University of Zurich[UZH]

Is traveling a luxury or a burden? Human mobility analysis during COVID-19

GEO 511 Master's Thesis

Author
Tao Peng
21-738-927

Supervised by
Dr. Cheng Fu

Faculty representative
Prof. Dr. Robert Weibel

30.09.2023
Department of Geography, University of Zurich

Universität Zürich[UZH]

Is traveling a luxury or a burden? Human mobility analysis during COVID-19

Tao Peng

21-738-927

GEO511 Master's Thesis

Supervised by: Dr. Cheng Fu
Faculty Representative: Prof. Dr. Robert Weibel

29. 09. 2023
Department of Geography, University of Zurich

Abstract

This thesis investigates the evolution of the relationship between human mobility and socioeconomic factors in the US. We find that poverty rate is negatively associated with human mobility before the pandemic outbreak. However, the association becomes positive after the outbreak. Similarly, we examine political partisanship and aging rate. The changes reveal that the social inequalities exacerbate during the COVID-19 pandemic.

Keywords: Human mobility, Socioeconomic factors, COVID-19

Acknowledgements

I would like to dedicate this work to all who made it possible.

It is a distinct honor to work with Dr. Cheng Fu and Prof. Robert Weibel who have encouraged me to grow a tiny research idea into the thesis, guided me through every challenge, shared much wisdom and bore with me. It is a great joy to walk this journey with all my classmates and friends who have become my brothers and sisters in this foreign land and made Zürich a home for me. I would never make it thus far if it were not for my parents who always support me even from afar and my husband who never ceases to cheer me on.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The COVID-19 pandemic marks a distinct division between the time before it and the time that follows. In the context of this thesis, we define the pre-pandemic period (January to December 2019), the pandemic period (February 2020 to April 2022), the post-pandemic period (April to October 2022). Over the three periods, we analyzed the relationship between human mobility and socioeconomic factors from a longitudinal and comparative point of view.

1.1 Motivation

Human mobility plays a critical role in the transmission of infectious diseases. Advances in technology and economy allow people to travel further and more frequently, expanding their influence over larger areas. Meanwhile, the frequency and the scale of Emerging Infectious Diseases have been increasing, for instance, the SARS pandemic in 2003, the H1N1 pandemic in 2009, the Ebola outbreak in 2013, COVID-19 outbreak in 2019 [Bambra, 2022]. The COVID-19 pandemic was declared by the World Health Organization as a major global health issue [Cucinotta and Vanelli, 2020]. In response, to curb the spread of the virus, various travel restrictions were implemented across countries to urge people to stay at home and avoid unnecessary travel.

Along with travel restrictions, the COVID-19 pandemic jeopardizes various aspects of socioeconomic well-being. The disproportionate impacts of COVID-19 on specific subsections of populations reveal increasingly salient social injustice [Bonotti et al., 2021]. In addition, the pandemic gives rise to conflicts due to the government's response to the pandemic. In the US, there is a wide variation in responses. For example, California implemented the Stay-at-home

Order on March 19th 2020, followed by many states in the following two weeks. Some states even imposed more stringent regulations. Whereas some states never issued any. In a particularly sensational case, members of a militia were arrested in relation to plotting to kidnap Michigan's Democratic Governor because they were opposed to the COVID-19 policies the governor imposed in early 2020 [Gregorian, 2021]. COVID-19 has also taken a significant toll on the economy, particularly due to the travel restrictions and lockdown measures. The World Bank reported the worst global economic recession since WWII [Felsenthal, 2020]. Many people lost their jobs. The transportation sector, tourism industry, and food service sectors faced huge challenges due to the travel restrictions.

Analysis of the relationship between human mobility and socioeconomic indicators plays an essential role in various fields. In epidemiology, understanding the characteristics of travel behavior among population groups is significant to establishing well-tailored public health policies to slow the transmission of diseases [Abdullah et al., 2020]. In social justice, the heterogeneity in mobility changes reveals the inequality arising from the pandemic. In transportation, analyzing how the pandemic influences travel patterns enhances customized mobility solutions to meet the needs of people from various backgrounds. Furthermore, human mobility patterns change over time, especially under the influence of COVID-19. Examining the evolution of the relationship between human mobility and socioeconomic indicators enables a deeper understanding of ongoing societal, epidemiology, and transportation issues and enhances policy efficacy.

1.2 Research Aim

This thesis aims to investigate how the COVID-19 pandemic has shaped the travel patterns of people across various socioeconomic backgrounds. To understand the changes brought about by the pandemic, we first examine the evolution of the relationship between human mobility and socioeconomic indicators from the pre-pandemic to the pandemic period. Second, we explore the evolution from the pandemic to the post-pandemic period. It has been more than three years since the initial outbreak of the pandemic. Travel restrictions were lifted and there are no more lockdowns. However, the effects of the pandemic will cast a long shadow into the future. Understanding the long-term influence of the pandemic is critical to identifying the communities that are still suffering, to informing decision-making and resource-allocating, and ultimately to improving social equality in the post-pandemic world.

1.3 Thesis Structure

In Chapter 2, we provide an overview of relevant literature, covering the fundamental aspects of human mobility analysis and the studies of the relationship between socioeconomic factors and traveling behaviors. We identify research gaps and propose research questions in Chapter 3, and present the data and preprocessing details in Chapter 4. In Chapter 5, we describe the methodology leveraged to answer the research questions. Three candidate regression models are described in detail. We present the results in Chapter 6 and discuss the findings and limitations in Chapter 7. In Chapter 8, we draw an overall conclusion and point out future research.

Chapter 2

Literature Review

The last decades have witnessed numerous research of human mobility. This section first presents the fundamental aspects of human mobility analysis, from data sources, and human mobility modeling, to human mobility patterns. Then we summarize current findings in the relationship between mobility and socioeconomic indicators before the pandemic and after the pandemic outbreak.

2.1 The Basics of Human Mobility Analysis

Human mobility refers to the movement of humans in space and time. It can be explored at individual or population level. This thesis focuses on population level mobility. Regarding spatial dimension, human mobility has been studied over scales as large as global, continent, country, and city, [Balcan et al., 2009] and as small as building [Zhao et al., 2008]. In terms of temporal dimension, granularities used in previous studies are day, week, and month [Kim and Kwan, 2021; Barbalat and Franck, 2022; Grossman et al., 2020]. In addition to the research questions, temporal granularity selection highly depends on the data collected.

2.1.1 Data Source

One essential aspect of human mobility analysis is the data collection technique [Asgari et al., 2013]. In the past, surveys, such as census or local travel surveys, were leveraged to take a glimpse of how people travel. However, both often fail to provide a dynamic picture of human mobility [Palmer et al., 2013; Barbosa et al., 2018]. The game-changing data is the mobile phone data, including CDR[1]-based and GPS[2]-based mobile phone data [Barbosa et al., 2018].

Mobile phone companies collect and maintain CDRs for billing purposes. Each record contains information about the time and the cell tower that the phone was connected to [Oliver et al., 2020]. One of the earliest works using large anonymous CDRs dataset is by [Gonzalez et al., 2008]. This groundbreaking research revealed the predictability of human mobility and popularized the use of CDRs in human mobility analysis.

