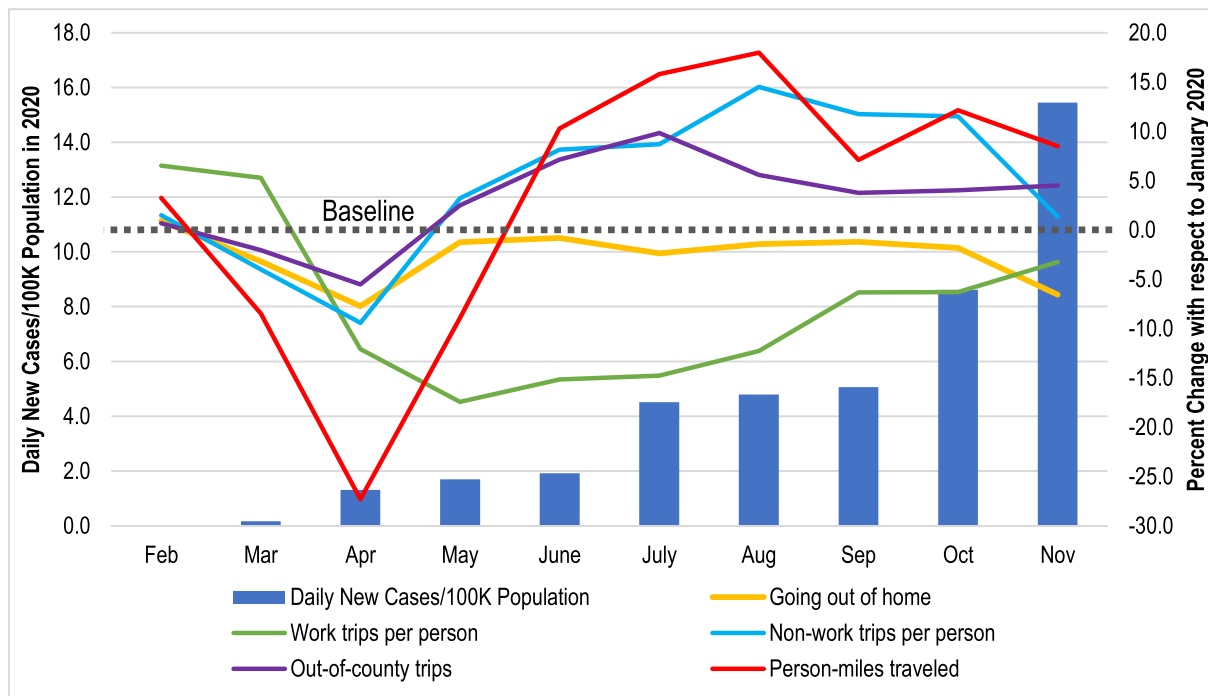
urban–rural classification data).

To account for the “spatial effect” of infection rate, a new variable was generated for each county that represented the infection rates of neighboring counties. Neighboring counties were defined as the counties that shared a common boundary with each county. These counties were identified using county-level geographic data from the U. S. Census Bureau (2019) and applying the ‘Spatial Join’ tool of ArcGIS Pro (ESRI, 2020). Note that this tool ensured the inclusion of a county as a neighbor even if the boundary only touches at a corner. On average there were six neighboring counties for each county and the maximum number reached 14 for San Juan County, Utah. After finding the neighboring counties, the average number of COVID-19 daily new cases per 100 K population in May 2020 was averaged over all neighboring counties for each respective county. Finally, we prepared the dataset for a total of 3,140 counties for our model.

Study timeframe and rationale for selection

With careful consideration, the timeframe for this study was selected as the month of May 2020 to capture the underlying interactions among the various indicators of mobility and COVID-19 infection rate. To understand this timeframe, we considered Fig. 1 that shows the trend of COVID-19 in the U.S. over 10 months. In terms of daily COVID-19 new cases, three peaks could be identified in this figure: at the beginning of April, at the end of July, and at the end of November. The selected timeframe started shortly after the first peak following the “first wave” chaotic movement in March – April 2020 during spring and prior to the summer (June) when a “second wave” began. Thus, the period appears to best capture the “flattening the curve” period of COVID infection rates and an expectation of relative stability in average mobility characteristics.

In addition, this timeframe constitutes a period when the virus had spread over 80 percent of U.S. counties, many response policies had been implemented, and most citizens and businesses appeared to accept the reality and to support public health guidelines. After this period, adherence to policy began to weaken at the beginning of the summer



Data Source: Maryland Transportation Institute (2020) and JHU COVID-19 data repository (Dong et al., 2020)

Fig. 2. Daily new COVID-19 cases and changes in mobility characteristics in 2020.

holidays, ultimately leading to the pandemic's second wave. All these events suggest that travel behavior began to stabilize in our study timeframe making it a suitable time window for developing the model and providing useful insights on the various interconnecting factors of COVID-19 infection rate and human mobility. Although this window appears stable at the aggregate level, it may not be the case in all counties in the U.S. A considerable degree of variation may well exist across counties and the county-level data indeed exhibited a discernable variation to enable estimating the postulated structural relationships between mobility and infection rate during the pandemic.

Changes in mobility characteristics during the pandemic

Fig. 2 shows the percentage changes in various mobility characteristics including the proportion of residents staying at home, average number of work and non-work trips made, person-miles traveled, and percentage of out-of-county trips in each month of the pandemic year 2020 with respect to the baseline value of January 2020 (pre-pandemic). Thus, a zero percent change of a mobility indicator suggests that its value is the same as it was in the pre-pandemic condition. The figure also shows the daily new infection cases per 100K population over the same period.

There was a sharp decline in all the mobility characteristics during the initial outbreak period of the pandemic followed by an increase towards the baseline. Then these characteristics were observed to exceed the baseline during the summer and fall 2020 period. The largest change was observed in person-miles traveled in April 2020 compared to other mobility features. An important distinction can be made between work and nonwork trips. Unlike other mobility characteristics, changes in work trips remained lower from the baseline throughout the year. This confirms the reduction in work trips due to the substitution of in-person work with telework by a considerably larger fraction of employed individuals and also due to the increase in the unemployment rate during the pandemic. Another concurrent study developed with a similar dataset also supported this finding by investigating a direct positive relationship between COVID-19 infection rate and working from home (Rafiq et al., 2022). On the other hand, after the initial decline, the changes in non-work trips increased gradually and then remained above the baseline value throughout the year. The increase in non-work trips might be due to the higher flexibility of scheduling non-work trips while working from home. Another reason might be due to the restriction of indoor social gatherings people tend to socialize more in outdoor locations, for example, parks, which ultimately resulted in more non-work travel.

Model specifications

Structural Equation Modeling (SEM) is a regression-based statistical modeling framework that can estimate the statistical relationships among a set of observed—as well as unobserved variables represented as latent factors—based on a specified theoretical model (Schumacker & Lomax, 2004; Kaplan, 2008). Such a structural model can capture the influences of the *exogenous* variables on the *endogenous* variables (regression effects) and the influences of endogenous variables on each other. SEM enables testing and evaluation of different models by specifying, estimating, and statistically testing hypothesized relationships among variables (Bentler, 1995; Zhang, 2018). The strength of SEM is that it examines the relationships between a variety of independent and dependent variables simultaneously. In addition, it allows for the possibility of examining the relationships among latent variables. The measurement error can also be incorporated in the models. Furthermore, SEM can simultaneously estimate the *direct* effect of one variable on another and the *indirect* effects through other mediating variables (the summation of the two that is called the *total* effects provides valuable insights on the interrelationships between variables).

Given the advantages of SEM methodology and the wide

proliferation of computer software packages capable of doing SEM construction and estimation (in Stata, R, or MPlus), SEM has become a popular method for testing or developing theories in non-experimental research (Byrne, 2006). Correlation, multiple regression, and analysis of variance (ANOVA) are other popular statistical techniques for analyzing relationships between observed variables; but none can estimate relationships between latent variables that SEM can do (Hoyle, 1995; Kline, 2016). SEM is widely used in travel behavior research: Golob (2003) outlined a comprehensive review of the application of SEM in various travel behavior research.

In this study, we use a general form of SEM, called the *Structural Regression (SR)* model (Kline, 2016). This kind of SEM model has two parts: (a) the *structural* part that represents hypotheses about direct and indirect effects among a set of *latent* and observed variables, and (b) the *measurement* part that denotes the relationships between a latent variable and a set of its associated *indicators* (Kline, 2016). The variables (observed or latent) can be either exogenous or endogenous in type (a simple way to separate them is that exogenous variables appear *only* on the right-hand side of an equation where endogenous variables can appear on both sides).

In this study, structural equation modeling (SEM) is used to quantitatively estimate direct and indirect effects between various mobility and socio-demographic factors and the rate of COVID spread across different counties in the United States for a certain time window. In addition to SEM model explaining the relationships between different factors, culminating in their effects (direct and indirect) on the infection rates, the constructed model can be utilized for predicting values of interest because the basic mechanism is regression and the variables of interest can be regressed from other variables that directly and indirectly affect the variables in question. In that, the model is able to predict the outcome in various scenarios, which we discuss in the results section later in the paper.

The conceptual model

The construction of the structural regression model involves conceptualizing possible relationships among a set of variables in a graphical construct where variables are represented as rectangles or ovals and the relationship between a pair of variables is denoted as directed arrows. Each of these arrows postulates a certain degree of effect from one variable to the other and the degree of these effects is determined by the respective *coefficients* (in their sizes, signs, and statistical significance). These coefficients are also known as factor loadings or model parameters. Given a graphical model, these model coefficients are estimated from data, which is referred to as the model estimation. The latent part of the structural regression model is called the *measurement model* where each of the latent factors (unobserved) is described by a set of indicator variables (observed), whereas the overall construct is called the *structural model* where the relationship among the set of exogenous and endogenous variables is represented by a series of regression equations. Fig. 3 depicts the conceptual structure of our proposed structural regression model, which represents the hypothesized relationships between COVID-19 infection rate and human mobility across U.S. counties.

In this construct, the “degree of human mobility” was represented as a *latent* factor characterized by three human mobility-related indicators. Here, this latent factor was considered as an *endogenous* variable as it was independent with respect to some variables and dependent with respect to other variables in the model (Acocck, 2013). Another endogenous variable in this hypothesized model was the “COVID-19 infection rate”. The infection rate of a county was measured in the unit of the number of COVID-19 daily new cases per 100 K population averaged over the entire month of May 2020 (the study period). Arguably, a number of variables might affect the infection rate either directly or indirectly via some other variables. We grouped these variables into three categories: (a) socio-demographic and location characteristics, (b)

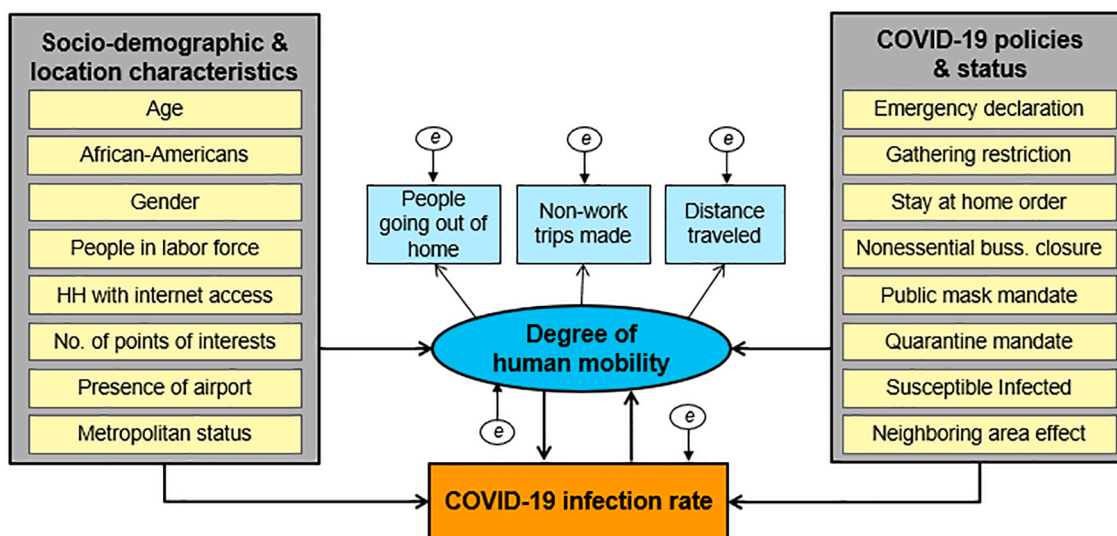


Fig. 3. Conceptual structural regression model.

