Geospatial Health 2022; volume 17(s1):1056

Cascading effects of COVID-19 on population mobility and air quality: An exploration including place characteristics using geovisualization

Samuel F. Atkinson,[1,2] Abhishek K. Kala,[2] Chetan Tiwari[3,4]

[1]Department of Biological Sciences, University of North Texas, Denton, TX; [2]Advanced Environmental Research Institute, University of North Texas, Denton, TX; [3]Department of Geosciences & Computer Sciences, Georgia State University, Atlanta, GA; [4]Center for Disaster Informatics and Computational Epidemiology, Georgia State University, Atlanta, GA, USA

Abstract

This study hypothesizes that public health responses to coronavirus disease 2019 (COVID-19), including a mandated restriction of activity (commonly called a 'lockdown') resulted in reduced transportation activities and changes in air quality in Texas, USA. This presented a natural experiment where population mobility and air quality before and after the lockdown could be compared. Changes in mobility were measured by SafeGraph mobility data (from opt-in smart phone applications that transmit location data) and air quality changes were based on NO_2 concentrations measured by the European Space Agency's Sentinel-5 Precursor satellite (from the TROPOspheric Monitoring Instrument). The changes in population mobility and NO_2 concentration between mid-March 2020 (lockdown initiated) and the end of 2020, as compared to the same time window in 2019, were the basis of exploring the lockdown hypothesis. Additionally, numerous socio-economic (place based) indicators were hypothesized to follow public health vulnerability assumptions based on COVID-19 incidence patterns. This hypothesis was subjected to geovisualization techniques in order to find potential patterns and insights into the complex combinations of these place-based data. Our results suggest that simultaneously visualizing COVID-19, mobility, air quality and socio-economic data yields insights in underlying spatial processes related to public health policy decisions. The hypothesis that the lockdown resulted in reduced mobility and NO_2 concentrations was found partially correct - this trend was observed in highly urbanized areas, but not in less populated areas. Data related public health vulnerability assumptions (*e.g.* a region's age, poverty, education, *etc.*) were agreed with in part, but disagreed with in part.

Correspondence: Samuel F. Atkinson, Department of Biological Sciences, University of North Texas, 1155 Union Circle #305220, Denton, TX, USA.
Tel.: +1.940.565.2694 - Fax: +1.940.565.4297.
E-mail: sam.atkinson@unt.edu

Key words: SafeGraph mobility data; Sentinel-5 tropospheric instrument; self-organizing maps; parallel coordinate plots; public health social vulnerability.

Acknowledgements: the authors would like to acknowledge the pain and suffering of victims of COVID-19 around the world. The University of North Texas provided funding from faculty research funding resources.

Received for publication: 19 November 2021.
Revision received: 14 December 2021.
Accepted for publication: 22 January 2022.

Geospatial Health 2022; 17(s1):1056
doi:10.4081/gh.2022.1056

Introduction

The movement of people across space influences how humans interact among themselves and with the environment. To limit the spread of coronavirus disease 2019 (COVID-19), various governments across the world enacted policy actions to restrict the movement of people across space. The sudden and drastic reduction in population movement because of such policies presents a unique opportunity to better understand how human interaction limits the spread of diseases (Chan *et al.*, 2020; Kraemer *et al.*, 2020; Li *et al.*, 2020) as well as our collective impacts on the environment. In the United States of America, social distancing measures were implemented by individual states with the goal of limiting the spread of the pandemic. In general, closure or non-physical interaction options (*e.g.*, product delivery only) were implemented for schools, restaurants, and public places of gathering. Businesses, workers, and types of activities that were deemed essential during the pandemic could either continue operating under strict protection measures (*e.g.*, personal protective equipment, masks) or switch to online operations. Non-essential businesses requiring physical presence and interaction (*e.g.*, hair salons, bars, gyms) were required to close completely.

While population restrictions were generally widespread in the United States, the level of compliance with these limitations tended to vary by geography. Political perspectives, socio-economic constraints, attitudes towards the pandemic, and various other factors led to varying levels of compliance across the United States. Many states announced some level of social distancing orders starting in mid-March 2020, often including a mandatory quarantine for people diagnosed with or showing likely COVID-19 symptoms. By the beginning of April, almost all states had a mandatory shelter-in-place or lockdown orders. In the USA, as a result of social distancing, states started to experience a dramatic decrease in personal transportation and mobility in general (Gao *et al.*, 2020). Personal vehicle transportation decreased by approximately 46% on average nationwide, while freight movement only decreased by approximately 13% (Pishue, 2020). Air traffic decreased significantly as well (Slotnick, 2020). In addition, the dynamics of population movement and interaction in urban areas tends to be different when compared to rural areas.

Texas covers a large area (approximately 695,000 km^2, an area slightly larger than France) and is the largest state in the contiguous United States. It contains a large population (over 29 million people - would be in the top 50 counties in terms of population), and on the date of this writing contained the second highest total number of COVID-19 cases in the United States. These statistics prompted our exploration of the hypothesis covering assumptions of place-based vulnerability to infectious disease. This exploration focused on the areas in Texas that clustered together because of similarities in total COVID-19 incidence and deaths, along with similar changes in population mobility and NO_2 concentration after the lockdown (as compared to the same time window in 2019), and ultimately several socio-economic factors that are used by the CDC as indicators of potential health vulnerability.

In Texas, population density as well as demographic and socio-economic indicators can be different as one moves from the more rural western half of the state towards the more urban eastern half. The areas along the US-Mexico border in Texas also present different demographic, socio-economic, and cultural dynamics when compared to other parts of Texas. In this study, geovisualization techniques were used to integrate and analyse complex and disparate datasets that capture different aspects of population movement in Texas along with measures of COVID-19 burdens and air quality. This project explored the relationships across Texas between public health events, daily COVID-19 incidence, mobility of the population, and atmospheric NO_2 concentrations during the 2020 calendar year, as compared to the 2019 calendar year. The Texas ‘lockdown’ and subsequent reopening mandates in response to COVID-19 presented a unique opportunity to explore Texas mobility trends before and after the lockdown, and whether air quality responded as would be expected. Major ambient air pollution sources include power generation, industry, traffic, and residential energy use (Lelieveld *et al.*, 2015; Crippa *et al.*, 2018), leading to a reasonably foreseeable presumption that large-scale reductions in population mobility should result in improved air quality. This presumed cascading of COVID-19 incidence leading to policies restricting population interactions, leading to reduced mobility, leading to improved air quality prompted additional analyses involved exploring the place characteristics (*i.e.* socio-economic data related to health vulnerability) using geovisualization tools.

Materials and methods

Overview of study methodology

Figure 1 shows the four components used to better understand how policy actions in Texas influenced complex interactions between COVID-19 outcomes, demographic and socio-economic characteristics, levels of population interaction, and changes in air quality due to sudden reductions in mobility. Geovisualization was used to combine these four sources of information that were compiled from multiple and disparate datasets obtained from a variety of public and proprietary data sources.

COVID-19 incidence

One of the first decisions made in this study involved whether to use Texas Department of State Department of Health COVID-19 incidence data (Texas Department of State Health Services, 2021), or Johns Hopkins CORONAVIRUS Resource Center incidence data (Johns Hopkins Coronavirus Resource Center, 2021). While the Texas data are focused solely on Texas and therefore presumably a better representation, the Johns Hopkins data are nationwide and therefore available for each state. The data sets are quite similar (Figure 2), but daily numbers of new cases were slightly different due to when data were reported and how backlogged data were handled; each added backlogged data on different days (primary difference).