The most accurate data on human movement is GPS-based mobile phone data [Barbosa et al., 2018]. It allows researchers to track the movement trajectories of individuals with a high degree of accuracy and temporal frequency. One of the most pioneering research projects is Reality Mining [Eagle and Pentland, 2006]. This project collected data by tracking participants' wearable devices equipped with GPS sensors to identify patterns and behaviors in human movement and interaction. Later, with the advances in information and communication technologies, the prevalence of smartphone usage enables the collection of mobility data through the GPS receiver in the mobile phone of users [Hu et al., 2022]. Mobile phone applications, such as Google Maps and Apple Maps, record the real-time locations of users once the Location Access is turned on. Billions of people carry their phones every day, which makes it possible to provide a large quantity of data on human movement. In addition to information technology companies, commercial companies like SafeGraph also provide mobility data.

In addition to these data sources, datasets from social media, Wi-Fi access points, and public transit systems also offer human mobility analysis with valuable information. However, the primary data source in COVID-19 studies is GPS-based mobile phone location data [Hu et al., 2022].

2.1.2 Human Mobility Modeling

Researchers often measure how far and how often an individual travels at the individual level. For example, Radius of Gyration conceptualizes the characteristic distance traveled. In addition, human mobility is associated with different kinds of places. It is also modeled by the places attached. For example, empirical measurements indicate that people have a tendency to return home on a daily basis [Gonzalez et al., 2008]. Some studies model the Home-dwelling-Time or Time-outside-Residential-Places to quantify mobility [Grossman et al., 2020; Huang et al., 2022].

[1] Call Data Records
[2] Global Positioning System

Individual travel patterns have a degree of stochasticity and mobility information at the individual level is subject to privacy issues. Consequently, many studies model human mobility by synthesizing individual-level mobility information to the population-level. Origin-Destination (OD) matrix is widely used in aggregated mobility modeling. OD matrix provides an estimation of the number of individuals traveling between locations during a certain period. In addition to the OD matrix, some studies utilize the summary statistics of individuals in a certain geographical area to measure population mobility, such as the medium travel distance of residents in a county [Kim and Kwan, 2021], the average home-dwelling time [Grossman et al., 2020].

2.1.3 Human Mobility Patterns

In general, aggregated human mobility has a strong temporal variability:

- Daily commuting pattern: Many people follow daily commuting patterns, traveling from residential locations to work or school [McKenzie and Rapino, 2011].
- Weekday-weekend pattern: Human mobility patterns exhibit repeatable diurnal behaviors that differ between weekdays and weekends [Dobler et al., 2021].
- Seasonal and holiday pattern: In the north hemisphere, weekly recorded movements reveal a peak in July and August, and the valley in January. There are also low points in holidays, such as Easter and Thanksgiving [Kraemer et al., 2020].

2.2 Human Mobility and Socioeconomic Indicators

Socioeconomic status plays an essential role in shaping people's travel behavior. During the past two decades, an increasing number of studies explored the relationship between human mobility and socioeconomic indicators. Existing studies mainly analyzed three categories, including demographic, composite, and political (Table 1). The Demographic indicators are commonly analyzed, covering income level, employment level, education level, age, population density, race/ethnicity, etc [Limtanakool et al., 2006; Silm and Ahas, 2014; Järv et al., 2015; Lamb et al., 2021; Sy et al., 2021; Bonaccorsi et al., 2020]. Some research utilized composite indicators. For instance, socioeconomic levels are measured from the combination of 134 indicators [Frias-Martinez et al., 2012]. With the outbreak of the COVID-19 pandemic, the association between socioeconomic indicators has drawn more attention. On top of the demographic and composite indicators, researchers also used political indicators, to distinguish between Democratic-leaning and Republican-leaning areas in the US [Grossman et al., 2020; Hsiehchen et al., 2020; Clinton et al., 2021; Zhang et al., 2023].

Table 1: Socioeconomic indicators used in human mobility analysis.

Indicator type	Indicator	Source
Demographic	Income level	[Pappalardo et al., 2015]
	Employment level	[Hanson and Hanson, 1981]
	Education level	[Pappalardo et al., 2015]
	Age	[Limtanakool et al., 2006]
	Gender	[Hanson, 1982]
	Race/Ethnicity	[Hu et al., 2022]
	Population density	[Sy et al., 2021]
	Essential workforce rate	[Garnier et al., 2021]
	Uninsured rate	[Sy et al., 2021]
Composite	Socioeconomic levels	[Frias-Martinez et al., 2012]
	Deprivation rate	[Pappalardo et al., 2015]
	Multidimensional poverty	Dueñas et al., 2021
	Income segregation	[Moro et al., 2021]
	Economic Deprivation	[Long and Ren, 2022]
Political	Percentage of Republican candidate votes in US Presidential Elections	[Grossman et al., 2020]
	Percentage of Democratic candidate votes in US Presidential Elections	[Grossman et al., 2020]
	US state government trifectas	[Barbalat and Franck, 2022]

2.2.1 Before the COVID-19 Pandemic

In the late 20th century, travel surveys were widely used to investigate travel behaviors of people across socioeconomic status. For instance, studies suggested that an individual's travel frequency is positively correlated with employment status by using travel diary data [Hanson and Hanson, 1981]. However, the studies were limited in covering only a small sample of the population during a short period of time due to the difficulties in data collection.

In the 21st century, mobile phone data has become a critical data source to study human mobility at very fine spatial-temporal granularity [Gonzalez et al., 2008; Iovan et al., 2013; Smith-Clarke et al., 2014]. An increasing number of studies examined the association between mobility and socioeconomic indicators. However, the conclusions are not universal. Some research stressed that better socioeconomic status is associated with larger mobility [Frias-Martinez et al., 2012]. Some presented the opposite result [Pappalardo et al., 2015]. Whereas, others indicated that there is no significant difference in mobility patterns across socioeconomic classes [Xu et al., 2018].

For example, a study suggests that better socioeconomic status is strongly linked to larger mobility [Frias-Martinez et al., 2012]. They computed six mobility indicators from mobile phone data (e.g., average weekly distance traveled, radius of gyration, etc.) and obtained a composite indicator (socioeconomic levels). On the contrary, a study indicated that better socioeconomic development is associated with smaller mobility [Pappalardo et al., 2015]. They analyzed two mobility indicators (entropy and radius of gyration) and four socioeconomic indicators (income, unemployment, education, and deprivation rate). However, another study reached a different conclusion from the aforementioned two studies [Xu et al., 2018]. By comparing Singapore and Boston metropolitan areas, they concluded that phone users across different socioeconomic statuses exhibit very similar characteristics in both cities. They utilized six mobility indicators (e.g., the number of activity locations, activity entropy, and travel diversity) and two socioeconomic indicators (housing price and per capita income).

The underlying reason for the different findings might be the complicated nature of human mobility. Using different mobility indicators may result in conflicting conclusions. The relationship varies across geographical regions and changes over time. Furthermore, the modifiable areal unit problem (MAUP) could also impact the analysis results.

2.2.2 During the COVID-19 Pandemic

The COVID-19 pandemic has brought about socioeconomic disturbance and has had a significant influence on travel patterns. A large number of studies investigated socioeconomic disparities in human mobility change. In addition to the demographic and composite indicators, political indicators are also investigated. Various mobility indicators are computed to model how far, how often, and how long people traveled, and the attached locations like residential neighborhoods, workplaces, transit stations. The widely used mobility indicators are travel volume, travel distance and radius of gyration, and home-dwelling time [Bonaccorsi et al., 2020; Iio et al., 2021; Dueñas et al., 2021; Garnier et al., 2021; Huang et al., 2022]. Existing research mainly focuses on the initial period of the pandemic (Table 2). The temporal granularity varies from day to week, and month [Garnier et al., 2021; Hu et al., 2022; Kim and Kwan, 2021]. The analyzed geographical area is as large as a country, or as small as a city, with spatial granularity from census block group to state [Grossman et al., 2020; Lamb et al., 2021].