COVID-19 policies and status, and (c) the degree of human mobility. The conceptual model considered the first two categories of variables as *exogenous* variables — variables that are not causally dependent on any other variables in the model. In Fig. 3, “socio-economic and location characteristics” and “COVID-19 policies and status” represent a collection of variables. When an arrow is directed from any of these two groups to another variable, it means that some but not necessarily all the variables in the group are directly connected to that particular variable (those exogenous variables that are directly connected to a particular variable are provided in the corresponding tables in the results section). Note that model variables were selected based on relevant prior work and data availability. The full list of all these variables with the associated data sources and relevant summary statistics is presented in Table 1. In the following, we described all the hypothesized connections of the conceptual model in detail.

Measurement model: Degree of human mobility

The degree of human mobility, a latent variable in the model, collectively represented the overall degree of human movement in the study area. We used three travel-related indicators to describe human mobility: percentage of residents going out of home, the number of non-work trips made per person per day, and the person-miles traveled. Here, the first indicator represented the proportions of people in a given county that went out of home for various purposes. This was calculated as the complement of the percentage of residents staying at home which accounted for the percentage of people who did not make any trips more than one mile away from home. The person-miles traveled denoted the average miles traveled per person per day on all modes. Positive effects were hypothesized between the mobility factor and all the three observed mobility indicators. We chose these three indicators to describe human mobility from different perspectives. The percentage of people going outside home represents overall out-of-home movement, the number of non-work trips reflects non-work travel frequency, and person-miles traveled describes the average distance traveled during the pandemic. Similar mobility descriptors are used in Hu et al. (2021).

We postulated a positive effect from the degree of human mobility to the COVID-19 infection rate anticipating that higher mobility of people in a certain area within a certain time window would associate with a higher spread of disease in that area since the disease is generally carried out by humans and transmitted through humans to humans when an infected individual (with or without symptoms) comes close to another individual (Centers for Disease Control and Prevention, 2020a).

Structure model: Effects on COVID-19 infection rate

In addition to the degree of human mobility, the COVID-19 infection rate was also hypothesized to be directly influenced by various county-level socio-demographic and location characteristics, statewide COVID-19 policies, and the status of the pandemic. “Socio-demographic characteristics”, expressed as proportions of the total population, included the number of males, African-Americans, and people with specific age groups. We considered four variables representing four different age groups: 18 – 24, 25 – 44, 45 – 64, and 65 + years. As a location variable, we included the metropolitan status of a county in the model. The influence of the proportions of the elderly population (aged 65 or above) in a county on the COVID-19 infection rate was hypothesized as twofold: counties with a greater proportion of elderly people might have a greater infection rate as more elderly people might have less immune power, making them more vulnerable to the disease (positive effect). On the other hand, elderly people might be more cautious about the disease and more abide by the mitigation strategies compared to the younger group, which might result in less infection rate (negative effect). A positive association is anticipated for metropolitan counties as higher density areas might pose a greater possibility of close human interactions, which significantly increases the risk of COVID-19 exposure. Regarding race, we hypothesized that counties having a higher fraction of African-Americans might be positively associated with higher infected counties because of their higher risk of COVID-19 exposure (Centers for Disease Control and Prevention, 2021a).

Another set of exogenous variables was related to various “COVID-related policies”, which were measured as the number of days from the first enactment of a policy until May 1, 2020 (starting date of the study period). The policies that are assumed to influence the infection rate of a county included stay at home order, public mask mandate, and quarantine mandate. All these policy variables are postulated to negatively affect the infection rate assuming that a longer duration of enactment of these policies would associate with a higher reduction in the infection rates.

We considered two county-level measures as “COVID-19 status” indicators. The *first* one was an interaction term, called the Susceptible-Infected (SI) term, which is inspired by the classical SIR (susceptible, infected, and recovered) model of epidemics. According to the SIR model (Wikipedia, 2021), the rate at which the number of infected people in an area increase is proportional to the product of the number of susceptible people (people who have not been infected with the virus yet, hence susceptible to contract the virus) and the number of currently infected people in that population (current active cases). This is aligned

with the observation that the virus only spreads from currently infected people to people who did not have it before (not the one who already got it and recovered). So, the rate of virus spread (the infection rate) is influenced by these two numbers jointly. To consider this effect, we incorporated the interaction SI term that is calculated as the product of the fraction of people susceptible in a country (total population minus the total number of people infected until May 1, 2020) and the fraction of active cases (total infected minus total recovered minus total deaths in May 2020) in that county. A positive association from the SI term to the infection rate is postulated. The *second* COVID-19 status indicator considered in the model was also an interaction term, called “spatial effect”. The infection rate of a county might be influenced by the infection situations of neighboring (shared common boundaries) counties because if a county was surrounded by other counties that have higher numbers of infected cases and if that county received infected cases from those highly infected counties, the subject county might become prone to have a higher infection rate. To accommodate this effect, we considered an interaction term, which was a product of the number of average COVID-19 daily new cases per 100 K population in neighboring counties for a respective county and the average number of daily external trips by infectious persons from out of state or county during the study period.

Structural model: Effects on human mobility

A set of socio-demographic and location characteristics as well as COVID-19 policy variables were postulated to affect human mobility in a county. The socio-demographic characteristics were the proportion of African-Americans, males, four age groups, labor force, and households with internet connections. The outbreak of the COVID-19 pandemic and its subsequent activity-travel restrictions led people to stay at home and work from home (telecommuting). Since a higher fraction of African-Americans were part of the essential workers and commuter groups during the pandemic (Rafiq and McNally, 2022; Centers for Disease Control and Prevention, 2021a), we postulated a positive connection between the proportion of African-Americans and mobility anticipating that more commuting by this group of people would involve more trips and consequently higher mobility. Similarly, we assumed a positive effect from male proportion to mobility because recent studies (Beck and Hensher, 2020; Rafiq and McNally, 2022; Brynjolfsson et al., 2020) reported that the commuter group consisted of a higher fraction of male workers during the pandemic. In addition, a higher proportion of the labor force in a county would involve more commute trips and higher mobility (positive effect). On the other hand, a higher proportion of households with internet access would associate with more work from home and online shopping, which would lead to less mobility (negative effect).

In addition to socio-demographics, three county-level location characteristics were hypothesized to influence mobility: number of points of interest for crowd gathering, presence of an airport, and metropolitan status. Being a metropolitan county was assumed to negatively affect mobility. This is because as more people worked from home in metropolitan areas during the pandemic (Rafiq and McNally, 2022), metro counties would experience a reduction in work trips and therefore, a reduction in mobility within these areas. Another reason might be between March and May 2020, the COVID-19 pandemic was highest among the residents of large metropolitan areas and the infection then shifted to a rapid surge to small metropolitan and non-metropolitan areas (Centers for Disease Control and Prevention, 2020c; Cuadros et al., 2021). Due to higher infection cases, metropolitan areas had comparatively more policy restrictions on movement during the study period. Similarly, metropolitan areas where there is a higher chance of having places with more points of crowd gatherings, as well as airports, might have similar negative associations with mobility.

Finally, several COVID-19 policy variables including gathering restrictions, stay at home order, emergency declaration, and non-essential business closure were hypothesized to negatively affect mobility

assuming that the longer the duration of policies in action, the lower the mobility in a respective county.

Structural model: Interactions between human mobility and COVID-19 infection rate

We postulated a positive effect between the degree of human mobility in a county and the infection rate of that county. A county having a lower degree of mobility (e.g., fewer people going out of home, fewer non-work trips made) might cause fewer opportunities for people to interact with other people, which might reduce the chances of getting infected by the disease. On the other hand, a higher degree of mobility would cause more people to get infected due to higher chances of interactions.

While human mobility might affect the spread of the infection rate in one way, the infection rate might influence human mobility in the other way but with a reverse effect (*negative* association). For example, a county having a higher infection rate might experience a reduction in the number of people going outside home as well as a decrease in the total distance traveled due to a more restrictive choice of movement to avoid the potential spread of the disease. The converse might also happen: a county experiencing fewer COVID-19 cases might exhibit a higher movement of people and consequently a higher fraction of people not staying home (negative association between infection rate and mobility). Since the structural regression model is a directional model, we accommodated the above-mentioned two associations by introducing a *feedback* connection from the COVID-19 infection rate to the human mobility latent factor (this kind of structure is known as a non-recursive model). With this feedback connection in place, the model simultaneously estimated how human mobility might influence the COVID-19 spread as well as how the disease spread could affect human mobility in reverse. As discussed, our model postulated a positive effect in the forward direction (mobility to infection rate) and a negative effect in the reverse direction (infection rate to mobility). Finally, two error-term covariances between the percentage of people going out of home and the number of non-work trips made and between the average distance traveled and the number of non-work trips made were added to the model.

The structural regression model

The conceptual model discussed in the previous section can be represented mathematically in terms of a set of simultaneous equations formulated for the endogenous variables. The equations for the two main endogenous variables, degree of human mobility and COVID-19 infection rate, in the structural model can be expressed with the help of Eqn. (1) and (2). These two equations capture the feedback effect between human mobility and infection rate by having their corresponding terms (η_m and Y_c) as a predictor in the other equations.

$$\eta_m = B_m Y_c + \Gamma_m X + \delta_m \quad (1)$$

and

$$Y_c = B_c \eta_m + \Gamma_c X + \delta_c \quad (2)$$

where

η_m vector of latent endogenous variables for degree of human mobility;

Y_c vector of observed endogenous variables for COVID-19 infection rate;

X vector of observed exogenous variables representing socio-demographics & location characteristics, and COVID-19 policies & status;

B_m matrix of coefficients representing direct effects from observed endogenous variable (Y_c) to latent endogenous variable (η_m);

B_c Matrix of coefficients representing direct effects from the latent endogenous variable (η_m) to the observed endogenous variable (Y_c);

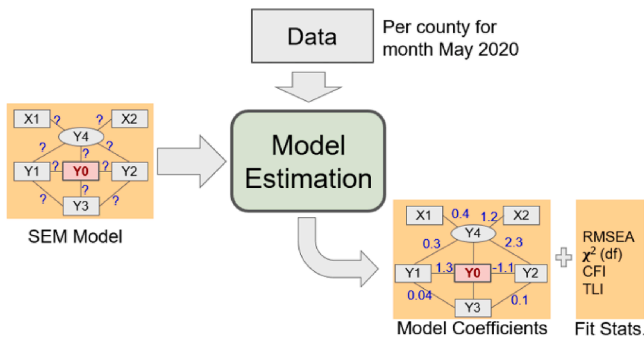


Fig. 4. SEM model estimation.

Γ_m matrix of coefficient representing direct effects from the observed exogenous variables to the latent endogenous variable (η_m);

Γ_c matrix of coefficient representing direct effects from the observed exogenous variables to the observed endogenous variable (Y_c);

δ_m, δ_c vector of error terms for degree of human mobility (m) and COVID-19 infection rate (c), respectively.

The set of equations for the three indicators in the measurement model of degree of human mobility (η_m) is given by Eqn. (3).

$$Z_i = \Lambda_i \eta_m + \epsilon_i \quad (3)$$

where

Z_i vector of observed variable (indicator) i for the latent variable (η_m);

Λ_i matrix of pattern coefficients for indicator i , representing its loading for the latent variable (η_m);

ϵ_i vector of measurement error terms for indicator i .