Texas COVID-19 incidence numbers on each day in 2020 were ranked from highest to lowest in both data sets. For example, day 364 (December 29th) represented the highest daily COVID-19 count in the Texas data, and the second highest daily COVID-19 count in the Johns Hopkins data. Day 183 (the mid-point of 2020, July 1st), Texas data showed the 68th highest count while Johns Hopkins data showed the 66th highest count. The coefficient of variation between the two ranked sets for all days was 0.93

Figure 1. Overview of exploring the cascading effects of COVID-19 in Texas, USA.

(P=0.09), (Figure 3). Therefore, in order to allow our approach to be replicated by researchers in other states, the Johns Hopkins data were selected.

Population mobility

The hypothesis that the lockdown cascaded into reduced population-wide mobility activities because millions of people were sheltering at home was explored using SafeGraph's Patterns dataset [from SafeGraph, a data company that aggregates anonymized location data from numerous applications in order to provide insights about physical places, via the SafeGraph Community. To enhance privacy, SafeGraph excludes census block group information if fewer than two devices visited an establishment in a month from a given census block group.]. This dataset

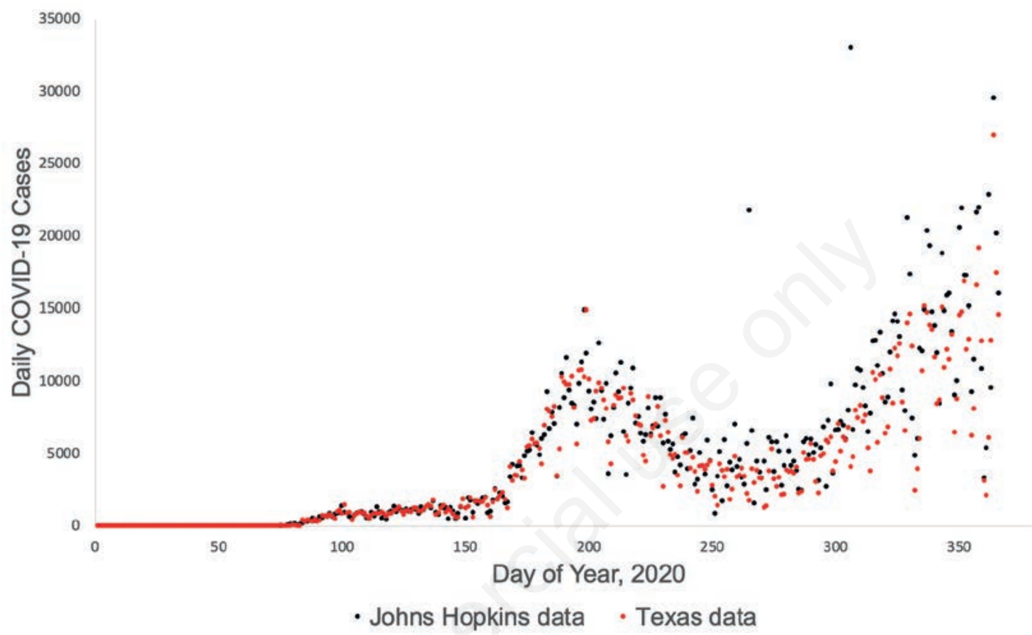

Figure 2. Daily COVID-19 new cases as reported by the Johns Hopkins CORONAVIRUS Resource Center (black points), and the Texas Department of State Health Services (red points).

Figure 3. Relationship between ranked daily Texas Department of State Health Services and Johns Hopkins CORONAVIRUS Resource Center COVID-19 data (R^2=0.9333, P=0.092).

contains approximately 4.5 million Points of Interest (POI) in the United States using a panel of mobile devices (Safegraph, 2021). The Patterns dataset was used to analyse and compare foot traffic for 450,936 POI locations in Texas across the baseline year 2019 and the pandemic year 2020. Information on daily visitor counts were extracted from the dataset for these POIs to produce approximately 130 million data points for 2019 and approximately 128 million data points for 2020. Information on movement patterns including frequency of visits, originating block group of visitors, and dwell times provide insights into population movement patterns. To evaluate these data for potential biases when comparing movement patterns across multiple years, we evaluated the data for consistency in the number of POIs captured by year and geography. The number of POIs captured during every month in 2019 and 2020 tended to be somewhat consistent at the county level. However, the number of POIs captured by SafeGraph data varied across different regions in Texas; as expected, rural areas had fewer POIs compared to urban areas. Further, urban areas captured a much larger variety of POIs compared to rural areas where some locations were very popular while others only had a handful of visits. For these reasons, exploring the hypothesis that the lockdown influenced population mobility patterns were based on using the 'most popular' POI in every census block group for each day of the two-year time-period from 2019 to 2020. The resulting dataset contained approximately 12 million records where each row represents the most popular POI for a specific date and census block group in Texas. In order to compare population mobility measure to satellite-derived NO_2 estimates, the POI dataset was compiled at the county-level (n=255) for 14-day time periods starting January 01, 2019 and ending on December 31, 2020 (n=24). The final dataset contained 12,240 records. Maps showing differences in mobility were produced by aggregating data from the daily, most popular POI table to the census block group level. The resulting table contains 18,638 records. Alteryx software was used to prepare raw data obtained from SafeGraph. The processed data files were then loaded into a PostgreSQL database for processing. For mapping purposes and clarity, the data were mapped based on mobility during 30 days prior to compared to the 30 days post lockdown, but for analytical purposes of the lockdown hypothesis, mobility was compared between a time window beginning of March and the end of December 2020, and the same time window in 2019.

Air quality

The hypothesis that the lockdown cascaded into reduced population mobility which cascaded into reduced transportation activity, cascading into less fossil fuel combustion and improved air quality, was explored by using European Space Agency Sentinel-5 satellite data. NO_2 was the focus of this research because it is one of the pollutants regulated at the federal level by the U.S. EPA as a priority pollutant under National Ambient Air Quality Standards. It can also be measured from satellite-borne remote sensing technology.

The spatial distribution of tropospheric NO_2 (from surface up to ~10 km) can be measured using the recently launched (13 October 2017) European Space Agency's Sentinel-5 Precursor satellite (S5P). The satellite has an onboard a high precision optical payload that is referred to as TROPOspheric Monitoring Instrument (TROPOMI) which provides a (near-) global coverage of different air pollutants including NO_2. The satellite operates in a sun-synchronous orbit at 824 km, has an orbital cycle of 16 days, and has a spatial resolution of 7 km^2. The Tropospheric Vertical Column Density (VCD) data of NO_2 is measured from TROPOMI which serves as an accurate proxy for ground level concentrations (Lamsal *et al.*, 2015). The VCD is defined as the number of molecules of a certain atmospheric gas between the on-board sensor of the satellite platform and the Earth's surface per unit area. For NO_2, VCD measurements have been successfully used to estimate trends and variations in atmospheric concentration (Castellanos *et al.*, 2012; Curier *et al.*, 2014), infer surface emissions (Ghude *et al.*, 2013; Streets *et al.*, 2013) and monitor emission changes at a given location (Kim *et al.*, 2006; Wang *et al.*, 2012).