Concerning the Demographic indicators, existing studies primarily examined the income level, the percentage of elderly citizens, race and ethnicity composition, and population density. Mobility exhibited significant but uneven reduction across those demographic landscapes [Hu et al., 2022]. Areas with higher income levels experienced a greater reduction in mobility. More precisely, higher-income groups saw a larger decrease in travel distances than lower-income

Table 2: Summary of existing studies in the association between human mobility and socioeconomic indicators during the pandemic.

Spatial (granularity)	**Temporal (granularity)**	**Mobility indicator**	**Source**
US (state)	Mar. - Apr.2020 (day)	Greatest percentage reduction in mobility	[Hsiehchen et al., 2020]
US (state)	Feb. - May.2020 (day)	Visits to transit station	[Barbalat and Franck, 2022]
US (county)	Mar. - May.2020 (day)	Travel distance Visitation rate Encounter rate	[Garnier et al., 2021]
US (county)	Mar. - Sep.2020 (month)	Travel distance	[Kim and Kwan, 2021]
US (county)	Mar. - Dec.2020 (week)	Visit change Time staying home	[Hu et al., 2022]
US (county)	Mar.2020 (day)	Median time spent at home	[Grossman et al., 2020]
US Metropolitan Areas (census block group)	Jan. - Aug.2020 (day)	Home-dwelling time	[Huang et al., 2022]
New York City (zip code)	Mar. - Apr.2020 (week)	A standardized change in subway use	[Sy et al., 2021]
New York City (zip code)	Apr.2020 (day)	Daily visits to points of interest	[Lamb et al., 2021]
Greater Houston Area (census tract)	Apr.2020 (month)	Total travel distance Radius of gyration Number of distinct visited locations Per-trip distance	[Iio et al., 2021]
Individuals	Apr. - Oct.2020 (day)	Reported activities	[Clinton et al., 2021]

groups in April 2020 [Iio et al., 2021]. Despite the similarity in increasing home-dwelling time, poor communities exhibited less time at home than wealthy communities from January to August 2020 [Huang et al., 2022]. It reveals the luxurious nature of staying at home. Racial disparities also present strong associations with mobility. Black populations showed a stronger reduction in daily travel distance and rate of visitation to nonessential places from February to May 2020 [Garnier et al., 2021]. Even though the elderly are at a higher risk of infection and death from COVID-19, they showed less decrease in traveling during the first pandemic wave [Hosseini and Gittler, 2020; Mallapaty, 2020].

Political Partisanship is also analyzed in many studies. It is more important than public health concerns in explaining people's willingness to avoid traveling during the pandemic in the US [Clinton et al., 2021]. The association between mobility and partisanship is mainly explored during the initial period of the pandemic. When looking at the association at aggregated geographical unit level, citizens living in Republican-leaning states were more mobile than those in Democratic-leaning states [Hsiehchen et al., 2020; Barbalat and Franck, 2022]. At county level, people living in Democratic-leaning counties tended to travel less while those in Republican-leaning counties traveled as usual [Kim and Kwan, 2021]. Governors' tweets about social distancing and staying home were more effective in reducing mobility in Democratic-leaning counties than in Republican-leaning counties [Grossman et al., 2020]. When looking at the individual level, Democrats were more likely to engage in social activities than Republicans [Clinton et al., 2021]. All the studies reveal the role of political affiliation in travel behavior.

Many countries implemented policies in response to COVID-19 to encourage people to stay at home, such as the reduction of public transit operations and workplace and school closures. Some research analyzed how the strictness of travel restriction policy is associated with mobility. The Stringency Index from Oxford COVID-19 Government Response Tracker was used in many studies [Hale et al., 2021].

Chapter 3

Research Gaps and Questions

3.1 Research Gaps

First, existing work did not assess how mobility changes over time during the pandemic in comparison to the pre-pandemic or post-pandemic time. Previous studies mainly analyzed human mobility during the initial period of the pandemic (Table 2), primarily from January to September 2020. On the one hand, limited studies compared human mobility during the pandemic with previous years. While some considered January 2020 as the pre-pandemic period, it introduced potential bias in the analysis of temporal mobility changes due to seasonal or holiday patterns. More robust conclusions about mobility changes can only be drawn by comparing with data from previous years. On the other hand, few studies explored mobility beyond 2020. Given that it has been over three years since the pandemic outbreak, it is worth examining the long-term effect of COVID-19 on mobility. After implementations and lifts of travel restriction policies, people might adapt to the 'new normal' of traveling [Emanuel et al., 2022]. Exploring the long-term effect of the pandemic on human mobility informs transport planners to customize strategies to meet travel demands and create a sustainable transport system in the post-pandemic world.

Second, previous work did not analyze the relationship between human mobility and socioeconomic indicators from a dynamic point of view. We lack the understanding of how the relationship evolved from the pre-pandemic to the pandemic, and to the post-pandemic time. Existing studies primarily examined the relationship statically by focusing on a single or several time points during the initial period of the pandemic. Considering the dynamic nature of human mobility and the social problems brought about by the rapidly changing pandemic, investigating the evolution of the relationship provides us with a deeper understanding of human mobility across various socioeconomic landscapes. Understanding the evolution of the relationship is critical to identify inequalities arising from the pandemic and persisting in the post-pandemic time. Therefore, policymakers could gain insights into allocating resources and enhancing social justice. It sheds light on both current and future pandemics and emergencies.

3.2 Research Questions

To address the research gaps, this thesis aims to explore the influence of COVID-19 pandemic on human mobility and its relationship with various socioeconomic backgrounds by answering the two research questions:

Research Question 1. How does COVID-19 shape human mobility?

We investigated the characteristics of aggregated human mobility during the pre-pandemic, pandemic, and post-pandemic periods. Then we conducted a comparative analysis of human mobility between the pre-pandemic and the pandemic period, the pandemic and the post-pandemic period, respectively. Two mobility indicators are utilized on a weekly basis and at county level in the US. Due to data availability, they do not cover the same period. The main mobility indicator, Travel Volume is from January 2019 to April 2021. The alternative mobility indicator, Residence Time is from October 2020 to October 2022. The study period is thereafter decomposed into the pre-pandemic period, pandemic period, and post-pandemic period. The definitions of the waves are based on the existing literature [Walensky, 2021; CDC, 2022].

- Pre-pandemic period (January – December 2019)
- Pandemic period (February 2020 - April 2022)
  - First wave (February to June 2020)
  - Second wave (June to October 2020)
  - Third wave (October 2020 to April 2021)
  - Fourth wave (April 2021 to October 2021)
  - Fifth wave (October 2021 to April 2022)
- Post-pandemic period (April – October 2022)

Research Question 2. How does the relationship between human mobility and socioeconomic indicators evolve over the course of the pandemic?

We first synthesized the socioeconomic indicators commonly analyzed in existing human mobility research. Then we conducted feature selection to choose the optimal combination of indicators, which includes Poverty Rate, Political Partisanship, and Aging Rate. Next, Multiple Linear Regressions were conducted to explore the relationship between mobility and socioeconomic indicators. We utilized the Ordinary Least Squares as baseline models to explore spatial dependence and then conducted model selection between the Spatial Lag Model and the Spatial Error Model. To investigate the evolution of the relationship, we performed the regression model per week and analyzed the change in regression coefficients. Specifically, we analyzed the change in the relationship from the pre-pandemic to the pandemic period using Travel Volume, and from the pandemic to the post-pandemic period using Residence Time.