Other parameter matrices include the covariance matrix of the measured exogenous factors Φ , the covariance matrix for the disturbances of endogenous factors on each other ψ , the covariance matrix of error terms of the two endogenous variables Θ_δ , the covariance matrix of error terms of the indicators Θ_ϵ . The population covariance matrix of the observed variables, denoted by Σ , can be expressed as a function of a set of parameters θ , shown in Eq. (4) (Lu and Pas, 1999).

$$\Sigma = \Sigma(\theta) \quad (4)$$

Here, θ represents the model parameters of B , Γ , Θ_ϵ , Λ_i , Φ , and Ψ . These unknown parameters are estimated such that the difference between the sample covariance matrix S and the model implied covariance matrix $\Sigma(\theta)$ is minimized.

The estimation of the model

After conceptualizing the model, we estimated the model coefficients using U.S. county-level data (as shown in Fig. 4). Before estimating the model, a Cronbach’s alpha or score reliability check was performed for the latent construct. This test measures the internal consistency reliability or the degree to which responses are consistent across the items of a latent measure. If internal consistency is low, then the responses may be very heterogeneous that the total score is not the best possible unit of analysis. On the other contrary, if internal consistency reliability is higher, then it represents the consistency among the responses. From the unstandardized solution, Cronbach’s alpha value is obtained as 0.7, which is considered ‘adequate’ by Kline (2016). We performed a confirmatory factor analysis (CFA) separately to check the goodness-of-fit of the latent construct of “degree of human mobility”. The fit statistics appeared satisfactory ($\chi^2 = 0.00$, RMSEA = 0.00, CFI = 1.0, and TLI = 1.0).

Based on our conceptual structure (Fig. 3) and the best possible combination of exogenous variables, we estimated the *Structural Regression (SR) Model* using the Maximum Likelihood Missing Value (MLMV) estimation method. Note that for this study, we used U.S.

county-level data for May 2020. We checked the distribution of all the model variables and tried to reduce the non-normality in data in appropriate cases (e.g., infection rate) by taking the logarithm of those variables. There are some missing values on COVID-19 policy variables, which are typically better addressed by the MLMV estimation method. Another advantage of this method is that it produces robust standard errors and provides both a mean- and variance-adjusted chi-squared test (Van Acker et al., 2014).

In this paper, we reported several goodness-of-fit measures. First, χ^2 statistic, which tests whether the observed covariance matrix and the model implied covariance matrix are equal. Smaller χ^2 value with high p -value (p -value greater than 0.05) indicates better model fit. However, χ^2 value tends to increase with sample size so models with larger sample sizes might show larger χ^2 value and subsequently may lead to rejection of an otherwise good model (Van Acker and Witlox, 2011; Acock, 2013; Kline, 2016). Since the χ^2 value increases with the sample size, it is not an appropriate measure of fit, but as the basis of other goodness-of-fit measures, it is always reported anyway (Byrne, 2001; Cao et al., 2007). With a comparatively larger sample size (3,140), we obtained a larger χ^2 value (1,477 at the degree of freedom 227) with a lower p -value (0.000).

Another fit statistic is Root Mean Square Error Approximation (RMSEA), which measures the estimated discrepancy between the model implied and true population covariance matrix controlling for sample size (Cao et al., 2007). In other words, it measures how much error there is for each degree of freedom. This measure penalizes for unnecessary added complexity. A complex model may fit better as it capitalizes on chance, but RMSEA adjusts for this (Acock, 2013). Therefore, RMSEA is often considered as a robust goodness-of-fit measure in covariance structure modeling (Byrne, 1998). We obtained an RMSEA value of 0.042 (less than 0.05) for our model, which indicates a satisfactory model fit. In addition, there are two incremental fit indices called Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI). These are the assessment of the improvement of the hypothesized model compared to the independence model with unrelated variables (Kline, 2016; Hu and Bentler, 1999). Compared to CFI, a TLI imposes a greater penalty for model complexity. Since CFI and TLI values are highly correlated only one of these two fit statistics is recommended to be reported (Kline, 2016; Kenny, 2020). The acceptable threshold value for TLI is greater than 0.9 (Acock, 2013; Van Acker and Witlox, 2011) and we got 0.93 from our model. Based on these fit measures, it can be concluded that our postulated model of causal structure fits reasonably well with the data and represents a close approximation of the population.

Results and discussion

The results from the estimated structural regression model are discussed in four broad sections: 1) Latent measure of human mobility 2) Effects on COVID-19 infection rate 3) Effects on human mobility factor and 4) Interactions between mobility and infection rate. The unstandardized and standardized coefficients of *direct* and *total* effects are presented for all the analyses. If not otherwise stated, the effects mentioned in the discussion represent direct effects. Note that an

Table 2
Estimated factor loadings for the latent measure of human mobility (N = 3,140).

Measurement model	Unstandardized coefficient	Standardized coefficient
Latent factor: Human mobility		
Indicators		
Percentage going out of home ^a	1***	0.671***
Avg. non-work trips made	0.132***	0.127***
Person-miles traveled	2.102***	0.679***

Note: ^a refers to fixed parameter and *** indicates statistical significance at 1%.

unstandardized coefficient for a variable pair represents the number of units of changes in the indicator variable due to one unit of change in the latent variable. On the other hand, standardized coefficients represent the same effect as that produced by the unstandardized coefficients, but in the units of standard deviations — change in the number of standard deviations in the indicator variables due to one standard deviation change in the latent variable (Kline, 2016).

Latent measure of human mobility

The measurement model estimated county-level human mobility based on indicators including the percentage of people going out of home, average number of non-work trips made, and average person-miles traveled. The estimated factor loadings, standardized and unstandardized, are shown in Table 2.

The unstandardized coefficients that are scaled with respect to the percentage of people going out of home (called the fixed-parameter) represent the unit change in the indicators for a unit increase in human mobility. In unstandardized solution, since the coefficients are not normalized, they can be interpreted as regression coefficients (Kline,

2016). For example, for the indicator “average number of non-work trips made”, the coefficient value of 0.132 indicates a 0.132 point increase in this indicator for every one-point increase in the “human mobility” factor. As anticipated, all the indicators have a positive association with mobility and are statistically significant at the 1% level of significance. The positive coefficients indicate that when mobility increases, the value of the indicator increases or vice versa.

Since standardized coefficients treat all variables as having a variance of one, they become independent of the scales of both indicator and latent variables, which means that standardized coefficients can be used to directly compare the relative strengths of the factor loadings (Kwan and Chan, 2011). The standardized coefficient for the variable “average non-work trips made” in the measurement model indicates that there is 0.127 standard deviation change in the number of non-work trips made when the “human mobility” latent factor changes by one standard deviation. Among the three indicators, this one appeared to be the weakest measure of human mobility. The magnitudes of the factor loadings of the other two indicators were very close (slightly higher for person-miles traveled). The standardized coefficient of person-miles traveled is 0.679, which implies that the latent factor, human mobility, explained 0.679² or 46 percent of the observed variance of person-miles traveled (the square of standardized coefficients are the proportion of explained variance) (Kline, 2016).

Table 3
Direct and total effects of variables on COVID infection rate (N = 3,140).

Structural model	Unstandardized coefficient		Standardized coefficient	
	Direct effect	Total effect	Direct effect	Total effect
Outcome: COVID-19 infection rate				
Predictors				
Human mobility	20.116***	18.191***	0.273***	0.246***
Socio-demographic characteristics				
Age 18–24 years	6.388***	1.060	0.090***	0.015
Age 25–44 years	16.377***	5.752***	0.220***	0.077***
Age 45–64 years	−1.437	−3.358*	−0.017	−0.039*
Age 65 years and older	−4.117**	−7.411***	−0.078**	−0.141***
African-Americans	1.936***	2.275***	0.115***	0.135***
Male	−13.115***	−4.604**	−0.128***	−0.045**
People in labor force	—	1.408***	—	0.035***
HH with internet access	—	−1.607***	—	−0.058***
Location characteristics				
Number of points of interest	—	−0.002***	—	−0.039***
Presence of airport	—	−0.266***	—	−0.039***
Metropolitan status	0.500***	0.332***	0.099***	0.066***
COVID-19 status				
^a Susceptible	0.0003***	0.0002***	0.150***	0.136***
Infected (SI)	—	—	—	—
^b Spatial effect	0.051***	0.046**	0.314***	0.284**
COVID-19 policies				
Emergency declaration	—	−1.406***	—	−0.017***
Gathering restriction	—	−0.029***	—	−0.002***
Stay at home order	2.942***	1.856***	0.198***	0.125***
Non-essential business closure	—	−0.606**	—	−0.013**
Public mask mandate	−0.477*	−0.431*	−0.037*	−0.034*
Quarantine mandate	0.145	0.132	0.013	0.012

Notes: — denotes no direct connections. *, **, and *** indicate statistical significance at 10%, 5%, and 1% respectively.

All the COVID-19 policy variables represent the number of days from the first enactment of policies until May 1, 2020

^a Susceptible Infected (SI) = number of susceptible people (people who have not been infected with the virus yet) × number of currently infected people in that population (current active cases).

^b Spatial effect = number of average COVID-19 daily new cases per 100 K population in neighboring counties × average number of daily external trips by infectious persons from out of state or county (Imported cases)

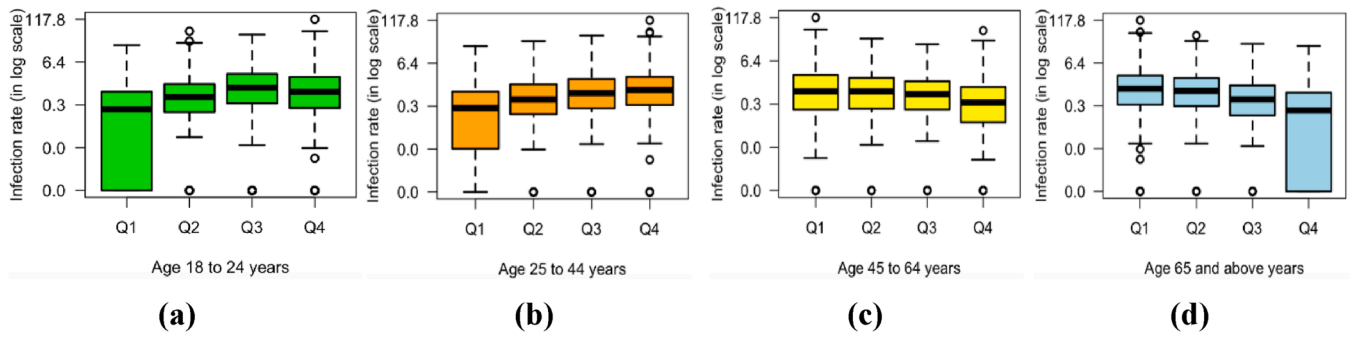
Structural model: Effects on COVID-19 infection

The unstandardized and standardized coefficients of the structural model variables are shown in Table 3. Both direct and total effects are reported, where the former only indicates the direct influence of a variable and the latter includes both the direct and indirect effects through other mediating variables.

The structural model infers a positive association between mobility and COVID-19 infection rate in a county. That means, the counties having a higher degree of mobility experienced a higher value of infection rate. As per our model construction, there was a feedback connection from infection rate to mobility. So, the interactions of mobility and infection are indeed both ways that call for a separate discussion, which is provided in a later subsection (Section 5.4). In the following, we discussed the effects of other variables of the model on the infection rate.

The presence of the younger population groups (age 18 – 24 and 25 – 44 years) had a positive effect on the infection rate whereas the other two population groups aged 45 – 64 and 65 or above years had negative effects on the infection rate. The results are consistent with prior studies. Boehmer et al. (2020) found that from May to July 2020, the COVID-19 cases increased the largest among the people aged less than 30 years and the median age of confirmed cases in May 2020 was 37 years.