The lockdown hypothesis was explored in this study using NO_2 concentrations across Texas for the baseline year 2019 compared to the pandemic year 2020, but analyses were focused on the same time window used for mobility analysis, the beginning of March through the end of December. The publicly available NO_2 data was retrieved and analysed using the Google Earth Engine API (Gorelick *et al.*, 2017), a cloud based and an open-source geospatial analysis platform. A JavaScript was composed to retrieve the global NO_2 data, clip it based on the Texas state boundary, calculate mean pollution levels, and produce a map for visualizing average NO_2 for the two years.

Geovisualization

The lockdown hypothesis of cascading effects of changes in mobility and NO_2 concentrations, and potential implications on public health vulnerability indicators can be explored using spatially explicit approaches. Geovisualization was used to identify patterns in the interactions of COVID-19 incidence, population mobility changes, air quality changes, and public health vulnerability. Geovisualization can be defined as the use of tools and techniques to support the analysis of large amounts of geospatial data through the use of interactive visualization. It is essentially a data mining process and is commonly used to identify the spatial context and associated relationships between a pre-defined set of potential explanatory variables. Geovisualization techniques may provide valuable insights for identifying variables and associated processes that contribute to variations in disease risk across space and time. Geovisualization was used in this study following procedures outlined in Kala *et al.* (2020) to provide a glimpse into the large number of potential variables influencing the COVID-19 cases and help distil them into a smaller number that might reveal hidden and unknown patterns.

Three techniques were utilized: self-organizing maps (SOM), parallel coordinate plots (PCP) and geographic mapping. SOM is a clustering method of data visualization that uses pre-specified sub-regions of a study area (*e.g.* counties) called 'elements', and groups those that are most similar in terms of the factors being investigated into a cluster. A cluster could therefore contain just one element, or it could contain many, but the elements included in a cluster are more similar to each other than to any element in another cluster. Clusters are displayed in a grid of hexagons, which are shaded light to dark to show the level of dissimilarity to neighbouring clusters.

Once a SOM has been created, a parallel coordinate plot is used to consider the data within the clusters. These are represented as a line graph with multiple vertical axes, one for each of the factors included in the analysis. All of the clusters are represented by a line on the graph, so if there are 30 clusters in the SOM then there are 30 lines on the PCP, and the thickness of the line indicates the number elements within the cluster. The point at which a line intersects an axis indicates the relative value of the given factor for that specific cluster. Relationships between factors can be determined by examining the points of intersection between each line (a cluster) and each vertical axis in the graph. The PCP uses a nested-means scaling on each axis, a nonlinear scaling method that recursively calculates several mean values and uses the values as break-points to divide each axis into equal-length segments. Thus, nested-means scaling puts the mean value at the centre of each axis and allows different axes defined by different units to be comparable.

Geographic mapping then displays where the elements of these clusters are on a map of the study area. When a cluster is selected all the elements of that cluster are highlighted on the map. In essence for this study, maps were produced that show which counties fall into which clusters because of statistical similarity across the factors analysed. These results could suggest approaches for further analyses or suggest areas of focus for public health interventions.

In addition to the lockdown hypothesis of cascading effects of COVID-19 incidence on population mobility and subsequent NO_2 emissions, the social vulnerability across Texas' population was explored. A community's overall social characteristics influence the type, degree and nature of public health interventions. These characteristics include socio-economic status, household composition, minority status, and housing type. These factors have been used to describe a community's social vulnerability to help public health officials and emergency response planners identify and map the communities that will most likely need support before, during and after a hazardous event (Agency for Toxic Substance and Disease Registry, 2020).

In the research presented here, the factors used to develop the PCP axes for exploratory analysis were: i) COVID-19 incidence rate; ii) COVID-19 deaths; iii) the difference in population mobility (year 2020 compared to 2019); iv) the difference in NO_2 concentration (year 2020 compared to 2019); v) percent population aged 65 years and older; vi) percent of population living in multi-unit structures; vii) percent of population that live in households with no vehicles; vii) percent of population classified as minority; ix) unemployment rate; x) percent of adult population without a high school diploma; xi) poverty rates.

Results

The first reported case of COVID-19 in Texas occurred on March 4th, 2020 (Texas Department of State Health Services, 2020), what is referred to herein as Event 1 (Figure 4). Fifteen days later, March 19th, Texas' Commissioner of Public Health declared a Public Health Disaster (Hellerstedt, 2020), and the Governor issued an Executive Order (Abbott, 2020a) intended to mitigate the spread of the virus, referred to as Event 2, and frequently considered to be the Texas lockdown. The lockdown ordered limits on the size of social gatherings, closing of non-essential business (*e.g.* restaurants, hair salons), ban on visitations of nursing homes and retirement/long-term facilities, and the closing of schools. Approximately one month later, April 17th, lockdown restrictions were eased by two Executive Orders (Abbott, 2020b, 2020c). These two Executive Orders, Event 3, expanded the definition of essential services beyond the U.S. Department of Homeland Security's Guidance on the Essential Critical Infrastructure, to also include religious services (conducted in churches, congregations and houses of worship) and retail services that were not previously defined as essential were allowed to provide services through pick-up, delivery by mail or delivered to a customer's doorstep.

Approximately five weeks later, on June 3rd, the Governor issued another Executive Order (Abbott, 2020d) that continued to reopen the State (Event 4), allowing all businesses to operate at 50% occupancy, some businesses at 75% occupancy (*e.g.* restaurants with alcoholic beverage sales less than 51% of total sales), and no occupancy limits for many other types of activities (*e.g.* religious services, local government operations, child-care services, youth camps). On that day, June 3rd, the day's reported number of confirmed COVID-19 cases in Texas was 1540 (Johns Hopkins Coronavirus Resource Center, 2021). Two weeks after Event 4, reported incidence of new confirmed daily COVID-19 cases had increased to 4258 cases (Johns Hopkins Coronavirus