Chapter 4

Data

4.1 Human Mobility Indicators

This thesis uses two human mobility indicators: Travel Volume and Residence Time. Both are open-sourced and GPS-based mobile phone data. Because of data availability, the two datasets do not cover the same period. The main mobility indicator, Travel Volume, is available from the pre-pandemic period to the third wave. The alternative mobility indicator, Residence Time, is available from the third wave to the post-pandemic period.

4.1.1 Travel Volume

Source Data. We retrieved the raw data of Travel Volume from a publicly available source [Kang et al., 2020], which contains daily origin-destination (OD) matrices at county level in the US. The data was collected from millions of anonymous mobile phone users' travel trajectories. The trajectories are provided by SafeGraph[1]. The users of SafeGraph account for approximately 10% of the total population in the US. SafeGraph indicated that their data is aligned with the US census data [SafeGraph, 2021]. In other words, it does not overrepresent or underrepresent individuals in counties with different socioeconomic backgrounds. Besides, the data has been used in many research to analyze travel patterns during COVID-19 in the US [Grossman et al., 2020; Hu et al., 2022; Huang et al., 2022].

OD matrix is a standard object in aggregated mobility studies [Barbosa et al., 2018]. Considering n counties in total, the element t_{ij} of an OD matrix T represents the estimated

[1] https://www.safegraph.com/products/places

number of people traveling from origin county i to destination county j. Therefore, T is an n-by-n matrix. For example, 7327 people traveled within county 1001, 35 people traveled from county 1001 to county 1003 on January 1st 2020 (Table 3). In the source data [Kang et al., 2020], three steps were performed to measure county-level OD matrices: (1) Home locations of anonymous phone users were estimated based on the common nighttime (6 pm – 7 am local time) during a six-week period. The home locations were aggregated at census block group (CBG[2]) level. The most frequent CBG is used as the 'home location' of each user. (2) Daily CBG to CBG flows were computed. The number of visits from the home CBG to the destination CBG was recorded every day. (3) The obtained OD matrices at CBG level were aggregated to county level.

Table 3: Sample records of the raw data for Travel Volume

geoid_o*	**geoid_d***	**date**	**travel_flows**
1001	1001	Jan. 1st 2020	7327
1001	1003	Jan. 1st 2020	35
1001	1005	Jan. 1st 2020	1

* geoid is a unique numeric identifier of a county. geoid_o and geoid_d represent the origin county and destination county, respectively.

Data Preprocessing. First, we calculated the daily out-flows O of each county. Daily out-flows O_i represents the total number of people traveling from county i to all the counties. Second, we calculated the baseline out-flows $\hat{O}$ of each county. The baseline calculation took the Google Community Mobility Reports as a reference [Aktay et al., 2020]. To eliminate the impact of weekly pattern, every county has 7 baselines for each day of the week (Monday to Sunday). The time span of the baseline is a 5-week period from January 3rd 2020 to February 6th 2020. We calculated 7 medians of the out-flows according to weekdays within the 5-week baseline period. For instance, the baseline of Monday is the median of the out-flows of the 5 Mondays.

Third, we calculated the main mobility indicator, Travel Volume, by comparing daily out-flows O to the corresponding baseline $\hat{O}$. As the populations vary among counties, the out-flows O range from 20 to over a million. All the counties are put at the same scale by comparing to their corresponding baselines. In this case, the Travel Volume lies near 1 for all the counties. Precisely, it ranges between 0.60 to 1.20. Fourth, as daily Travel Volume takes large computational power in the regression analysis in Chapter 5, we calculated the weekly average

[2] A Census Block Group is a geographical unit used by the United States Census Bureau which is between the Census Tract and the Census Block. https://tinyurl.com/2k5hkrpm

Travel Volume of each county. Thereafter, the Travel Volume referred to weekly measurements in this thesis.

4.1.2 Residence Time

Source Data. The raw data of Residence Time were retrieved from the Google Community Mobility Report [Google, 2022]. This dataset has been leveraged in the study of mobility change over time in response to COVID-19 policies. It is aggregated and anonymized from users who turn on the location history settings in Google Maps. This dataset provides the time series of 6 categories of mobility data: (1) transit stations; (2) retail and recreation places; (3) workplaces; (4) groceries and pharmacies; (5) residential places; (6) parks. Precisely, each of the 6 types of mobility data is compared to the baseline level [Aktay et al., 2020]. The method to calculate baselines is the same as Travel Volume. A main limitation of this dataset is the large missing data.

Data Preprocessing. In this thesis, we calculated the alternative mobility indicator, Residence Time, based on the residential places category from Google Community Mobility Reports [Google, 2022], as this category has relatively less missing data. In addition, Residence Time is also a weekly average measurement.

Residence Time is only available in 998 out of 3108 counties continuously from October 2020 to October 2022. However, the 998 counties are not representative of all the three socioeconomic indicators (Section 4.2) in all the 3108 counties as suggested by the result from the Cramér-von-Mises-Test [Genest et al., 2006]. Therefore, we sampled 300 representative counties. In the 300 counties, the Cramér-von-Mises-Test suggested that the three socioeconomic indicators follow the same distribution as those of the 3108 counties. For the sample selection, we first conducted re-sampling to estimate the non-parametric distribution of the socioeconomic indicators of all the 3108 counties. Then a weight was assigned to each county based on the estimated distribution. We performed importance sampling to sample from the 998 counties. We tried to sample as many counties as possible. As a result, 300 is the maximum number of counties that are representative of all the 3108 counties. In the analysis of the relationship between Residence Time and socioeconomic factors, we only used these 300 counties (Section 6.2.3).

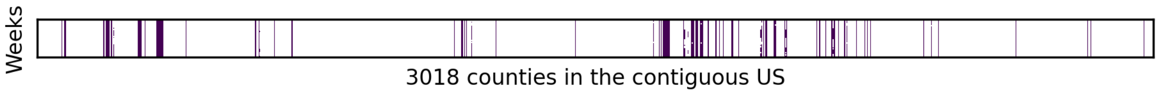


Figure 1: Missing data of Residence Time. The white space represents missing data. The X-axis represents all the 3108 counties. The Y-axis represents every week from October 2020 to October 2022.

4.2 Socioeconomic Indicators

Source Data. We synthesized eight socioeconomic indicators from existing research (Table 4). Most data were retrieved from the American Community Survey (ACS) 5-year estimates (2016-2020) by the US Census Bureau [Census, 2022], including the data on the aging rate, poverty rate, education rate, uninsured rate, median income, and population density. The data on racial and ethnic minority status is a composite indicator retrieved from the US Centers for Disease Control and Prevention[3]. The source data of political partisanship is the US 2020 Presidential Election results retrieved from the Harvard Dataverse[4]. political partisanship is measured by the percentage difference in the votes for the Republican and the Democratic candidates [Jung et al., 2017]. A positive value indicates a higher level of support for the Democratic party, suggesting a county with a Democratic-leaning tendency. On the contrary, it suggests the Republican-leaning tendency of a county.