Prior studies reported two major reasons behind the higher risk of exposure to the COVID-19 infection by younger adults: their jobs in frontline occupations and highly exposed industries (Rho et al., 2020; Dey et al., 2020) and their less tendency to follow various mitigation measures (Czeisler et al, 2020; Nagata, 2020). The proportion of elderly people (65 or above years) was directly and negatively associated with the infection rate. Findings from other studies supported this observation with the claim that people in the older age group had lower exposure to spread the disease than the younger population (Oster et al., 2020; Park et al., 2020). To make an empirical validation of age effects based on data, we constructed four quantile boxplots, as shown in Fig. 5. Each of the boxplots show the infection rates of counties with a certain proportion of age group. In each plot, counties were split into four groups based on the ranges of their quantile values of population proportion; for example, Q1 denotes counties having below 25-percentile value, Q2 contains above 25-percentile but below 50-percentile, and so on (each box effectively contains the equal number of counties). It is observed that counties having a larger share of 18 – 24 and 25 – 44 age groups had indeed higher infection rates whereas counties having a



Data Source: U.S. Census Bureau (2018) and JHU COVID-19 data repository (Dong et al., 2020)

Fig. 5. Relationships between age and infection rate across counties by four age groups: (a) 18 – 24 yrs. (b) 25 – 44 yrs. (c) 45 – 64 yrs. (d) 65 yrs. or above.

larger share of 45–64 and 65 + age groups had lower infection rates as the median values (the central line inside the box) increased in higher quantile boxes in the first two plots (Fig. 5a and 5b) but declined in the last two plots (Fig. 5c and 5d). This validates our results on the relationships between age groups and infection rate.

As anticipated, the infection rate of a county was positively associated with the proportion of African-Americans in a county. This is consistent with the findings that African-Americans were at higher risk of COVID-19 infection due to their employment in essential jobs (e.g., grocery stores, public transit, and health care facilities), crowded living condition, healthcare disparities, and reliance on public transportation (Centers for Disease Control and Prevention, 2021a; Kullar et al., 2020; Wright and Merritt, 2020; Kemp et al., 2020). Counties with a higher proportion of male population had a lower infection rate. According to the CDC COVID data tracker (Centers for Disease Control and Prevention, 2021b) the percentage of infected cases was higher for women compared to men (52.3% vs. 47.7%) whereas men had higher death cases than women (54.3% vs. 45.7%). The fraction of households with internet connections positively affected infection rate indirectly via the mobility factor, which implies that people living in technologically advanced counties were able to do more teleworking and online shopping, consequently traveled less and spread the disease less.

Regarding the effect of the metropolitan status on the infection rate, we observed that compared to non-metropolitan counties, metropolitan counties were associated with a higher infection rate during our study period (May 2020). Centers for Disease Control and Prevention (2020c) reported similar findings for the U.S. counties. Carozzi et al. (2020), Kadi and Khelifaoui (2020), and Sy et al. (2020) also found a positive association between density and infection cases in different countries.

As expected, the two interaction terms representing COVID-19 status positively affected the infection rate of a county. It suggests that a county’s infection rate was positively influenced by the proportion of the susceptible population in that county (susceptible infected) and the inflow of infected people coming from its highly infected neighboring counties (spatial effect). Considering the magnitude of the standardized total effect of the “spatial effect” variable on the infection rate it can be concluded that this variable in fact had the highest impact on the infection rate of a county. The relationship suggests that spaces are connected and what happens in one county does not remain contained within the boundary of that country, rather the effect spills over onto the nearby other counties. A larger ramification of this effect in policies is that counties need to coordinate their countermeasures and actions, at least with neighboring counties, so that the spatial entanglement can be taken care of (e.g., a school closure decision by some county may not bring forth the desired result in reducing COVID-19 spread unless a similar action is taken by other adjacent counties). This is an important observation of our study.

Regarding COVID-19 policies, counties with a longer period of public

Table 4

Direct and total effects of variables on human mobility (N = 3,140).

Structural model	Unstandardized coefficient		Standardized coefficient	
	Direct effect	Total effect	Direct effect	Total effect
Outcome: Human Mobility				
Predictors				
COVID-19 infection rate	-0.005***	-0.005***	-0.388***	-0.351***
Socio-demographic Characteristics				
Age 18–24 years	-0.259***	-0.265***	-0.270***	-0.276***
Age 25–44 years	-0.498***	-0.528***	-0.495***	-0.525***
Age 45–64 years	-0.113***	-0.095***	-0.098***	-0.083***
Age 65 years and older	-0.203***	-0.164***	-0.285***	-0.230***
African-Americans	0.029***	0.017***	0.126***	0.074***
Male	0.399***	0.423***	0.287***	0.304***
People in labor force	0.077***	0.070***	0.143***	0.129***
HH with internet access	-0.088***	-0.080***	-0.234***	-0.212***
Location Characteristics				
Number of points of interest	-0.0001***	-0.0001***	-0.158***	-0.143***
Presence of airport	-0.015***	-0.013***	-0.156***	-0.141***
Metropolitan status	-0.007***	-0.008***	-0.096***	-0.122***
COVID-19 status				
^a Susceptible	—	-1.290e-6	—	-0.053
Infected				
^b Spatial effect	—	-0.0002***	—	-0.110***
COVID-19 policies				
Emergency declaration	-0.077***	-0.070***	-0.070***	-0.063***
Gathering restriction	-0.002	-0.001	-0.007	-0.006
Stay at home order	-0.044***	-0.054***	-0.228***	-0.269***
Non-essential business closure	-0.033**	-0.030**	-0.051**	-0.046**
Public mask mandate	—	0.002*	—	0.013*
Quarantine mandate	—	-0.001	—	-0.005

Notes: — denotes no direct connections. *, **, and *** indicate statistical significance at 10%, 5%, and 1% respectively.

All the COVID-19 policy variables represent the number of days from the first enactment of policies until May 1, 2020.

^a Susceptible Infected (SI) = number of susceptible people (people who have not been infected with the virus yet) × number of currently infected people in that population (current active cases).

^b Spatial effect = number of average COVID-19 daily new cases per 100 K population in neighboring counties × average number of daily external trips by infectious persons from out of state or county (Imported cases).

mask mandate had a lower infection rate. This is evidence of the effectiveness of mask usage for mitigating community transmission of the COVID-19 disease. Other studies also reported that face mask usage could result in a large reduction in the spread of the infection (Kaufman et al., 2020; Chu et al., 2020). The direct effect of quarantine mandate on infection rate did not appear significant. All the other policy variables including the emergency declaration, gathering restrictions, and non-essential business closure had negative indirect effects on the infection rate via the mobility factor. It suggests that counties that enacted these policies for longer periods had relatively lower mobility and consequently lower infection rates during the study period, which necessarily translates into the efficacy of these measures. However, we observed a positive direct effect from the longer period of stay at home order and infection rate, which is counterintuitive. Detailed discussion on this is provided later in the policy section 7.2.

Structural model: effects on human mobility

The direct and total effects of socio-demographic and location characteristics and COVID-19 policies on the latent human mobility factor as well as on various observed mobility indicators (e.g., trips per person and person mile traveled) are discussed in this section.

Effects on latent mobility factor

The unstandardized and standardized coefficients of direct and indirect effects of variables on human mobility are shown in Table 4. Counties with a higher proportion of people at any age group and with more households with internet access were less likely to make trips, in other words, more likely to stay at home during the study period (negative effects on mobility). On the other hand, as hypothesized, counties with a higher proportion of African-Americans, males, and people in the labor force were more likely to travel (positive effects on

mobility). Hu et al. (2021) reported similar findings for African-Americans. As anticipated, metropolitan counties, counties with an airport, and more points of interest of crowd gatherings were more likely to follow the stay at home order and other movement restriction strategies compared to other counties (negative effect on mobility). The COVID-19 policies including emergency declaration, stay at home order, and non-essential business closure were negatively associated with mobility. It suggests that counties with longer periods of community mitigation strategies in action were more likely to abide by the orders and less likely to travel, more specifically in May 2020. This finding is consistent with Moreland et al. (2020).

Effects on observed mobility indicators

In addition to obtaining the impact of variables on the latent mobility factor, we obtained the effects of the variables on the observed mobility

Table 6
Interactions between human mobility and COVID infection (N = 3,140).

Structural model	Unstandardized coefficient		Standardized coefficient	
	Direct effect	Total effect	Direct effect	Total effect
Outcome: COVID infection rate				
Predictors: Human mobility	20.116***	18.191***	0.273***	0.246***
Outcome: Human mobility				
Predictors: COVID infection rate	-0.005***	-0.005***	-0.388***	-0.351***

Note: *, **, and *** indicate statistical significance at 10%, 5%, and 1% respectively.

Table 5
Total effects of variables on mobility indicators (N = 3,140).

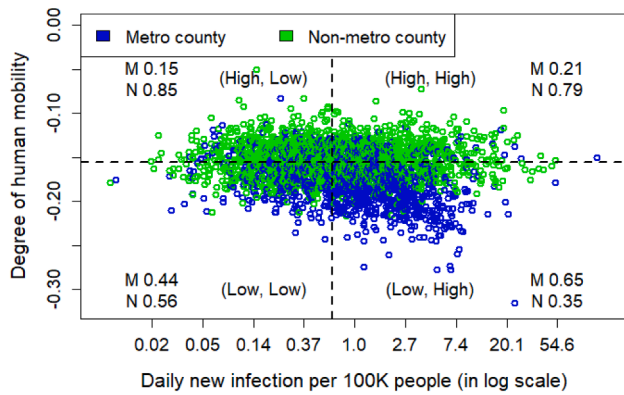
Predictors	Outcome variables					
	Going out of home	Non-work trips	Person-miles traveled	Going out of home	Non-work trips	Person-miles traveled
	Unstandardized Coefficient			Standardized Coefficient		
COVID infection rate	-0.005***	-0.001***	-0.010***	-0.236***	-0.045***	-0.238
Human mobility	0.904***	0.119***	1.901***	0.607***	0.115***	0.614
<i>Socio-demographic Characteristics</i>						
Age 18–24 years	-0.265***	-0.035***	-0.557***	-0.185***	-0.035***	-0.188
Age 25–44 years	-0.528***	-0.070***	-1.110***	-0.352***	-0.067***	-0.356
Age 45–64 years	-0.095***	-0.013***	-0.201***	-0.055***	-0.011***	-0.056
Age 65 years and older	-0.164***	-0.022***	-0.344***	-0.154***	-0.029***	-0.156
African-Americans	0.017***	0.002***	0.035***	0.049***	0.009***	0.050
Male	0.423***	0.056***	0.889***	0.204***	0.039***	0.206
People in labor force	0.070***	0.009***	0.147***	0.087***	0.016***	0.088
HH with internet access	-0.080***	-0.011***	-0.168***	-0.142***	-0.027***	-0.144
<i>Location Characteristics</i>						
Number of points of interest	-0.0001***	-1.48e5***	-0.0002***	-0.096***	-0.018***	-0.097
Presence of airport	-0.013***	-0.002***	-0.028***	-0.095***	-0.018***	-0.096
Metropolitan status	-0.008***	-0.001***	-0.018***	-0.082***	-0.016***	-0.083
<i>COVID-19 status</i>						
^a Susceptible Infected	-1.290e-6	-1.700e-7	-2.71e6***	-0.035	-0.007	-0.036
^b Spatial effect	-0.0002***	-3.19e5***	-0.001***	-0.074***	-0.014***	-0.075
<i>COVID-19 policies</i>						
Emergency declaration	-0.070***	-0.009***	-0.147***	-0.043***	-0.008***	-0.043
Gathering restriction	-0.001	-0.0002	-0.003	-0.004	-0.001	-0.004
Stay at home order	-0.054***	-0.007***	-0.113***	-0.180***	-0.034***	-0.182
Non-essential business closure	-0.030**	-0.004**	-0.063**	-0.031**	-0.006**	-0.031
Public mask mandate	0.002*	0.0003*	0.005*	0.009*	0.002*	0.008
Quarantine mandate	-0.001	-9.140e-5	-0.001	-0.003	-0.001	-0.003

Notes: *, **, and *** indicate statistical significance at 10%, 5%, and 1% respectively.

All the COVID-19 policy variables represent the number of days from the first enactment of policies until May 1, 2020.

^a Susceptible Infected (SI) = number of susceptible people (people who have not been infected with the virus yet) × number of currently infected people in that population (current active cases).