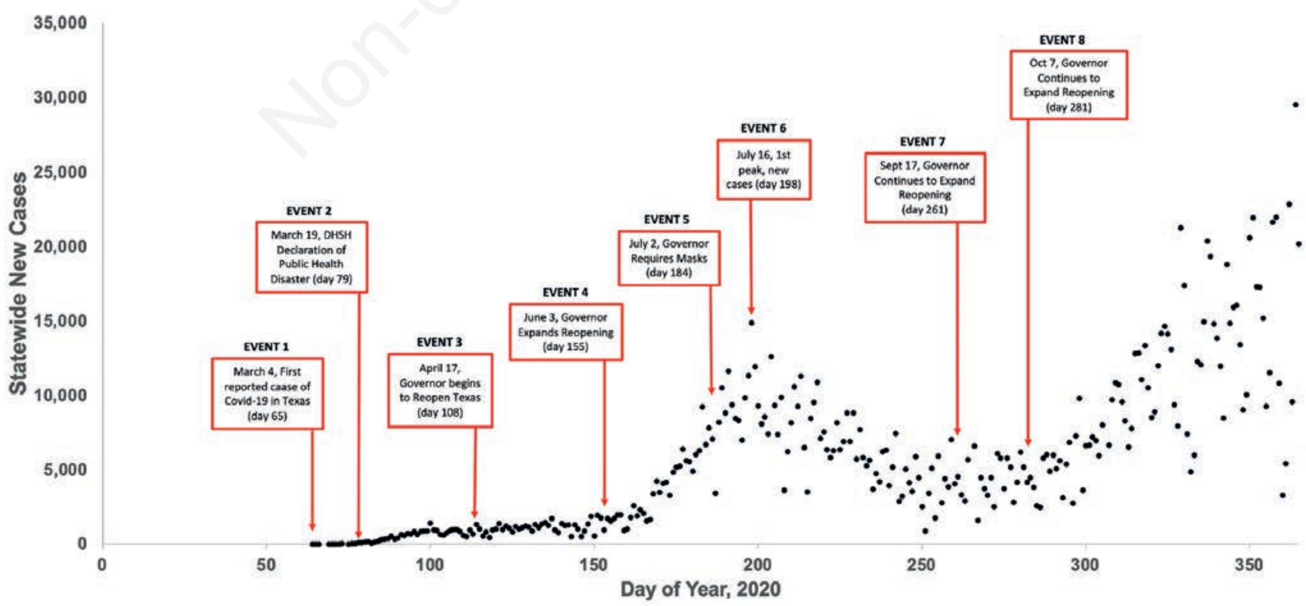

Figure 4. Reported number of new daily COVID-19 cases in Texas, 2020, with various events noted.

Resource Center, 2021). Daily COVID-19 incidence continued to rise, prompting Event 5, an Executive Order issued on July 2nd (6725 cases) that mandated the wearing of face coverings in commercial establishments, other buildings and space open to the public (Abbott, 2020e). Multiple exemptions to wearing face coverings were provided, including, for example, youth under 10 years of age, those consuming food or drink in a restaurant, anyone providing or obtaining religious worship, or anyone voting, assisting voters, or serving as poll watchers.

Daily new confirmed COVID-19 cases continued increasing until the new peak daily incidence occurred on July 16th, at a level of almost 15,000 cases per day (Event 6, Johns Hopkins Coronavirus Resource Center (2021). Almost exactly 2 weeks after the face covering mandate, daily incidence reversed course for the remainder of the summer and continued dropping through early September to daily reported incidence as low as 867 (September 7th). On September 17th, the Governor issued another Executive Order, Event 7 (Abbott, 2020f), that continued reopening Texas, when the new daily COVID-19 incidence was reported to be 4,565 cases (Johns Hopkins Coronavirus Resource Center, 2021). Another Executive Order (Event 8) was issued 3 weeks later on October 7th that continued reopening the state (Abbott, 2020g). Event 8 was issued on a day when 3776 new cases of COVID-19 were reported. New cases of COVID-19 trended upward after Event 8, albeit with much higher variability, until the end of the year. The highest daily incidence of new COVID-19 cases, 269,535, was reported for Texas on December 29th, 2020 (269,535 new cases) (Johns Hopkins Coronavirus Resource Center, 2021).

The response to these policy actions, including the lockdown, was different across Texas. Some parts of the state, mostly rural, showed an increase in mobility in 2020 compared to 2019 (Figure 5A) suggesting different population-level responses to the lockdown and other recommendations of social distancing. This partially agrees with the lockdown hypothesis, and partially disagrees. In general, activity in the three major metropolitan areas in Texas - Houston (Figure 5B), Dallas-Fort Worth (Figure 5C), and the Austin-San Antonio corridor (Figure 5D) - showed decreases in relative mobility between 2020 and 2019. However, it is important to note that these generalized observations of increasing or decreasing mobility across Texas cannot simply be attributed to the rural-urban divide. For example, rural areas in West Texas showed decreases in mobility while many sub-urban areas around the large urban centres showed patterns of overall increase in mobility. Population-level responses to these policy measures is a complex process that is driven by a multitude of factors, including demographic structures, socio-economic conditions, and political outlooks. Exploratory data analysis techniques to explore the lockdown hypothesis, such as geovisualization, was used to yield more nuanced insights into how and why populations responded differently in Texas.

Figure 5. Changes in population mobility - units represent difference in total visits recorded at most visited Point of Interest in each census block when March through December 2019 is compared to the same time window in 2020. A - State of Texas; B - Houston area; C - Dallas area; D - Austin/San Antonio area.

The lockdown hypothesis of cascading effects of policy decisions cascading into impacts on population mobility and air quality was assessed using two surfaces showing the patterns of NO_2 estimates across Texas in 2019 compared to 2020. Zonal statistics of the NO_2 concentrations were computed for the baseline year 2019 and the pandemic year 2020 (Figure 6). The reduction in the mean NO_2 concentration differences for each county was then calculated and further analysed. It is evident that not all counties in Texas experienced an improvement in air quality during the pandemic year 2020. The results in part agree with the lockdown hypothesis and disagree in part. Out of 254 counties in Texas, approximately 30% experienced a reduction in air pollution between the baseline year and the pandemic year. With the backdrop of COVID-19 incidence and the public policy events associated with the pandemic's effects in Texas, the hypothesis of cascading effects of the statewide lockdown and subsequent reopenings on Texan's mobility (as measured by SafeGraph anonymous opt-in consent location data) and air quality (as measured by satellite measured nitrogen-dioxide [NO_2] column density) were examined using geovisualization.

The grey and blue bars in Figure 7 show mobility (top panel) and NO_2 (bottom panel) patterns for baseline year 2019 and pandemic year 2020 respectively. Mobility patterns in Texas show a noticeable decline in volume soon after the lock down was announced in early March 2020. Although the SafeGraph data shows gradually increasing volumes in mobility starting in mid-April and levelling off towards the end of 2020, mobility was substantially lower than volumes observed in 2019. Mobility observed during the early part of 2020 (January through mid-March) was higher than that observed during the same time window in 2019. But the effect of the lockdown in mid-March was evident in the SafeGraph data, showing a distinct reversal in mobility volumes between mid-March and end of year 2019 compared to the same time window in 2020 as suggested by the lockdown hypothesis. In comparison, the differences between 2019 and 2020 NO_2 data are not quite as apparent, with both years experiencing similar cyclical patterns (highest NO_2 concentrations in the winter and lowest in the summer). It is important to note that the original spatial resolution of the mobility and the NO_2 datasets are markedly different. Mobility data are collected using individual, device-level data which are then aggregated up to the census block group and then the county level. In comparison, the NO_2 data are measured at a coarser spatial and temporal scale (3.5×7 km, daily coverage). This may partially explain why the mobility data agrees with the lockdown hypothesis than does the NO_2 data.

Texas contains 254 counties (geographic subdivisions, each with its own administrative jurisdiction). When examining mobility and NO_2 reductions at the county level there is a clear overlap between the counties with the largest reductions in mobility and those with the largest reductions in NO_2 concentrations. As shown in Figure 8, six of the top seven counties in terms of mobility drop, are in the top nine counties in terms of NO_2 drop.