Data Preprocessing. Feature selection is conducted to minimize multicollinearity of the socioeconomic indicators. We first preprocessed the data by dealing with outliers and normalizing all the indicators from 0 to 1, to put them in the same scale. To detect the multicollinearity among the indicators, the correlation matrix and Variance Inflation Factor (VIF) are analyzed. The value of VIF larger than 10 indicates a high degree of multicollinearity, which can make it difficult to interpret the individual effects of an independent variable on the dependent variable. Based on both the results from the correlation matrices (Figure 2) and VIF, three socioeconomic indicators are selected, namely, Poverty Rate, Political Partisanship, and Aging Rate (Table 5). The VIF is smaller than 5 for each of the three indicators.

[3]https://www.atsdr.cdc.gov/placeandhealth/svi/index.html
[4]https://www.dataverse.harvard.edu

Table 4: Summary statistics of the raw data of the socioeconomic indicators.

Socioeconomic Indicators	Description	Mean	St.d.	Min.	Median	Max.
Aging Rate	The percentage of population over 65 years old	19.28	4.75	3	18.9	57.8
Poverty Rate	The percentage of population below national poverty level	14.636	6.1776	0	13.7	58.9
Education Rate	The percentage of population with bachelor's degree or higher	16.28	7.05	0	14.61	60.49
Insured Rate	The percentage of population with health insurance	90.52	5.03	57.4	91.6	99.5
Median Income	Median income in 1000 US dollars	54.84	14.58	11.29	52.66	147.1
Population Density	Population density in 10000 persons/sq. mile	88.56	495.9	0.066	16.482	18676
Racial & Ethnic Minority Status	The ranking of the percentage of racial & ethnic minority population	23.89	19.891	0	16.7	99
Political Partisanship	The percentage difference of the votes for the Republican and the Democratic candidates	-31.82	31.284	-92.03	-38.4	86.76

Figure 2: The correlation matrix of all the socioeconomic indicators.

Table 5: Summary of selected socioeconomic indicators.

Indicators	Mean	SD	Min.	Median	Max.	VIF
Poverty Rate	0.23	0.1	0	0.23	1	4.21
Political Partisanship	0.34	0.17	0	0.3	1	4.91
Aging Rate	0.3	0.09	0	0.29	1	3.99

Note: all the three indicators were normalized.

4.3 Policy Stringency

We leveraged the indicator, Policy Stringency, to quantify the strictness of travel restriction policies, as people's travel behavior is potentially influenced by the governments' policies. The raw data of Policy Stringency was retrieved from the Oxford COVID-19 Government Response Tracker [Hale et al., 2021]. It is a composite indicator based on 9 responses, including school closures, workplace closures, cancellation of public events, restrictions on public gatherings, closures of public transport, stay-at-home requirements, public information campaigns, restrictions on internal movements, and international travel controls.

As there is no publicly available county-level data, we calculated the indicator, Policy Stringency, by applying the state-level data to counties. All the counties within the same state share the same travel restriction policy. In addition, we calculated the weekly average for each county and normalized the data. The change in the Policy Stringency is illustrated in Figure 3. The Policy Stringency is weekly measurements from January 2020 to October 2022. A higher score indicates a stricter government response to control COVID-19.

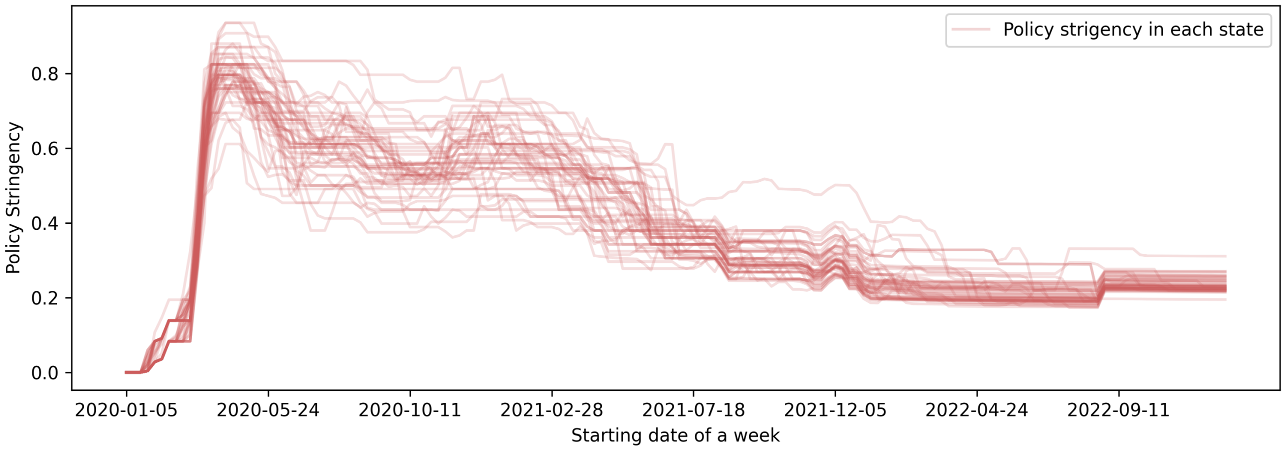


Figure 3: The temporal change of Policy Stringency in each state.

Chapter 5

Methodology

This thesis examines human mobility at county level in the Contiguous US[1]. The summary of variables can be found in Table 6. The workflow of the research methodology is illustrated in Figure 4. First, we performed exploratory data analysis to answer Research Question 1, how has COVID-19 shaped human mobility? We analyzed the temporal change of Travel Volume from the pre-pandemic period to the third wave, and Residence Time from the third wave to the post-pandemic period. Second, we utilized linear regression models to answer research question 2, how has the relationship between human mobility and socioeconomic indicators evolved over the course of the pandemic? In each regression model, we took a mobility indicator as the dependent variable and the three socioeconomic indicators as the independent variables. In addition, the Policy Stringency is the confounder in the regression models, considering the potential influence of the policies on people's travel choices. To analyze the evolution of the relationship, we fitted the regression models every week. The mobility indicators change by week. The socioeconomic indicators are constant over time. We use Travel Volume to explore the evolution of the relationship from the pre-pandemic period to the third wave, and Residence time from the third wave to the post-pandemic period.

[1]Contiguous US consists of the 48 adjoining US states and the District of Columbia. https://www.nrel.gov/comm-standards/editorial/contiguous-united-states-continental-united-states-and-conus.html