^b Spatial effect = number of average COVID-19 daily new cases per 100 K population in neighboring counties × average number of daily external trips by infectious persons from out of state or county (Imported cases).



Data Source: Ingram and Franco (2013) and JHU COVID-19 data repository (Dong et al., 2020)
 Note: Values for degree of mobility were estimated by the model

Fig. 6. County distribution based on human mobility and infection rate by metropolitan status (dotted lines are the median lines).

indicators. Table 5 lists the predictors and their unstandardized and standardized total effects on the mobility indicators.

The coefficients of primary interest are the ones that represent the effects of infection rate. Since the infection rate was log-transformed, the estimated coefficients could be interpreted as the one-hundredth change in the mobility indicators due to a 1% increase in the infection rate. For instance, the unstandardized coefficient for person-miles traveled -0.010 suggested that it would reduce by 0.0001 miles if the infection rate increased by 1%. All the coefficients of infection rate were significant at the 1% level and denoted its negative influence on all the mobility indicators. A similar magnitude of impacts was observed for the percentage of people going out-of-home and person-miles traveled. The lowest impact was observed on the average number of non-work trips.

Structural Model: Human mobility and COVID-19 infection interactions

The conceptual SEM model postulated a bidirectional effect between mobility factor and COVID-19 infection rate. In that, in addition to mobility affecting county-level COVID-19 spread, COVID-19 spread affected county-level human mobility. In the fitted model, we observed a positive effect of human mobility on the infection ($inf \xrightarrow{+ve} mob$) and a negative effect of the infection on human mobility ($mob \xrightarrow{-ve} inf$). That means, mobility influenced infection positively (lower mobility induced lower infection and higher mobility correlated to higher infection). On

Table 7
 Four groups of counties with varying interactions between mobility and infection.

Effect	Group of counties	Explanation
$inf \xrightarrow{+ve} mob$ (low, low) ("pink" in Fig. 9)	<i>Proactively restrictive</i> (cautious, took actions before surge in cases)	These counties had a lower infection as well as lower mobility indicating a proactive restriction on mobility as a precaution. They belonged to mostly East and West coast regions (e.g., northern portion of California, Oregon, Washington, Vermont, Maine, New Mexico) and are mostly metropolitan counties (54% of them).
$inf \xrightarrow{+ve} mob$ (high, high) ("green" in Fig. 9)	<i>Loose enforcement</i> (less reactive to rising cases)	These counties experienced a higher number of infections and at the same time, a higher degree of mobility, suggesting that they might be behind activating mobility restrictions. These were mostly rural counties spanning mostly in the Southeast region. Notably, counties from Georgia, Alabama, and Mississippi were included here and these states did not continue 'stay at home' order after April 30.
$mob \xrightarrow{-ve} inf$ (low, high) ("blue" in Fig. 9)	<i>Comfortably relaxed</i> (no surge, no action)	These counties had a lower infection but with higher mobility. That means, no mobility restriction was in place in those counties (or people perhaps did not comply with them as there were fewer reported cases). These were mostly rural counties located in the U.S. Midwest, West and Southwest regions.
$mob \xrightarrow{-ve} inf$ (high, low) ("orange" in Fig. 9)	<i>Reactively restrictive</i> (respond after surge in cases)	These coastal areas metro counties experienced higher infection and lower mobility. They got severe infection rate by the pandemic and mobility restrictions were realized.

Note: Values for degree of mobility were estimated by the model

the other hand, the effect of infection on mobility was negative: counties having lower infection rates observed a higher degree of human mobility and counties with higher infection rates experienced lower mobility. Table 6 lists the coefficient values of these bidirectional relationships.

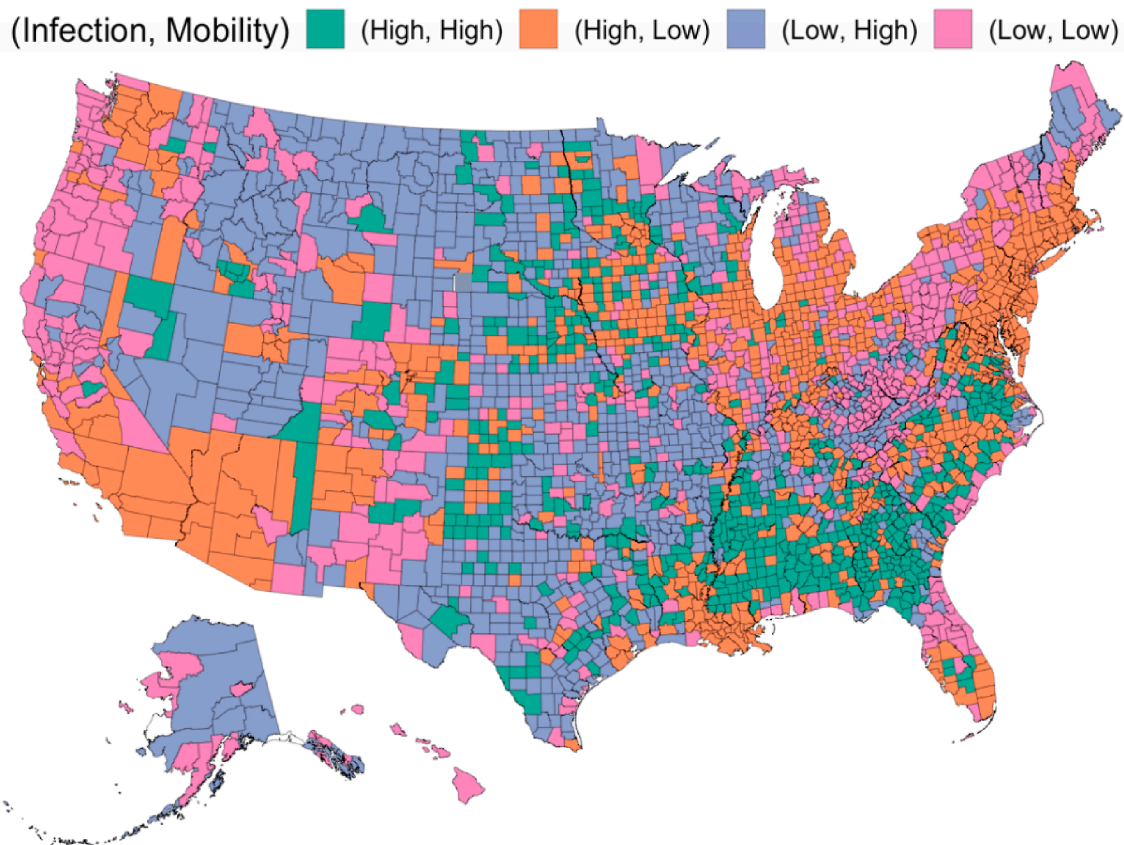
To analyze this bi-directional effect between mobility and infection more closely and to seek validation from the data, we partitioned the U. S. counties into *four groups* based on their infection rates and the degree of human mobility. We tagged a county to be in the "low" category if the value of the corresponding variable (infection or mobility) fell below the median value of that particular variable, otherwise, it was flagged as "high". In that, we obtained four quadrants: (low, low), (high, low), (high, high), and (low, high). We graphically showed this information in Fig. 6, where X-axis denotes the infection rate, Y-axis represents the degree of human mobility and each dot represents one U.S. county. Note that since the degree of human mobility was a latent factor, its values per county were estimated by the model from its observed indicators. The counties were color-coded based on whether they were metro counties (blue color) or non-metro counties (green color). The classification of counties into metro and non-metro areas was adopted from Ingram and Franco (2013).

The estimated positive association between mobility and infection ($inf \xrightarrow{+ve} mob$) was demonstrated by (low, low) and (high, high) quadrants whereas the negative association between infection and mobility ($mob \xrightarrow{-ve} inf$) was manifested by (low, high) and (high, low) quadrants in Fig. 6. We observed that metro counties usually had lower mobility: (low, low) and (low, high) categories, whereas non-metro counties had higher mobility: (high, low) and (high, high) categories. A varying nature of interactions between mobility and infection rate during our study period was observed among these four groups of counties, which is summarized in Table 7. In addition, Fig. 7 marks these four groups of counties on the U.S. map.

Model applications

Validation of the model

The structural model is validated using data from June 2020 to determine how well the model performs in terms of its predictive power when applied to a different dataset. In the month of June, the United States experience a decline in the new COVID-19 cases. As such, some of the areas started reopening in different phases and easing restrictions on activities while others were waiting for the daily new cases to drop further (Curry, 2020; The Economic Tracker, 2020b). Hence, this month represents a scenario that is quite different from the month of May but



Note: Values for degree of mobility were estimated by the model

Fig. 7. Distribution of four groups of U.S. counties in May 2020.

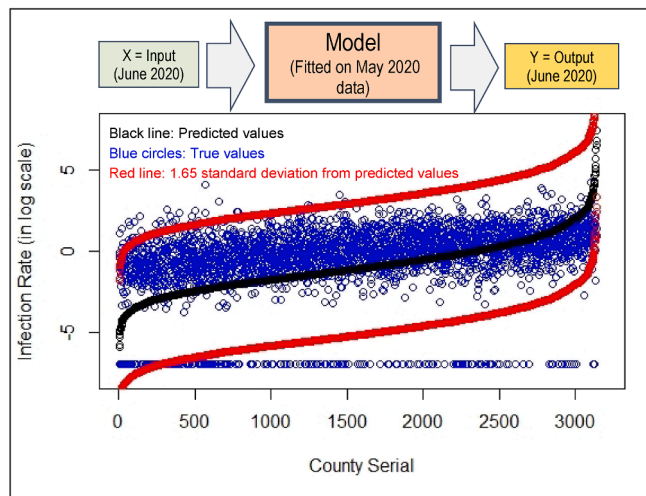


Fig. 8. SEM model validation.

also represents a large variation in the mobility, COVID-19 status, and policy circumstances. All these characteristics make this month better suited for the model validation.

The result is shown in Fig. 8. The predicted values (black line) and true infection rate for the month of June 2020 (blue circles) are plotted against the counties. In addition, two lines (red) are drawn to represent the 95% confidence interval of the predicted values, which are generated by calculating the 1.65 standard deviation of the absolute

prediction errors found from the May 2020 dataset. As apparent, most true data points are within the 95% confidence interval of the predicted data points. The calculated root means squared error (RMSE) between the model predicted and true value is close to the RMSE calculated using the May 2020 – 2.259 vs 2.174. All of these observations indicate that the model predicts satisfactorily within reasonable ranges when applied on a different dataset pertaining to a different timeframe. Hence it can be applied to other datasets for county-level prediction of infection rate or other endogenous variables within a satisfactory margin of error. It is important to note that the predictor representing the infection rate in the neighboring counties is developed using the same data used for the outcome variable. Hence it would be impractical to predict the data which are also used as an input. To avoid this problem, the neighboring county infection rates for June are replaced with the data from May. Also, note that the constant scatter of infection rate at the bottom of the graph is due to the 226 counties representing zero COVID-19 cases during June 2020.