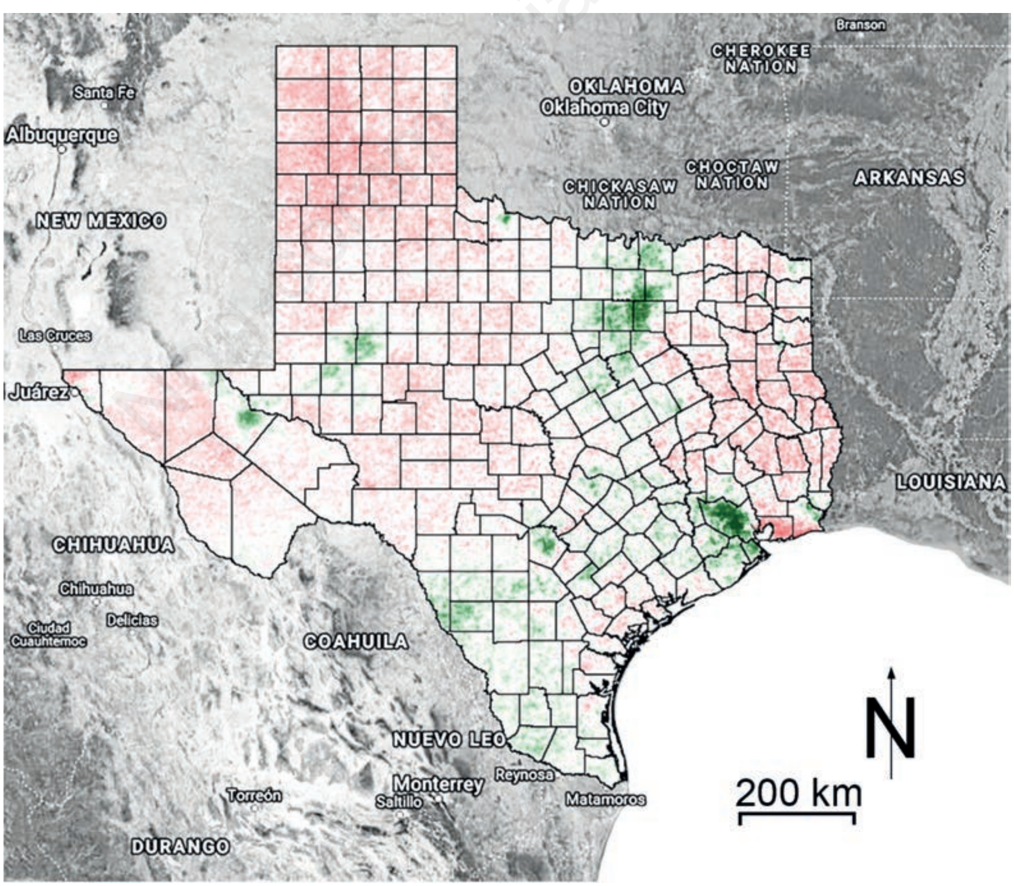

Figure 6. Annual average change in NO_2 concentration between 2019 and 2020. Green represents decrease in NO_2; red represents increase.

Discussion

Public health policy responses to COVID-19 varied globally, but it is overwhelmingly clear that the pandemic prompted multiple strategies to combat the effects of the disease. This study explored the lockdown hypothesis of cascading effects of the response to the virus in Texas on mobility and air quality, as well as demographic place characteristics of clusters of COVID-19 cases across Texas using geovisualization. Multiple insights can be gleaned from this exploration of data.

Insight 1 - Face coverings provide public health benefits

The COVID-19 trends shown in Figure 4 suggest that each Event documented above, those that involved both the lockdown itself and the reopening of the state, were followed by trends of increasing incidence of daily COVID-19 cases. The only Event that was eventually followed by decreasing numbers of daily cases was the face covering mandate (Event 5); two weeks after the mandate COVID-19 incidence peaked (Event 6), there was a marked reversal of daily COVID-19 incidence, even though mobility counts continued to increase during that same time. The decrease in COVID-19 incidence continued to decrease until the next Executive Order was issued (Event 7) that eased restrictions, after which numbers began increasing again. Further, public health research suggesting that the latency period between contracting the virus and onset of symptoms is between 2 and 14 days appears to be corroborated by the 2-week lag between Texas' face covering order and the onset of decreasing cases at the population level.

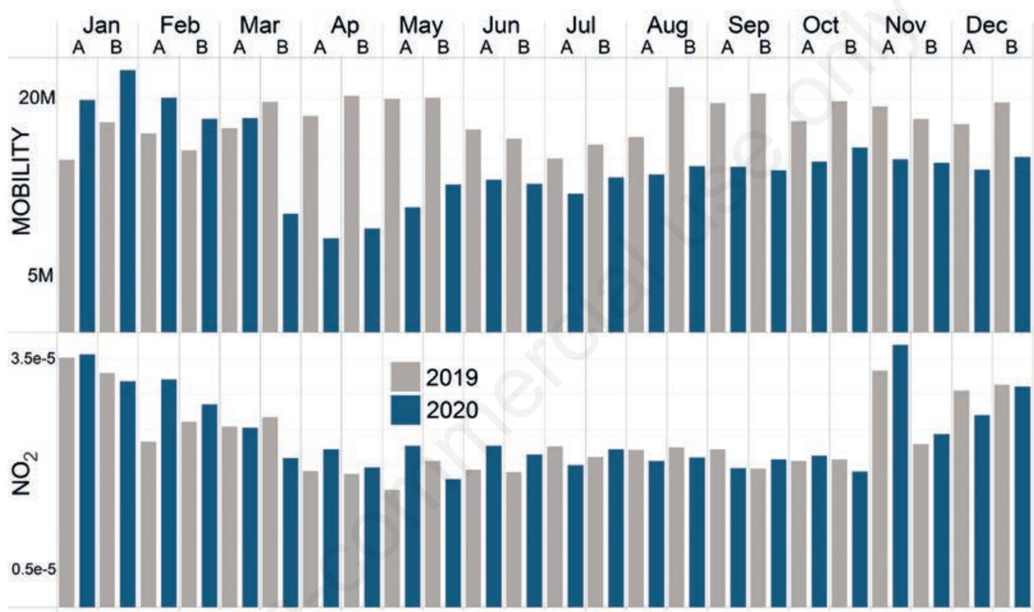

Figure 7. Comparison of 2019 (grey) *versus* 2020 (blue) in 2-week increments for: population mobility (upper graph), and; NO_2 concentration for Texas (lower graph).

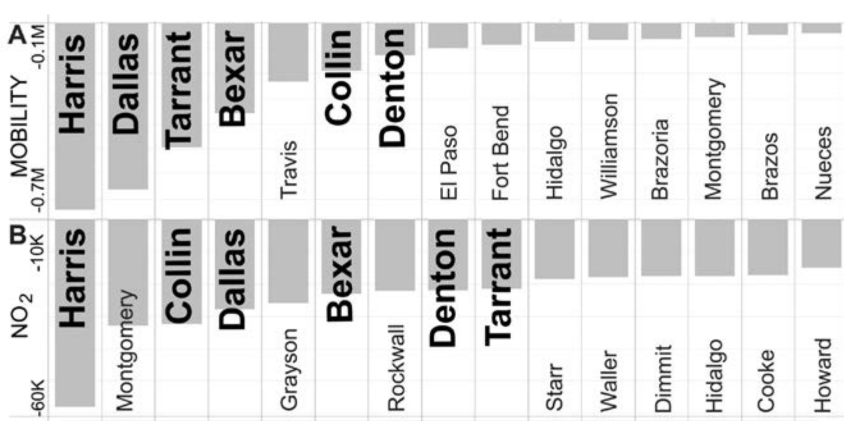

Figure 8. Top 15 counties (out of 254) with maximum drop in: (A) population mobility, or; (B) drop in NO_2 air quality between 2019 and 2020 in the March through December window. Six of the top 9 counties are in common.