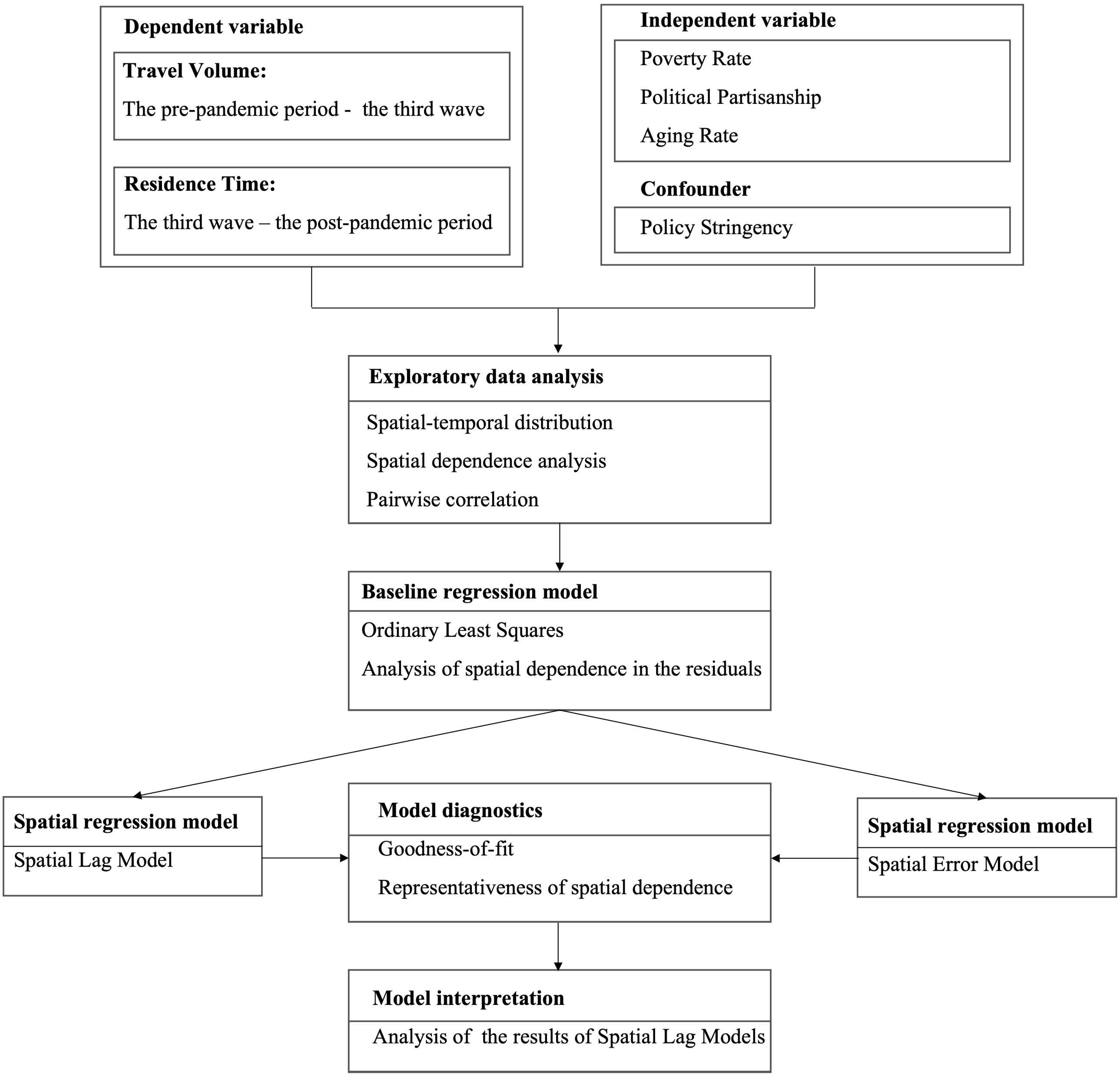


Figure 4: The workflow of research methodology.

Table 6: Summary of variables.

Variable	Time span	Temporal granularity	Counties	Data Source
(Dependent variable)				
Travel Volume	Pre-pandemic, first - third wave (Jan. 2019 - Apr. 2021)	week	3018	[Kang et al., 2020]
Residence Time	Third - fifth wave, post-pandemic (Oct. 2020 - Oct. 2022)	week	300	[Google, 2022]
(Independent variable)				
Poverty Rate	-	-	3108	[Census, 2022]
Political Partisanship	-	-	3108	[Census, 2022]
Aging Rate	-	-	3108	[Census, 2022]
(Confounder)				
Policy Stringency	Jan. 2020 - Oct. 2022	week	3018	[Hale et al., 2021]

5.1 Exploratory Data Analysis

We first explored the temporal change of Travel Volume from the pre-pandemic to the third wave. Then we analyzed the characteristics of Travel Volume in each wave. To minimize the influence of seasonal and holiday patterns, we measured the deviation of Travel Volume during the pandemic from the same weeks during the pre-pandemic period. Second, we took a snapshot of Travel Volume each period to investigate the spatial distribution (the week of March 24th 2019, March 22nd 2020, August 16th 2020, March 21st 2021). Furthermore, Moran's I statistics were used to gauge the spatial dependence of Travel Volume. It is a starting point to leverage spatial regression models in Section 5.2.2 and Section 5.2.3. Third, the pairwise correlation coefficients between Travel Volume and each socioeconomic indicator were calculated every week, as establishing correlation is a prerequisite for linear regression, which we conducted in Section 5.2.

Regarding the alternative mobility indicator, Residence Time, we investigated the temporal change from the third wave to the post-pandemic period due to data availability. To investigate the characteristics of each period and eliminate the influence of seasonal and holiday patterns, the fifth wave was compared with the third wave, and the post-pandemic period was compared with the fourth wave, since they cover the same months of a year. Then, we also conducted a pairwise correlation analysis between Residence Time and each socioeconomic indicator.

5.2 Linear Regression Models

To examine the relationship between mobility and socioeconomic indicators, we leveraged three multiple linear regression models, namely, Ordinary Least Square (OLS), Spatial Lag Model (SLM), and Spatial Error Model (SEM). In these regression models, the dependent variable is a mobility indicator, either Travel Volume or Residence Time. The independent variables are the three socioeconomic indicators, Poverty Rate, Political Partisanship, and Aging Rate. The confounder is Policy Stringency. To analyze the temporal evolution of the relationship, we fitted the regression models each week over the study period.

OLS serves as a benchmark. It is often fitted before considering more advanced regression models. OLS might overfit when there is spatial dependence in the data. When spatial dependence is diagnosed among the residuals of OLS, as well as the human mobility and socioeconomic indicators, the solution is to expand OLS to spatial regression [Chi and Zhu, 2019]. SLM and SEM are two basic types of spatial regression models. In terms of how to choose between them, there are two main approaches, the data-driven approach [Voss and Chi, 2006] and the theory-based approach [Doreian, 1980]. In this thesis, we conducted a data-driven approach by implementing both SLM and SEM based on seven different weight matrices (Section 5.2.4). We then selected the one that has better goodness-of-fit R^2, and better represents the spatial dependence using the spatial coefficients ρ.

5.2.1 Ordinary Least Squares (OLS)

OLS estimates the regression coefficient by minimizing the residual sum of squares, the sum of the squared difference between the observed and estimated values of the dependent variable. OLS assumes that the observations are independent and does not consider spatial dependence. It has the general form as in Equation 1. In the context of this thesis, it can be expressed with matrix notation in Equation 2.

$$Y = X\beta + S\gamma + \epsilon \tag{1}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & x_{n3} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} s_1 \\ s_2 \\ \dots \\ s_n \end{bmatrix} \gamma + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \dots \\ \epsilon_n \end{bmatrix} \tag{2}$$

where:

- Y is an n-by-1 vector comprising a mobility indicator of n counties in a week (n = 3108 of the Travel Volume; n =300 of the Residence Time).
- X is an n-by-4 matrix, the first column is comprised of constant 1 for the intercept. The rest of the matrix represents the 3 socioeconomic indicators (Poverty Rate, Political Partisanship, Aging Rate) of n counties.
- β is a 4-by-1 vector, β_0 is the intercept and β_{1-3} is the regression coefficients.
- S is an n-by-1 vector, representing the confounder, Policy Stringency, in n counties in a week.
- γ is the coefficient associated with the confounder.

- ϵ is an n-by-1 vector of the residuals.