Scenario testing

In this section, the estimated SEM model is applied to the dataset to test different scenarios of mobility restriction and infection rates. A major advantage of SEM over other modeling techniques is that it enables including multiple endogenous variables, which therefore allows the prediction of multiple parameters under the same scenario. Two scenarios are developed considering the practicality and importance of the control and outcome variables. The first scenario involves forecasting the infection rate corresponding to a specified change in selected mobility indicators. The second scenario makes use of the reciprocal relationship between infection rate and human mobility and predicts the

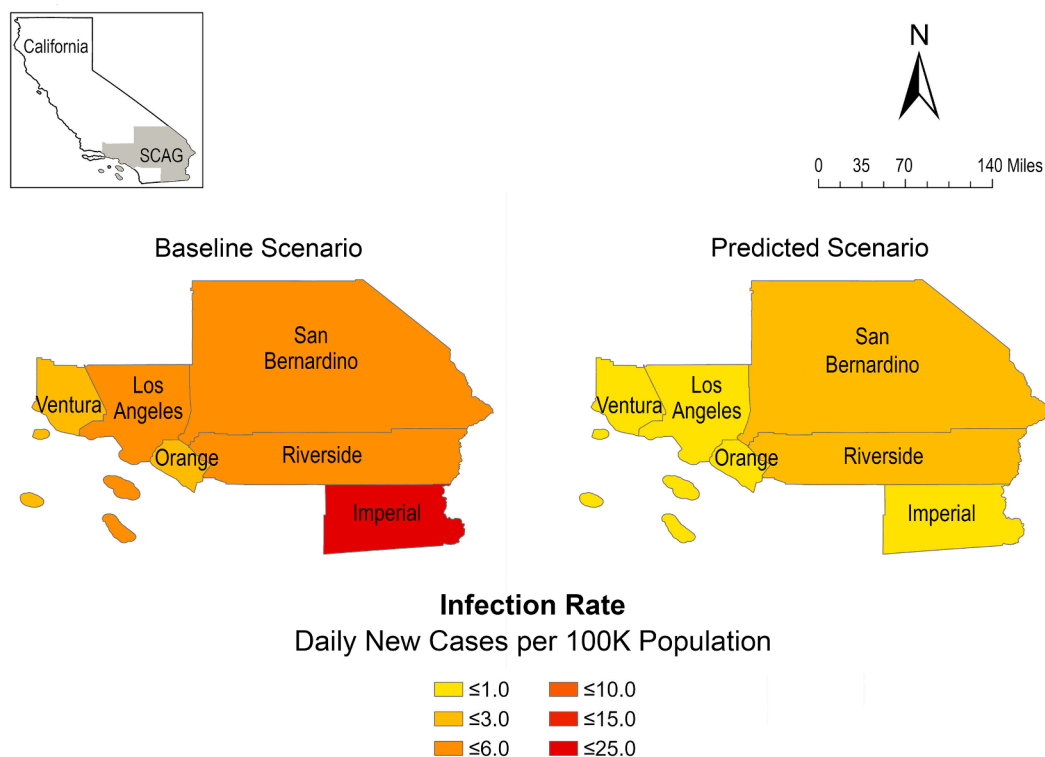


Fig. 9. Baseline and predicted infection rate for Scenario 1 in SCAG region.

mobility indicators for a specified change in infection rate. Both of these scenarios and their results are discussed in detail in the following sections.

Scenario 1: Forecasting COVID-19 infection rate by mobility indicators

This scenario is developed to simulate the circumstances when a set of mobility indicators need to be controlled to achieve a lower infection rate in the future. To conduct this scenario testing, the estimated model is applied on the June 2020 dataset for the Southern California Association of Governments or SCAG region with a 10% reduction in mobility. SCAG is the largest metropolitan planning organization in the United States and serves six counties in Southern California – Imperial, Los Angeles, Orange, Riverside, San Bernardino, and Ventura. The change in mobility reduction is attained by reducing people going out of home, non-work trips and person-miles traveled by 10%. By making these changes in June, the model is expected to provide an estimate of the change in infection in July compared to June as there is always a lag between the effect and detection of COVID-19 cases (Centers for Disease Control and Prevention, 2020b). Due to the same reason discussed in section 6.1, data for the average infection rate of neighboring counties is obtained from the month of May. This is practical because in reality future infection rate is unknown and more importantly this is the data that is predicted by the model.

The result of the forecasting is illustrated in Fig. 9 in which the predicted infection rate in the SCAG region is compared with its baseline value. A significant change in the number of cases is observed for all six counties, where Imperial County has the higher reduction (96.8%) in infection rate. The lowest reduction is observed for San Bernardino County (63.1%), while the average reduction is 76.5%. While this scenario assumes all factors will remain unchanged, this is rarely the case in reality. But it helps policy makers to understand the direct contribution of a change in the mobility indicators on the infection rate. By channeling the effects of the indicators through a latent construct of human mobility, it allows obtaining a combined effect while also excluding the measurement errors from the estimation. The use of the estimated model

in this scenario analysis exemplifies its usefulness in making informed decisions on mobility-oriented policy actions.

Scenario 2: Forecasting level of mobility by COVID-19 infection rate

In this scenario, the same three mobility indicators in the previous scenario are predicted based on an expected change in the infection rate. This prediction utilizes the feedback effect which translates the effects of the predictors to the mobility indicators through the latent construct. The goal is to understand the extent to which each aspect of human mobility would be affected corresponding to a 10% increase in COVID-19 infection rate. The same June 2020 dataset is used in this scenario without replacing the neighboring county infection rate data since it is used only as a predictor. Clay county, Iowa is selected to conduct the scenario analysis.

The chart in Fig. 10 presents the results. The value of each indicator suggests the percentage change from the June 2020 value due to the 10% increase in the infection rate in July 2020. The model suggests that going out of home would reduce by 10.7%, while non-work trips and person-miles traveled are expected to reduce by 15.6% and 25.0%, respectively. This prediction can inform policy makers about the extent of the impact on mobility caused by a possible change in the infection rate. The analysis can also be useful in setting and monitoring mobility status to make an expected reduction in the infection rate. For instance, it draws attention towards person-miles traveled which need to be controlled significantly more than the other indicators to control the infection rate. Moreover, the relative proportion of changes in the mobility indicators will help to formulate appropriate policies and allocate resources for a more equitable outcome.

Conclusions and policy implications

As in the case of other pandemics throughout history, human mobility played an important role in determining the direction and magnitude of spread. On the other hand, the spread of the pandemic made tremendous disruptions to people’s day-to-day schedules and

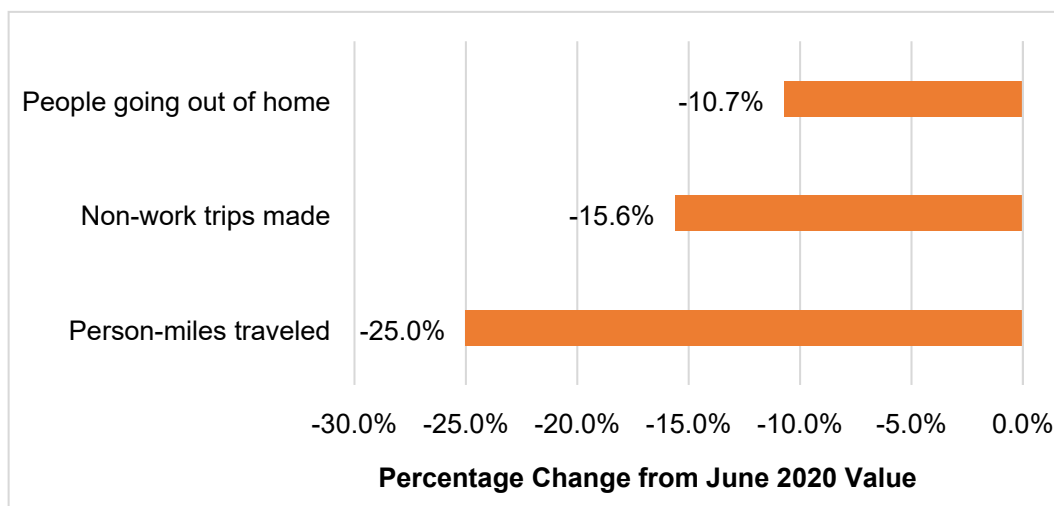


Fig. 10. Predicted changes in mobility indicators for Scenario 2 in Clay County, Iowa.

mobility patterns. This study is an attempt to shed some light on this bi-directional relationship between mobility and COVID-19 spread. In this regard, we developed a Structural Regression (SR) model across U.S. counties by conceptualizing a bi-directional relationship between human mobility and COVID-19 infection along with other influencing factors that can potentially affect both mobility and infection, such as socio-demographic and location characteristics, COVID-19 status, and policies. We considered one month (May 2020) of cross-sectional data during the early phase of the pandemic. The major study findings, their policy implications, and limitations of our study are discussed below.

Major findings

The structural regression model resulted in some notable findings:

- The degree of human mobility had a positive effect on the COVID-19 infection rate in a county. On the other hand, the infection rate negatively influenced people's mobility.
- The mobility factor was largely manifested by the fraction of people going out-of-home and the average distance traveled.
- Counties that had a larger fraction of African-Americans, younger people (18 – 44 years), and the labor force, experienced higher infection rates.
- Compared to non-metropolitan counties, metropolitan counties had higher infection rates in May 2020.
- Counties having highly infected neighboring counties and more external trips experienced a higher infection rate.
- Counties with a longer duration of stay at home order, emergency declaration, and non-essential business closure enacted had lower mobility. On the other hand, counties with a longer duration of public mask mandate had lower infection rates.

Policy implications

The preliminary ideas, analysis, and findings of this study provide important insights on the exposure to COVID-19 infection and equity issues, quantitative measure of two-way relationships between infection and mobility, geographical variations of disease spread, the effectiveness of COVID-19 policies at an aggregate scale. The policy implications of our findings are discussed below:

(a) Exposure to COVID-19 infection and equity issues

The findings of this study provide information on the socio-

demographic characteristics of people who are at high risk of exposure to COVID-19 infection and related equity issues. For example, results suggested that counties with a higher proportion of younger age groups (18 – 24 and 25 – 44) had higher infection rates whereas counties with a higher proportion of older adults (45 – 64 and 65 or above) had lower infection rates. Since younger groups were employed in essential jobs (more in-person work, more human interactions), they were at high risk of infection. Similarly, counties with a higher proportion of African-Americans, females, workers had higher infection rates. Since a higher fraction of African-Americans are employed in low-wage essential jobs, they are at higher risk of infection (Reyes, 2020). Other factors also increase their risk of exposure, such as crowded housing, lower access to healthcare facilities, and reliance on public transportation (Centers for Disease Control and Prevention, 2021a). Policies need to address these equity issues by considering frequent COVID-19 testing facilities and low-cost healthcare options. Policies need to also consider safe and easy access to jobs, transportation facilities, and healthcare systems for all the workers who work in essential jobs.

As observed from our findings, technologically advanced counties (higher internet access) experienced lower mobility and lower infection. It implies that policies need to be taken to reduce the digital divide particularly between rural and urban counties and to increase ICT coverage everywhere, which might contribute to reductions in mobility and consequently reductions in exposure to infection during the pandemic.

(b) The degree of the COVID-19 spread in relation to human mobility

Our developed model will help to quantify the two-way relationships between infection and mobility. There are two related policy implications. First, the model can be utilized to forecast the infection rate of a county as a function of mobility indicators and other variables that can help the county officials to take required mobility-oriented policy actions. For example, if we plan to make changes in mobility indicators, say 10% reductions in the fraction of people going outside home, the model can generate the amount of reductions that we could expect in infection rate in the future for a certain area. Second, the model can also be utilized to understand the extent to which each aspect of human mobility would be affected corresponding to a 10% increase in COVID-19 infection rate. For example, it can estimate the amount of possible reductions in non-work trips or average distance traveled if the infection rate increased, say 5%, of its current value.

(c) Geographical variations in the spread of the pandemic and human mobility

Throughout the pandemic, considerable heterogeneity existed among U.S. counties, both in terms of the spread of infection and the degree of human mobility. Based on the heterogeneity in the spread of the infection and degree of mobility, we divided counties into four categories: proactively restrictive, loose enforcement, comfortably relaxed, reactively restrictive. Counties with lower infection and lower mobility were considered “proactively restrictive”, they took action before the surge in infection cases. These counties belong mostly to the east and west coast regions, particularly in the northern portion of California, Oregon, Washington, Vermont, Maine, and New Mexico states. On the other hand, counties with higher infection rates and higher mobility were termed as “loose enforcement” counties as were less reactive to rising infection rates. Counties from Georgia, Alabama, and Mississippi were mostly included in this group. Counties that had lower infection rates and higher mobility were tagged as “comfortably relaxed”. These were primarily rural counties located in the Midwest, West, and Southwest regions. Counties that responded after the surge in infection rates (high infection and low mobility) were considered as “reactively restrictive”. Coastal metro counties mostly belonged to this group. This information of variations in the COVID-19 spread and human mobility across U.S. geographies will help policy makers in making better plans for resource mobilization and taking effective community mitigation measures for the current and future pandemic situations.