Insight 2 - Public interaction policies

The initial lockdown on 19 March 2020 (total number of new COVID-19 cases reported to be 109) was followed by a steadily increasing number of cases averaging just under 600 daily new cases over the next month. April 17th (Event 3) brought the first of the official reopenings of Texas, followed by an expanded reopening on June 3rd (Event 4); the face covering mandate was issued on July 2nd (Event 5). Between the expanded reopening of Texas on June 3rd and the mask mandate on July 2nd the incidence of COVID-19 cases rose to an average of over 3700 new daily cases during those 29 days, reaching 7831 cases on the day the face covering mandate was issued.

Insight 3 - Mandated lockdown influenced population mobility differentially

When comparing the counts of the most visited POIs in each county for corresponding days between 2019 and 2020 for the lockdown hypothesis, it was clear that major urban areas saw large decreases in numbers of visits after the lockdown, while smaller urban areas along with rural areas saw less of an effect. When comparing the first 3 months of 2019 to 2020, mobility increased statewide (see Figure 7). However, beginning with the second half of March (after the lockdown), there was distinctly less mobility in 2020 than in 2019. The largest decrease in mobility between 2019 and 2020 was observed in the second half of March and first half of April, and while there was a steady increase in mobility after the April bottom, mobility never reached baseline year 2019 levels for the remainder of the pandemic year 2020. Statistical clustering further shows that the counties with the highest incidence of COVID-19 cases also had the largest decrease in mobility, suggesting that people living in areas with large numbers of infections intentionally avoided traveling, while the mobility of people in areas with lower infection numbers was less likely to be affected.

Insight 4 - Air quality is related to mobility, especially in highly urbanized areas

Air pollutant emissions from the transportation sector have been well studied, showing that transportation related pollutants include particulate matter, volatile organic compounds and nitrogen oxides. While nitrogen oxides in the atmosphere can be formed from natural causes, the primary anthropogenic source of nitrogen oxide, measured as NO_2, is from fossil fuel combustion processes. The U.S. Environmental Protection Agency estimates that over 55% of total nitrogen oxide emissions in the U.S. are from the transportation sector (U.S. Environmental Protection Agency, 2018). When comparing NO_2 concentrations in Texas for corresponding days between 2019 and 2020 for the lockdown hypothesis, the two years were quite similar for the state as a whole (see Figure 7). However, when focusing on counties that had the largest decreases in NO_2 (Figure 8), these were the same areas that had the largest decreases in mobility. Texas has 254 counties: six of the top twelve counties in Texas in terms of NO_2 decrease between the two years were also in the top seven counties in terms of decreases in mobility. These are the counties that encompass the largest metropolitan areas including Houston, Dallas-Fort Worth and San Antonio.

Insight 5 - Health vulnerability characteristics don't tell the full story

There is a general public health hypothesis that socio-economic data will be descriptive of the health vulnerability of an exposed population, and often public health policies are developed with this tenant in mind. This hypothesis seems to hold very well under the lockdown hypothesis for the cluster that contained the counties with lowest number of COVID-19 cases), and fairly well for the cluster that contained the counties with highest number of COVID-19 cases (Figures 9 and 10). When focusing on the Texas cluster with the lowest incidence of COVID-19 cases under the lockdown hypothesis, there is a very clear trend: i) it also had the lowest numbers of COVID-19 deaths; ii) its mobility was essentially unaffected; iii) its NO_2 concentration did not decrease; iv) it had some of the highest percentages of people aged 65 years and older; v) it had low numbers of high-density housing, and; vi) it had the lowest rates of households without a vehicle, the lowest minority populations rates, the lowest unemployment rates, low rates of

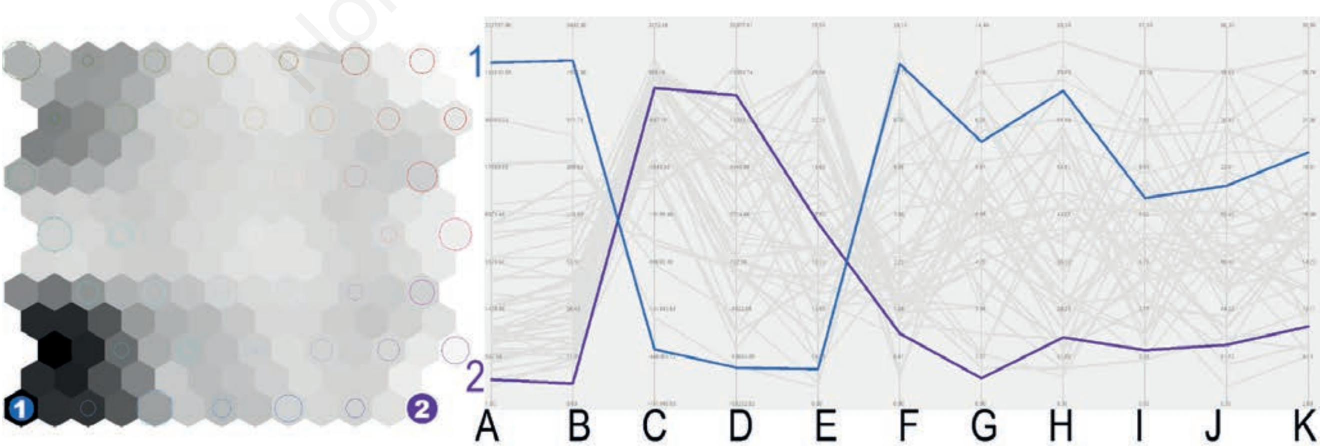

Figure 9. Self-organizing map (SOM) and parallel coordinate plot (PCP) showing the highest (Cluster 1 - blue) and lowest (Cluster 2 - purple) COVID-19 incidence clusters. *Variables determining clusters: A-COVID-19 incidence; B-COVID-19 deaths; C-change in population mobility; D-change in NO_2 concentration; E-percent of population over 65 years; F-percent of housing with 10 or more units per unit (e.g. apartments); G-percent of households with no vehicles; H-percent minority population; I-unemployment rate; J-percent of population without high school diploma (>25 years of age); K-percent of population below poverty line.*

individuals not having a high school diploma, and low rates of poverty. Other than containing counties with relative high rates of people aged 65 and over, all of the other socio-economic factors followed assumptions of public health vulnerability nearly exactly in this cluster. This cluster covers 6 rural counties in the western half of Texas.

For the cluster with the highest incidence of COVID-19 cases, there is also a clear trend, albeit not quite as clear as the cluster with the lowest COVID-19 incidence. The highest COVID-19 cluster shows: i) it had the highest numbers of COVID-19 deaths; ii) it had the largest decrease in mobility; iii) it had the largest decrease in NO_2 concentration; iv) it had the lowest percentage of people aged 65 and older; v) it had the highest rate of high-density housing; vi) it had high levels of households without a vehicle; vii) it had high minority population rates; and viii) it had moderate unemployment rates, moderate rates of individuals with no high school diplomas, and moderately high poverty rates. The socio-economic factors in this cluster also agree with current public health vulnerability assumptions quite closely. This cluster covers 6 counties, all highly urbanized counties in the eastern half of Texas.