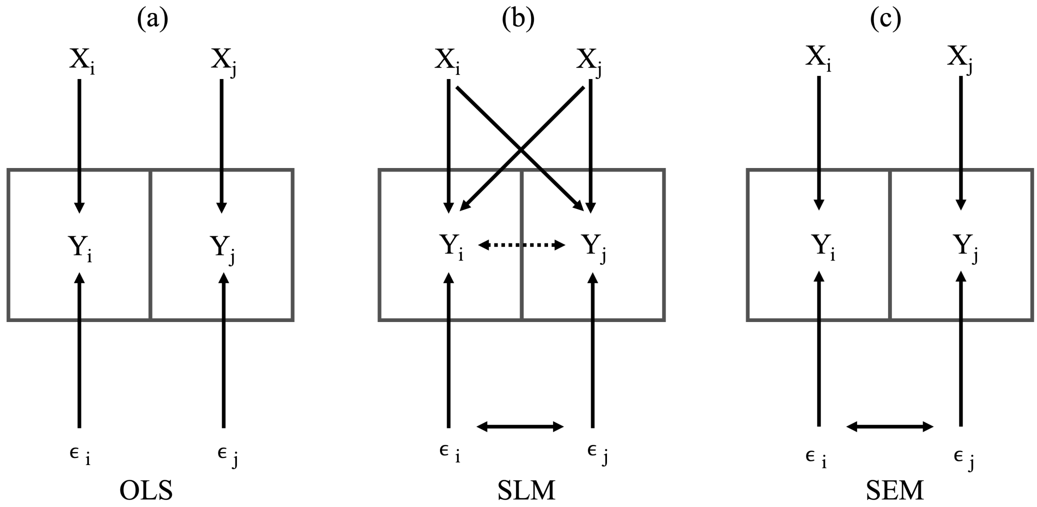


Figure 5: Comparison among OLS, SLM, and SEM [Yum, 2022].

5.2.2 Spatial Lag Model (SLM)

A spatial linear regression model extends an OLS by incorporating spatial dependence [Anselin and Griffith, 1988]. SLM accounts for spatial lag dependence (Figure 5.b). For instance, the Travel Volume y_i of a county i, is associated with the Travel Volume, y_j, of a neighboring county j. Besides, in county i, the Travel Volume is associated with not only the Poverty Rate, Political Partisanship, and Aging Rate in the county, but also with those in the neighboring counties. SLM has a general form as in Equation 3. In the context of this thesis, it can be expressed with matrix notation in Equation 4.

$$Y = X\beta + \rho WY + S\gamma + \epsilon \tag{3}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & x_{n3} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \rho \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \dots & \dots & \dots & \dots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} + \begin{bmatrix} s_1 \\ s_2 \\ \dots \\ s_n \end{bmatrix} \gamma + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \dots \\ \epsilon_n \end{bmatrix} \tag{4}$$

where:

- ρ is a scalar of the spatial dependence parameter.

- W is an n-by-n spatial weight matrix.

5.2.3 Spatial Error Model (SEM)

SEM accounts for spatial error dependence, the residual of county i, is correlated with the residual in county j (Figure 5.c). SEM is constructed by decomposing the residual in OLS into a residual and an associated spatially lagged residual [Anselin and Bera, 1998]. SEM has a general form as in Equation 5. In the context of this thesis, it can be expressed with matrix notation in Equation 6.

$$Y = X\beta + \rho WU + S\gamma + \epsilon \tag{5}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} \\ 1 & x_{21} & x_{22} & x_{23} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_{n1} & x_{n2} & x_{n3} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \rho \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \cdots \\ u_n \end{bmatrix} + \begin{bmatrix} s_1 \\ s_2 \\ \cdots \\ s_n \end{bmatrix} \gamma + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \cdots \\ \epsilon_n \end{bmatrix} \tag{6}$$

where:

- ρ is a scalar of the spatial dependence parameter.
- U is an n-by-1 vector of the spatial component of the residuals.

5.2.4 Identifying Spatial Weight Matrix

A spatial weights matrix is an n-by-n positive symmetric matrix W with element w_{ij} at location i to location j for n locations (Equation 4 and Equation 6). There are many types of spatial weight matrices with different specifications. Different spatial weight matrices result in distinguished parameter estimates [Zhou and Lin, 2008]. One approach to choose a better spatial weight matrix approach is to select the one that encompasses the highest spatial dependence of the dependent variable in combination with the high statistical significance [Chi and Zhu, 2008].

We implemented SLM and SEM based on seven spatial weight matrices. The illustrations of 100-kilometer kernel distance, 8-nearest neighbors, and queen contiguity weight matrices can be found in Figure 6 and Figure 7.

- Distance-based spatial weight matrix: It assigns weights to neighbors based on a kernel function that gives higher weights to closer neighbors. We utilized the Gaussian kernel function with bandwidth, 50km, 100km, 200 km.
- K-nearest Neighbor (KNN) spatial weight matrix: For each county, it contains the k closed neighboring counties. We changed k from 8 to 16 (k = 8, 12, 16).
- Contiguity-based spatial weight matrix: We utilized the queen contiguity matrix. Counties are considered neighbors if they share an edge or a vertex. A neighbour of a county is assigned the weight 1. Otherwise, it is 0.