(d) Effectiveness of COVID-19 policies on restricting human mobility and reducing infection

In this study, we investigated how mobility and infection changed in response to various policy enactments, such as emergency declaration, gathering restrictions, stay at home order, non-essential business closure, public mask mandate, and quarantine mandate. Counties that had a longer duration of stay at home orders, emergency declarations, and non-essential closures had lower mobility, whereas counties with a longer duration of public mask mandate experienced lower infection rates. However, we obtained a positive association between stay-at-home order enactment and COVID-19 infection cases, which might indicate that counties that had experienced a high surge of COVID-19 cases (higher infection) at the beginning of the pandemic had to maintain more restrictive policies leading to enacting stay at home earlier (longer duration of stay at home order). In our study window (May 2020), not all policies had been equally effective. For example, quarantine mandate and gathering restrictions did not significantly influence human mobility. This will help policy makers to relatively assess one policy over the other in controlling human mobility, and consequently reducing COVID-19 spread.

Study limitations

There are several limitations in this study, mostly pertaining to data unavailability and modeling technique. First, several other possible variables and interconnections could be added to the framework of the SEM model estimated in this study. But we only considered the most frequently cited variables in the literature related to mobility, socio-demographic, and location characteristics. This partly helped to avoid data availability issues, ensured inclusion of important variables in the model, and enabled us to compare the results with prior studies. In addition, the endogeneity of infection rate can be extended onto other variables beyond only mobility. For example, there can be a direct positive connection from infection rate to several COVID-19 policy-related variables, such as gathering restriction and non-essential business closure. To avoid making the model complex and to focus mostly on the relationship between mobility and infection, we limited the

connection only from the infection to mobility, not towards any other variables. With the availability of relevant data in the future, the authors plan to incorporate additional variables and connections to investigate their effects on the various components of the model.

Second, we conducted a structural analysis using a four-week span of cross-sectional data from the initial Spring outbreak period (first wave) of the pandemic. This ‘static’ analysis for a specific time frame can be extended for the other time periods, for example, Summer 2020 (second wave) and Fall 2020 (third wave), which can provide valuable insights on how the relationships between human mobility and COVID-19 spread evolved throughout the year. In addition to this, time-series analysis can be done to portray the dynamic relationship between mobility and COVID, along with the lagged effect of variables related to mobility, policy measures and COVID-19 status, which is one of our future research interests.

Third, the currently available dataset does not contain information on infected persons who are asymptomatic. As these persons do not show any symptoms, they are less likely to be tested and appear on the record. This could lead to a significant underestimation of COVID-19 cases. Moreover, as asymptomatic infected person can still transmit the disease to other persons, but since there is no record, it can be difficult to identify the true cause of a surge in cases. Some of the earliest spikes in COVID-19 cases were linked to international travel and ‘super-spreader’ events (Gu et al., 2020; Lakdawala and Menachery, 2021). But our model could not incorporate these factors due to the unavailability of the data on the occurrence, time and location of these events. All of the aforementioned discrepancies have the potential to disrupt the relationship presented in the model.

Fourth, data aggregated at the county level may lead to issues with obscuring the variations at the census tract or even at the city level. However, one of the advantages of using a higher-level spatial unit data, that is county-level, is that the distribution of the data gets more accurate with a lower margin of error, particularly for American Community Survey data (Spielman and Folch, 2015). In that, a county-level analysis will help to understand the influence of aggregated behavior on the spread of COVID-19 to take effective policy actions at the county or state level administrative units.

Fifth, this study was conducted in a period when the spread of COVID-19 was in the initial phase without any fully developed preventive health measures. But now after eight months of COVID-19 vaccine administration, some of the relationships in this study might not hold. Moreover, the newest ‘delta’ variant of COVID-19 has shown distinct characteristics both in the way of spread and severity. Thus, a recalibration of the model with such factors and additional sample might be required before proceeding with model application in new and developing scenarios.

CRedit authorship contribution statement

Rezwana Rafiq: Conceptualization, Methodology, Data curation, Formal analysis, Investigation, Software, Visualization, Writing – original draft, Writing – review & editing. **Tanjeeb Ahmed:** Conceptualization, Methodology, Data curation, Formal analysis, Investigation, Software, Visualization, Writing – original draft, Writing – review & editing. **Md Yusuf Sarwar Uddin:** Conceptualization, Methodology, Data curation, Formal analysis, Investigation, Software, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

A preliminary version of this paper was submitted to the C3.ai COVID-19 Grand Challenge and won third place in the competition (<https://c3.ai/c3-ai-covid-19-grand-challenge/>). Another version of this paper is accepted for presentation at the 2022 TRB Annual Meeting (presentation number TRBAM-22-00339). The authors also acknowledge the anonymous reviewers for their constructive comments and suggestions that helped them substantially improve the paper.

References

- Acock, A.C., 2013. *Discovering structural equation modeling using Stata*. Stata Press books.
- Badr, H.S., Du, H., Marshall, M., Dong, E., Squire, M.M., Gardner, L.M., 2020. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *Lancet Infect. Dis.* 20 (11), 1247–1254. [https://doi.org/10.1016/S1473-3099\(20\)30553-3](https://doi.org/10.1016/S1473-3099(20)30553-3).
- Beck, M.J., Hensher, D.A., 2020. Insights into the impact of COVID-19 on household travel and activities in Australia-The early days under restrictions. *Transport policy* 96, 76–93.
- Bentler, P.M., 1995. *EQA structural equations program manual*. Multivariate Software, Encino, CA.
- Bhouri, M.A., Costabal, F.S., Wang, H., Linka, K., Peirlinck, M., Kuhl, E., Perdikaris, P., 2021. COVID-19 dynamics across the US: A deep learning study of human mobility and social behavior. *Comput. Methods Appl. Mech. Eng.* 382, 113891. <https://doi.org/10.1016/j.cma.2021.113891>.
- Boehmer, T.K., DeVies, J., Caruso, E., van Santen, K.L., Tang, S., Black, C.L., Hartnett, K. P., Kite-Powell, A., Dietz, S., Lozier, M., Gundlapalli, A.V., 2020. Changing age distribution of the COVID-19 pandemic—United States, May–August 2020. *Morbidity Mortality Week Rep* 69 (39), 1404–1409.
- Bryant, P., Elofsson, A., 2020. Estimating the impact of mobility patterns on COVID-19 infection rates in 11 European countries. *PeerJ* 8, e9879. <https://doi.org/10.7717/peerj.9879>.
- Brynjolfsson, E., Horton, J.J., Ozimek, A., Rock, D., Sharma, G., TuYe, H.Y., 2020. COVID-19 and remote work: An early look at US data (No. w27344). National Bureau of Economic Research.
- Byrne, B.M., 2001. *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming* Lawrence Erlbaum Associates Inc. Mahwah, New Jersey.
- Byrne, B.M., 1998. *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Lawrence Erlbaum Associates Publishers, Mahwah, N.J., p. 1998
- Byrne, B.M., 2006. *Structural equation modeling with EQS: Basic concepts, applications, and programming*, 2nd ed. Lawrence Erlbaum Associates, Mahwah, NJ.
- Cao, X., Mokhtarian, P.L., Handy, S.L., 2007. Do changes in neighborhood characteristics lead to changes in travel behavior? A structural equations modeling approach. *Transportation* 34 (5), 535–556.
- Carozzi, F., Provenzano, S., Roth, S., 2020. Urban density and COVID-19. IZA Inst, Labor Econ.
- Carteni, A., Di Francesco, L., Martino, M., 2020. How mobility habits influenced the spread of the COVID-19 pandemic: Results from the Italian case study. *Sci. Total Environ.* 741, 140489. <https://doi.org/10.1016/j.scitotenv.2020.140489>.
- Centers for Disease Control and Prevention (CDC), 2020a. COVID-19 and Animals [WWW Document]. URL <https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/animals.html> (accessed 12.15.20).
- Centers for Disease Control and Prevention (CDC), 2020b. Options to Reduce Quarantine for Contacts of Persons with SARS-CoV-2 Infection Using Symptom Monitoring and Diagnostic Testing [WWW Document]. URL <https://www.cdc.gov/coronavirus/2019-ncov/more/scientific-brief-options-to-reduce-quarantine.html> (accessed 12.13.20).
- Centers for Disease Control and Prevention (CDC), 2020c. COVID-19 Stats: COVID-19 Incidence, by Urban-Rural Classification — United States, January 22–October 31, 2020. *MMWR Morb Mortal Wkly Rep* 2020 (69), 1753. <https://doi.org/10.15585/mmwr.mm6946a6>.
- Centers for Disease Control and Prevention (CDC), 2021a. Health Equity Considerations and Racial and Ethnic Minority Groups. Available at <https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/race-ethnicity.html>. (accessed 10.30.2021).
- Centers for Disease Control and Prevention (CDC), 2021b. Demographic Trends of COVID-19 cases and deaths in the US reported to CDC. Available at https://covid.cdc.gov/covid-data-tracker/?utm_source=morning_brew#demographics (accessed 10.30.2021).
- Chang, M.-C., Kahn, R., Li, Y.-A., Lee, C.-S., Buckee, C.O., Chang, H.-H., 2021. Variation in human mobility and its impact on the risk of future COVID-19 outbreaks in Taiwan. *BMC Public Health* 21 (1). <https://doi.org/10.1186/s12889-021-10260-7>.
- Chinazzi, M., Davis, J.T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Pastore y Piontti, A., Mu, K., Rossi, L., Sun, K., Viboud, C., Xiong, X., Yu, H., Halloran, M.E., Longini, I.M., Vespignani, A., 2020. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* 368 (6489), 395–400.
- Chu, D.K., Akl, E.A., Duda, S., Solo, K., Yaacoub, S., Schünemann, H.J., et al., 2020. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *The Lancet* 395 (10242), 1973–1987.
- Cuadros, D.F., Branscum, A.J., Mukandavire, Z., Miller, F.D., MacKinnon, N., 2021. Dynamics of the COVID-19 epidemic in urban and rural areas in the United States. *Ann. Epidemiol.* 59, 16–20.
- Curry, A., 2020. Which U.S. States Are Reopening? The Complete List of Open and Closed States. URL <https://www.zenefits.com/workest/which-us-states-are-reopenin-g-the-complete-list-of-open-and-closed-states/> (accessed 11.16.20).
- Czeisler, M.E., Tynan, M.A., Howard, M.E., Honeycutt, S., Fulmer, E.B., Kidder, D.P., Robbins, R., Barger, L.K., Facer-Childs, E.R., Baldwin, G., Rajaratnam, S.M.W., Czeisler, C.A., 2020. Public attitudes, behaviors, and beliefs related to COVID-19, stay-at-home orders, nonessential business closures, and public health guidance—United States, New York City, and Los Angeles, May 5–12, 2020. *Morbidity Mortal. Weekly Report* 69 (24), 751–758.
- Dey, M., Loewenstein, M.A., Piccone, D.S., Polivka, A.E., 2020. Demographics, earnings, and family characteristics of workers in sectors initially affected by COVID-19 shutdowns. *Monthly Labor Review* 1–29.
- Dong, E., Du, H., Gardner, L., 2020. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* 20 (5), 533–534. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1).
- Engle, S., Stromme, J., Zhou, A., 2020. Staying at Home: Mobility Effects of COVID-19. SSRN Electron. J. <https://doi.org/10.2139/ssrn.3565703>.
- ESRI, 2020. Spatial Join (Analysis) [WWW Document]. accessed 12.15.20. <https://pro.arcgis.com/en/pro-app/latest/tool-reference/analysis/spatial-join.htm>.
- Federal Aviation Administration (2021). Airport data and contact information. https://www.faa.gov/airports/airport_safety/airportdata_5010/ (accessed on 10.1.2021).
- Fullman, N., Bang-Jensen, B., Reinke, G., Magistro, B., Castellano, R., Erickson, M., Amano, K., Wilkerson, J., Adolph, C., 2020. State-level social distancing policies in response to COVID-19 in the US. Version 1.100. URL <http://www.covid19statepolicy.org> (accessed 10.29.20).
- Gatto, M., Bertuzzo, E., Mari, L., Miccoli, S., Carraro, L., Casagrandi, R., Rinaldo, A., 2020. Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proc. Natl. Acad. Sci. U. S. A.* 117 (19), 10484–10491. <https://doi.org/10.1073/pnas.2004978117>.
- Golob, T.F., 2003. Structural equation modeling for travel behavior research. *Transp. Res. Part B Methodol.* 37 (1), 1–25. [https://doi.org/10.1016/S0191-2615\(01\)00046-7](https://doi.org/10.1016/S0191-2615(01)00046-7).
- Gu, W., Deng, X., Reyes, K., Hsu, E., Wang, C., Sotomayor-Gonzalez, A., Federman, S., Bushnell, B., Miller, S., Chiu, C.Y., 2020. Associations of Early COVID-19 Cases in San Francisco With Domestic and International Travel. *Clin. Infect. Dis.* 71, 2976–2980. <https://doi.org/10.1093/cid/ciaa599>.
- Hamidi, S., Ewing, R., Sabouri, S., 2020a. Longitudinal analyses of the relationship between development density and the COVID-19 morbidity and mortality rates: Early evidence from 1,165 metropolitan counties in the United States. *Heal. Place* 64, 102378. <https://doi.org/10.1016/j.healthplace.2020.102378>.
- Hamidi, S., Sabouri, S., Ewing, R., 2020b. Does Density Aggravate the COVID-19 Pandemic? *J. Am. Plan. Assoc.* 86 (4), 495–509. <https://doi.org/10.1080/01944363.2020.1777891>.
- Hoyle, R.H., 1995. The structural equation modeling approach: Basic concepts and fundamental issues. In: Hoyle, R.H. (Ed.), *Structural equation modeling: Concepts, issues, and applications*. SAGE Publications, London, pp. 1–15.
- Hu, L.-T., Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Eq. Model. Multidiscipl. J.* 6 (1), 1–55.
- Hu, S., Xiong, C., Yang, M., Younes, H., Luo, W., Zhang, L., 2021. A big-data driven approach to analyzing and modeling human mobility trend under non-pharmaceutical interventions during COVID-19 pandemic. *Transport. Res. Part C: Emerg. Technol.* 124, 102955.
- IBM infographics., 2020. Big data and analytics hub. URL <https://www.ibmbigdatahub.com/infographic/extracting-business-value-4-vs-big-data> (accessed 12.14.20).
- Ingram, D.D., Franco, S.J., 2013. NCHS urban-rural classification scheme for counties. *Vital Heal. Stat. Ser. 2 Data Eval. Methods Res* 166.
- International Monetary Fund, 2021. Global Prospects and Policies, in: *World Economic Outlook, April 2021: Managing Divergent Recoveries*. International Monetary Fund, Washington, DC. 10.5089/9781513575025.081.
- Johns Hopkins, 2020. New Cases of COVID-19 in World Countries [WWW Document]. accessed 11.15.20. <https://coronavirus.jhu.edu/data/new-cases>.
- Johns Hopkins University & Medicine, 2020. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) [WWW Document]. URL <https://coronavirus.jhu.edu/map.html> (accessed 12.3.20).
- Kadi, N., Khelifaoui, M., 2020. Population density, a factor in the spread of COVID-19 in Algeria: statistic study. *Bull. Natl. Res. Cent.* 44, 138. <https://doi.org/10.1186/s42269-020-00393-x>.
- Kaplan, D., 2008. *Structural equation modeling: Foundations and extension*. (Vol. 10). SAGE Publications, Inc.
- Kaufman, B.G., Whitaker, R., Mahendratnam, N., Smith, V.A., McClellan, M.B., 2020. Comparing associations of state reopening strategies with COVID-19 burden. *J. Gener. Int. Med.* 35 (12), 3627–3634.
- Kemp, E., Price, G.N., Fuller, N.R., Kemp, E.F., 2020. African Americans and COVID-19: beliefs, behaviors and vulnerability to infection. *Int. J. Healthcare Manage.* 13 (4), 303–311.
- Kenny, D.A., 2020. Measuring model fit. URL <http://davidakenny.net/cm/fit.htm> (accessed on 11.1.21).