Things get more complicated when focusing on clusters in the middle of COVID-19 incidence numbers. Of the 49 clusters generated in this study, the 5 clusters in the middle of the COVID-19 numbers presented demographic data that varied widely and somewhat unexpectedly from the public health vulnerability assumptions. These 5 clusters are represented by 32 counties which are spread across Texas. The 5 clusters all have similar numbers of COVID-19 cases and deaths, as well as change in mobility, but as shown in Figure 11, they span the gamut of NO_2 change, are in the lower half of clusters with rates of people aged 65 and older, have moderate rates of high-density housing, a wide range of households without vehicles, and nearly cross the entire range of minority population, school diplomas, and poverty rates. The counties that fall within these 5 clusters are widespread across Texas (Figure 12), but none are in the heavily populated areas of the state. This pattern suggests that while public health vulnerability assumptions concerning socio-economic data hold valid for the low and high extremes of COVID-19 incidence, they cannot be consistently applied towards clusters with mid-level numbers of incidence. The implication here is that health vulnerability indicators may need to be selected more discriminately for moderate levels of incidence, but more generally selected for the lows and highs of incidence.

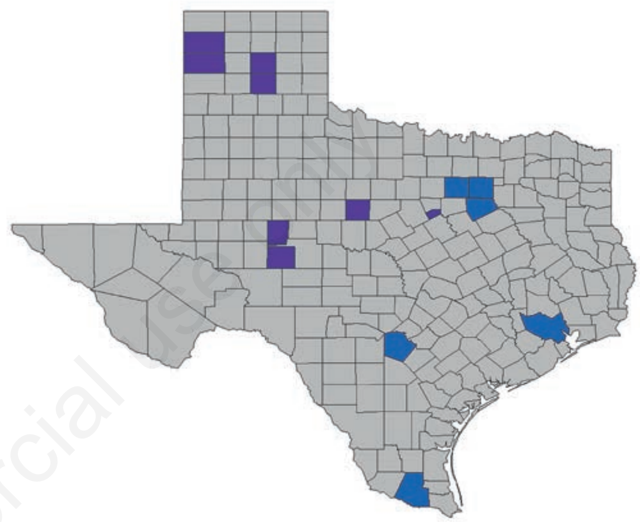

Figure 10. Texas counties contained in the highest (Cluster 1 - blue) and lowest (Cluster 2 - purple) COVID-19 incidence clusters.

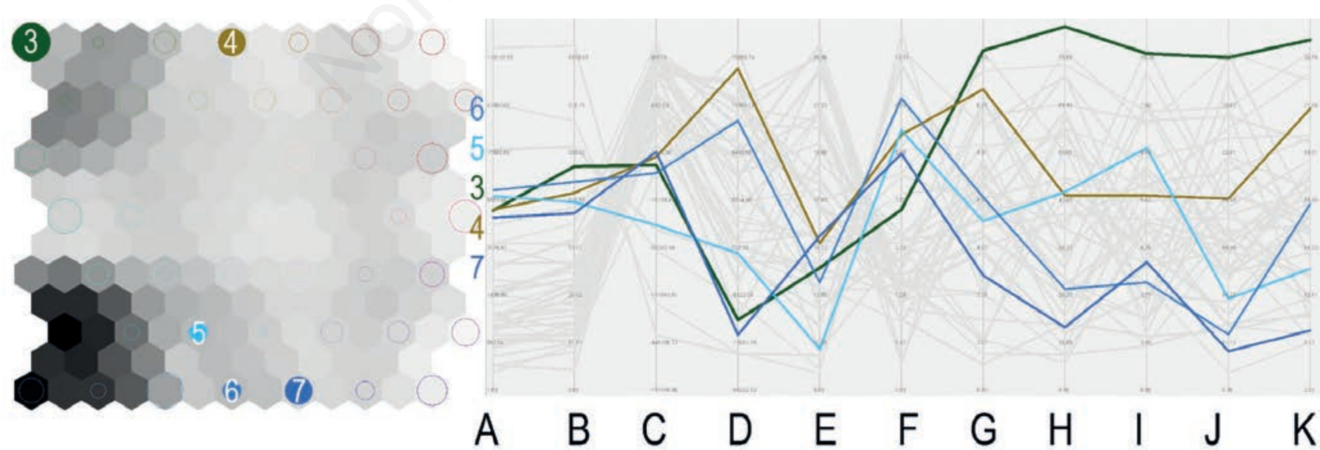

Figure 11. Self-organizing map (SOM) and parallel coordinate plot (PCP) showing in the mid-range of COVID-19 incidence clusters. ***Variables determining clusters: A-COVID-19 incidence; B-COVID-19 deaths; C-change in population mobility; D-change in NO_2 concentration; E-percent of population over 65 years; F-percent of housing with 10 or more units per unit (e.g. apartments); G-percent of households with no vehicles; H-percent minority population; I-unemployment rate; J-percent of population without high school diploma (>25 years of age); K-percent of population below poverty line.***

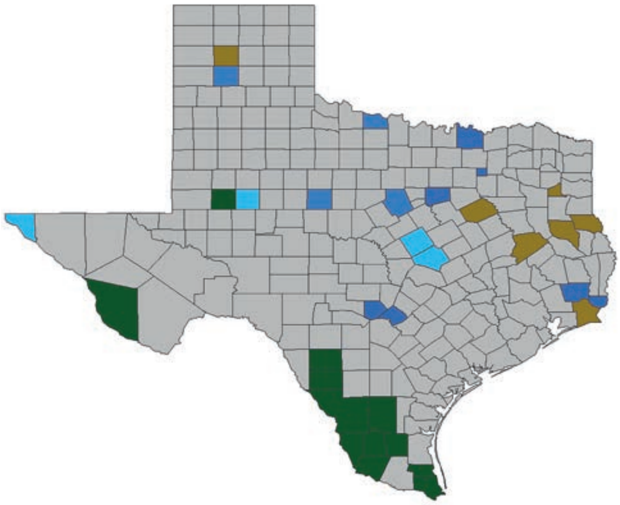

Figure 12. Texas counties contained in the mid-range of COVID-19 incidence clusters.

Conclusions

The enormous global effect of COVID-19 is difficult to comprehend. With all of the adverse consequences that have occurred, it is important to examine those things that might yield insights that could reduce, ameliorate or mitigate future pandemics. This study took advantage of a natural experiment when a 'lockdown' of a large population allowed an exploration of the cascading effects of that lockdown on population mobility and air quality. Those data allowed a further exploration of a few health vulnerability indicators that characterize spatial clusters of COVID-19 case numbers. Our data exploration shows partial agreement and partial disagreement with lockdown hypothesis. This natural experiment in Texas suggests that the restrictions imposed in Texas had a direct influence on the population mobility of Texans in highly urbanized areas, which cascaded into a reduction in air pollution emissions. Rural areas experienced less of these cascading effects. Demographic and socio-economic health vulnerability indicators across Texas seemed to follow public health assumptions in areas with the very lowest and the very highest COVID-19 incidence but were not consistent in areas with COVID-19 incidence numbers in the middle of the range across Texas. Finally, the Texas data strongly suggest that face coverings protect public health at the population level. Further work is needed to determine whether Texas is unique in these characteristics, or if other states in the USA, or perhaps other countries, experience similar patterns.

References

Abbott G, 2020a. Executive Order GA 08 Relating to COVID-19 preparedness and mitigation. Office of the Texas Governor March 19. Available from: https://gov.texas.gov/uploads/files/press/EO-GA_08_COVID-19_preparedness_and_mitigation_FINAL_03-19-2020_1.pdf

Abbott G, 2020b. Executive Order GA-15 Relating to hospital capacity during the COVID-19 disaster. Office of the Texas Governor April 17. Available from: https://gov.texas.gov/uploads/files/press/EO-GA-15_hospital_capacity_COVID-19_TRANS_04-17-2020.pdf

Abbott G, 2020c. Executive Order GA-16 Relating to the safe, strategic reopening of select services as the first step to Open Texas in response to the COVID-19 disaster. Office of the Governor April 17. Available from: https://gov.texas.gov/uploads/files/press/EO-GA-16_Opening_Texas_COVID-19_FINAL_04-17-2020.pdf

Abbott G, 2020d. Executive Order GA 26 Relating to the expanded opening of Texas in response to the COVID-19 disaster. Office of the Texas Governor June 3. Available from: https://gov.texas.gov/uploads/files/press/EO-GA-26_expanded_opening_COVID-19.pdf

Abbott G, 2020e. Executive Order GA 29 Relating to the use of face coverings during the COVID-19 disaster. Office of the Texaas Governor July 2. Available from: https://gov.texas.gov/uploads/files/press/EO-GA-29-use-of-face-coverings-during-COVID-19-IMAGE-07-02-2020.pdf

Abbott G, 2020f. Executive Order GA 30 Relating to the continued response to the COVID-19 disaster as Texas reopens. Office of the Texas Governor 5 pgs. Available from: https://gov.texas.gov/uploads/files/press/EO-GA-30_expanded_openings_COVID-19.pdf

Abbott G, 2020g. Executive Order GA 32 Relating to the continued response to the COVID-19 disaster as Texas reopens. Office of the Texas Governor October 7. Available from: https://gov.texas.gov/uploads/files/press/EO-GA-32_continued_response_to_COVID-19_IMAGE_10-07-2020.pdf

Agency for Toxic Substance and Disease Registry, 2020. CDC SVI 2018 Documentation. U.S. Department of Health and Human Services. Available from: https://www.atsdr.cdc.gov/placeandhealth/svi/documentation/SVI_documentation_2018.html

Castellanos P, Boersma KF, 2012. Reductions in nitrogen oxides over Europe driven by environmental policy and economic recession. Sci Rep 2:1-7.

Chan JF-W, Yuan S, Kok K-H, To KK-W, Chu H, Yang J, Xing F, Liu J, Yip CC-Y, Poon RW-S, 2020. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet 395:514-23.

Crippa M, Guizzardi D, Muntean M, Schaaf E, Dentener F, Van Aardenne JA, Monni S, Doering U, Olivier JG, Pagliari V, 2018. Gridded emissions of air pollutants for the period 1970-2012 within EDGAR v4. 3.2. Earth Sys Sci Data 10:1987-2013.

Curier R, Kranenburg R, Segers A, Timmermans R, and Schaap M, 2014. Synergistic use of OMI NO2 tropospheric columns and LOTOS-EUROS to evaluate the NOx emission trends across Europe. Remote Sens Env 149:58-69.

Gao S, Rao J, Kang Y, Liang Y, Kruse J, 2020. Mapping county-level mobility pattern changes in the United States in response to COVID-19. SIGSpatial Sp 12:16-26.

Ghude SD, Pfister GG, Jena C, Van Der A R, Emmons LK, Kumar R, 2013. Satellite constraints of nitrogen oxide (NOx) emissions from India based on OMI observations and WRF□Chem simulations. Geophys Res Lett 40:423-8.

Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R, 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sens Env 202:18-27.

Hellerstedt JW, 2020. Declaration of a Public Health Disaster in the State of Texas, Department of State Health Services. Secretary of State March 19. Available from: https://www.dshs.state.tx.us/coronavirus/docs/DECLARATION-PublicHealthDisaster.pdf

Johns Hopkins Coronavirus Resource Center, 2021. Data

Download. The COVID Tracking Project updated daily. Available from: https://covidtracking.com/data/download
Kala AK, Atkinson SF, Tiwari C, 2020. Exploring the socio-economic and environmental components of infectious diseases using multivariate geovisualization: West Nile Virus. PeerJ 8:e9577.
Kim SW, Heckel A, Mckeen S, Frost G, Hsie EY, Trainer M, Richter A, Burrows J, Peckham S, Grell G, 2006. Satellite-observed US power plant NOx emission reductions and their impact on air quality, Geophys Res Lett 33.
Kraemer MU, Yang C-H, Gutierrez B, Wu C-H, Klein B, Pigott DM, Du Plessis L, Faria NR, Li R, Hanage WP, 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. Sci 368:493-7.
Lamsal LN, Duncan BN, Yoshida Y, Krotkov NA, Pickering KE, Streets DG, and Lu Z, 2015. US NO2 trends (2005-2013): EPA Air Quality System (AQS) data versus improved observations from the Ozone Monitoring Instrument (OMI). Atmos Env 110:130-43.
Lelieveld J, Evans JS, Fnais M, Giannadaki D, Pozzer A, 2015. The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature 525:367-71.
Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KS, Lau EH, Wong JY, 2020. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med [Epub ahead of print].
Pishue B, 2020. COVID-19's impact on freight: an analysis of long-haul freight movement during a pandemic. INRIX Res [Epub ahead of print].
Safegraph, 2021. Patterns Dataset. Available from: https://docs. safegraph.com/docs/monthly-patterns; & https://safegraphcovid19.slack.com/archives/C0109NPA543/p1585946233411800
Slotnick D, 2020. Coronavirus demolished air travel around the globe. Bus Insider 22 [Epub ahead of print].
Streets DG, Canty T, Carmichael GR, De Foy B, Dickerson RR, Duncan BN, Edwards DP, Haynes JA, Henze DK, Houyoux MR, 2013. Emissions estimation from satellite retrievals: A review of current capability. Atmos Env 77:1011-42.
Texas Department of State Health Services, 2020. DSHS Announces First Case of COVID-19 in Texas. Texas Health and Human Services March 4. Available from: https://www.dshs.texas.gov/news/releases/2020/20200304.aspx
Texas Department of State Health Services, 2021. Texas COVID-19 Data. Texas Health and Human Services updated daily. Available from: https://dshs.texas.gov/coronavirus/additional-data.aspx
U.S. Environmental Protection Agency, 2018. 2014 National Emissions Inventory Report. Available from: https://www.epa.gov/air-emissions-inventories/2014-national-emissions-inventory-nei-data#doc
Wang S, Zhang Q, Streets D, He K, Martin R, Lamsal L, Chen D, Lei Y, Lu Z, 2012. Growth in NO x emissions from power plants in China: bottom-up estimates and satellite observations. Atmos Chem Phys 12:4429-47.