- Kim, J., Kwan, M.-P., 2021. The impact of the COVID-19 pandemic on people's mobility: A longitudinal study of the U.S. from March to September of 2020. *J. Transp. Geogr.* 93, 103039. <https://doi.org/10.1016/j.jtrangeo.2021.103039>.
- Kline, R.B., 2016. Principles and practices of structural equation modelling, 4th ed. The Guilford Press, New York.
- Kraemer, M.U.G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D.M., du Plessis, L., Faria, N.R., Li, R., Hanage, W.P., Brownstein, J.S., Layman, M., Vespignani, A., Tian, H., Dye, C., Pybus, O.G., Scarpino, S.V., 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 368 (6490), 493–497.
- Kullar, R., Marcelin, J.R., Swartz, T.H., Piggott, D.A., Macias Gil, R., Mathew, T.A., Tan, T., 2020. Racial disparity of coronavirus disease 2019 in African American communities. *J. Infect. Diseases* 222 (6), 890–893.
- Kwan, J.L.Y., Chan, W., 2011. Comparing standardized coefficients in structural equation modeling: A model reparameterization approach. *Behav. Res. Methods* 43 (3), 730–745. <https://doi.org/10.3758/s13428-011-0088-6>.
- Lakdawala, S.S., Menachery, V.D., 2021. Catch Me if You Can: Superspreading of COVID-19. *Trends Microbiol.* 29 (10), 919–929. <https://doi.org/10.1016/j.tim.2021.05.002>.
- Li, H., Liu, S.-M., Yu, X.-H., Tang, S.-L., Tang, C.-K., 2020a. Coronavirus disease 2019 (COVID-19): current status and future perspectives. *Int. J. Antimicrob. Agents* 55 (5), 105951. <https://doi.org/10.1016/j.ijantimicag.2020.105951>.
- Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., Shaman, J., 2020b. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* (80-. 368 (6490), 489–493.
- Lu, X., Pas, E.L., 1999. Socio-demographics, activity participation and travel behavior. *Transp. Res. Part A Policy Pract.* 33 (1), 1–18.
- Markel, H., 2007. Contemplating Pandemics: The Role of Historical Inquiry in Developing Pandemic-Mitigation Strategies for The Twenty-First Century, in: Lemon, S.M., Hamburg, M.A., Sparling, P.F., Choffnes, E.R., Mack, A. (Eds.), *Ethical and Legal Considerations in Mitigating Pandemic Disease: Workshop Summary*. National Academies Press (US), Washington, DC. 10.17226/11917.
- Maryland Transportation Institute, 2020. University of Maryland COVID-19 Impact Analysis Platform. URL <https://data.covid.umd.edu> (accessed 10.29.20).
- Moreland, A., Herlihy, C., Tynan, M.A., Sunshine, G., McCord, R.F., Hilton, C., et al., 2020. Timing of state and territorial COVID-19 stay-at-home orders and changes in population movement—United States, March 1–May 31, 2020. *Morbidity and Mortality Weekly Report* 69 (35), 1198–1203.
- Nagata, J.M., 2020. Supporting young adults to rise to the challenge of COVID-19. *J. Adolesc. Health* 67 (2), 297–298.
- Oster, A.M., Caruso, E., DeVies, J., Hartnett, K.P., Boehmer, T.K., 2020. Transmission Dynamics by Age Group in COVID-19 Hotspot Counties — United States, April–September 2020. *MMWR Morb. Mortal. Wkly. Rep.* 69 (41), 1494–1496.
- Park, Y.J., Choe, Y.J., Park, O.k., Park, S.Y., Kim, Y.-M., Kim, J., Kweon, S., Woo, Y., Gwack, J., Kim, S.S., Lee, J., Hyun, J., Ryu, B., Jang, Y.S., Kim, H., Shin, S.H., Yi, S., Lee, S., Kim, H.K., Lee, H., Jin, Y., Park, E., Choi, S.W., Kim, M., Song, J., Choi, S.W., Kim, D., Jeon, B.-H., Yoo, H., Jeong, E.K., 2020. Contact Tracing during Coronavirus Disease Outbreak, South Korea, 2020. *Emerg. Infect. Dis.* 26 (10), 2465–2468.
- Rafiq, R., McNally, M.G., 2022. A Comparison of Time-use for Telecommuters, Potential Telecommuters, and Commuters during the COVID-19 Pandemic. in the 101th Transportation Research Board (TRB) Annual Meeting, Washington D.C., Jan 9–13, 2022.
- Rafiq, R., McNally, M.G., Uddin, Y.S., Ahmed, T., 2022. Flattening the Curve by Working from Home: Changes in Activity-Travel Behavior during the COVID-19 Pandemic. in the 101th Transportation Research Board (TRB) Annual Meeting, Washington D.C., Jan 9–13, 2022.
- Reyes, M.V., 2020. The disproportional impact of COVID-19 on African Americans. *Health Human Rights* 22 (2), 299.
- Rho, H.J., Brown, H., Fremstad, S., 2020. A basic demographic profile of workers in frontline industries. In: *Center for economic and policy research*, p. 7.
- Schumacker, R.E., Lomax, R.G., 2004. A beginner's guide to structural equation modeling, 2nd ed. Lawrence Erlbaum Associates, Mahwah, NJ.
- Spielman, S.E., Folch, D.C., 2015. Reducing Uncertainty in the American Community Survey through Data-Driven Regionalization. *PLoS One* 10 (2), e0115626.
- Sy, K.T., White, L., Nichols, B., 2020. Population density and basic reproductive number of COVID-19 across United States counties. *MedRxiv*. <https://doi.org/10.1101/2020.06.12.20130021>.
- The Economic Tracker, 2020a. Daily Reported COVID-19 Cases, Deaths and Tests per 100,000 People [WWW Document]. URL <https://tracktherecovery.org/> (accessed 12.3.20).
- The Economic Tracker, 2020b. Percent Change in Number of Small Businesses Open [WWW Document]. URL <https://tracktherecovery.org/> (accessed 11.16.20).
- Tian, H., Liu, Y., Li, Y., Wu, C.-H., Chen, B., Kraemer, M.U.G., Li, B., Cai, J., Xu, B.o., Yang, Q., Wang, B., Yang, P., Cui, Y., Song, Y., Zheng, P., Wang, Q., Bjornstad, O.N., Yang, R., Grenfell, B.T., Pybus, O.G., Dye, C., 2020. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* 368 (6491), 638–642.
- Van Acker, V., Witlox, F., 2011. Commuting trips within tours: how is commuting related to land use? *Transportation* 38 (3), 465–486.
- Van Acker, V., Mokhtarian, P.L., Witlox, F., 2014. Car availability explained by the structural relationships between lifestyles, residential location, and underlying residential and travel attitudes. *Transport Policy* 35, 88–99.
- U.S. Census Bureau, 2018. 2014–2018 American Community Survey 5-year Estimates. URL <https://www.census.gov/programs-surveys/acs/news/data-releases/2018/release.html> (accessed 11.23.20).
- U.S. Census Bureau, 2019. TIGER/Line Geodatabases. URL <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html> (accessed 10.31.20).
- Wikipedia, 2021. Compartmental models in epidemiology. URL https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology (accessed 10.15.21).
- Wright, J.E., Merritt, C.C., 2020. Social equity and COVID-19: The case of African Americans. *Public Administr. Rev.* 80 (5), 820–826.
- Yoo, S., Managi, S., 2020. Global mortality benefits of COVID-19 action. *Technol. Forecast. Soc. Change* 160, 120231. <https://doi.org/10.1016/j.techfore.2020.120231>.
- Zhang, L., 2018. *Metacognitive and Cognitive Strategy Use in Reading Comprehension: A Structural Equation Modelling Approach*. Springer Nature Singapore Pte Ltd. 10.1007/978-981-10-6325-1.