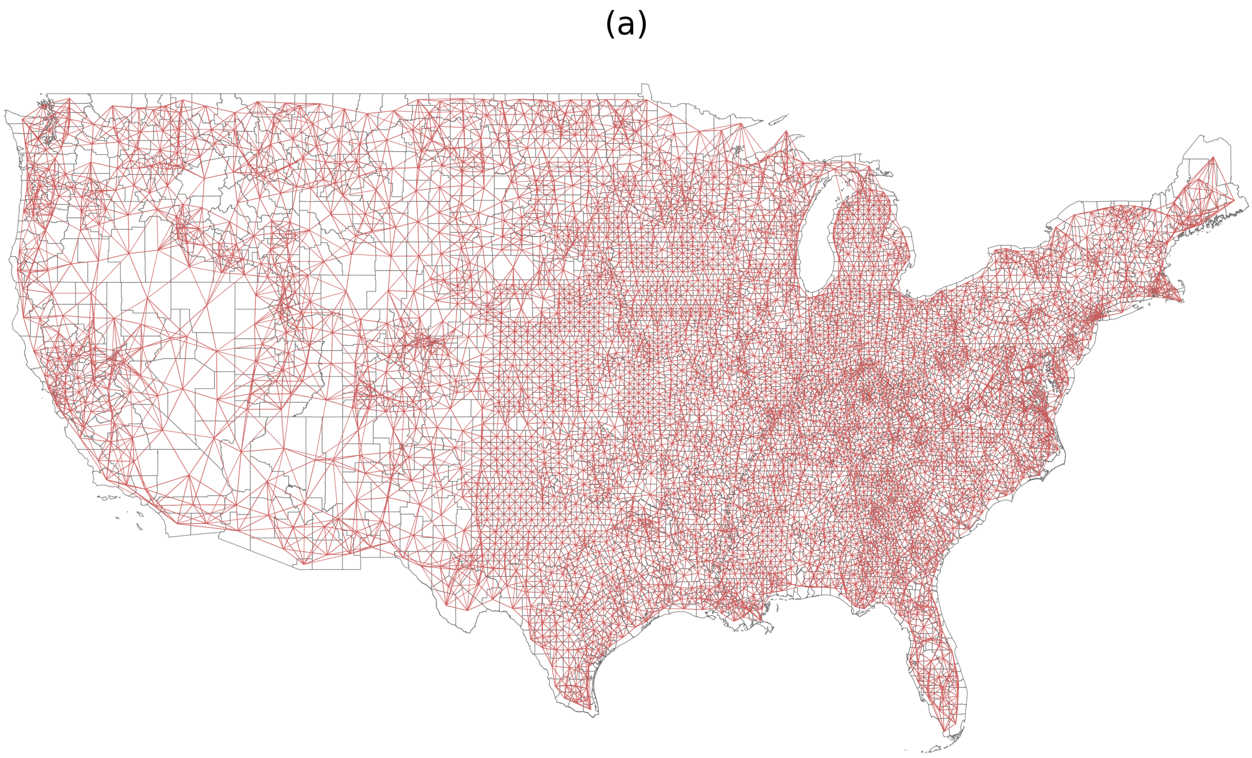
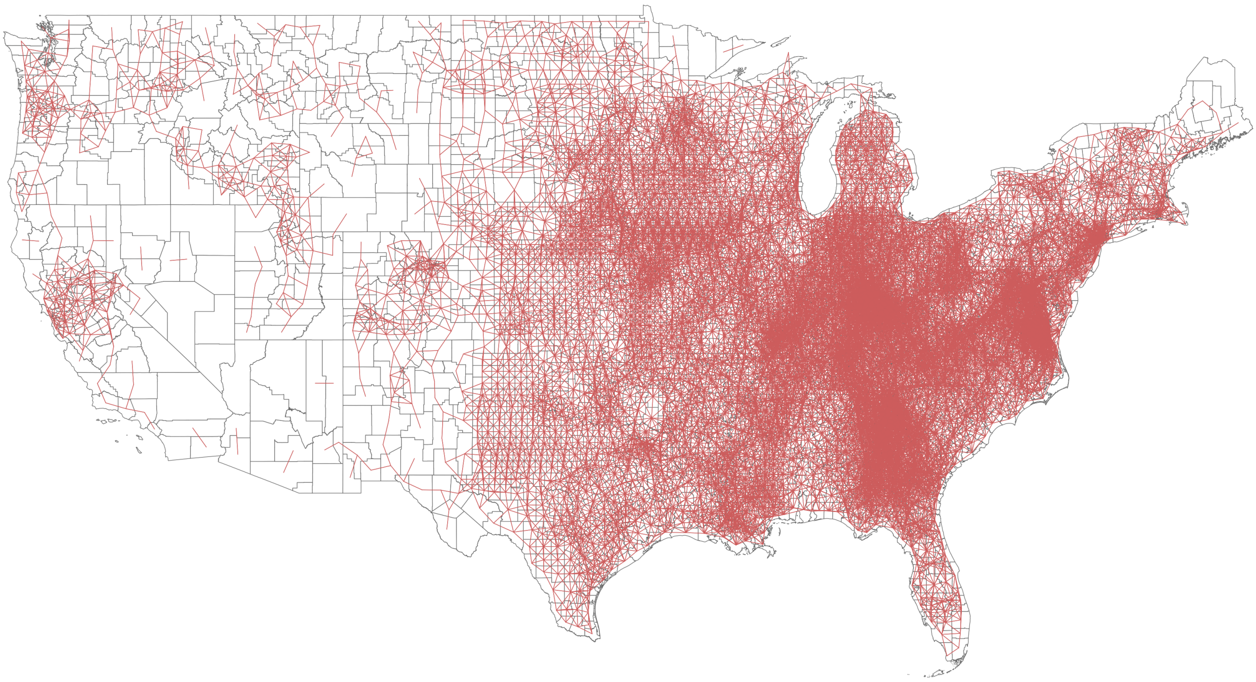


Figure 6: Spatial weight matrices. (a) 100-kilometer kernel density weight matrix.

(b)



(c)

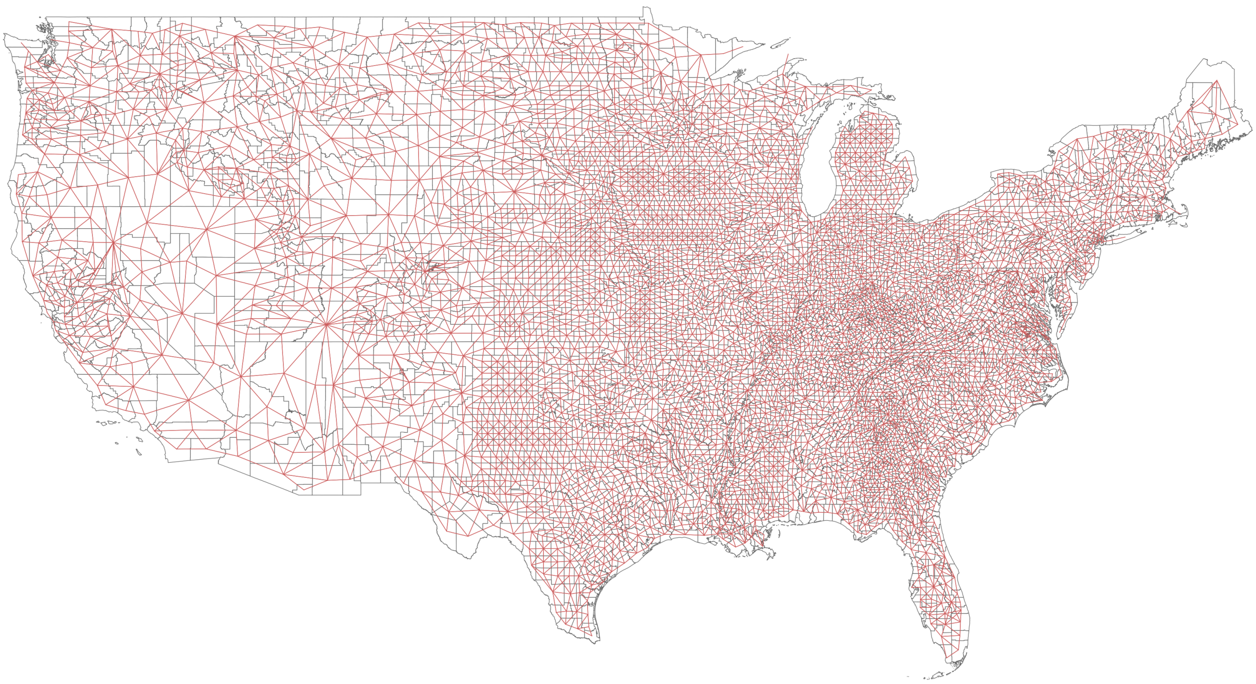


Figure 7: Spatial weight matrices. (b) 8-nearest neighbors spatial weight matrix; (c) Queen contiguity spatial weight matrix.

Chapter 6

Results

6.1 Exploratory Data Analysis

6.1.1 The Result of Travel Volume

Figure 8.a shows the temporal change of Travel Volume. Note that Travel Volume measures weekly mobility. Overall, Travel Volume has larger values but a smaller variability in 2019 than in 2020. The annual average Travel Volume in all the 3018 counties is 0.95 in 2019 and 0.92 in 2020. The standard deviation is 0.002 and 0.004, respectively.

In 2019, Travel Volume exhibits seasonal and holiday patterns. The lowest point occurs in January, while the highest point is in October. From January to October, Travel Volume shows a gradual increase, followed by a decline from October to December. Additionally, throughout 2019, Travel Volume exhibits valleys during weeks linked to holidays, with varying levels of decrease among different holidays. For instance, Thanksgiving experiences a greater reduction compared to Labor Day. In 2020, Travel Volume exhibits a different pattern. Both the lowest point and the highest point are in March. Travel Volume declines steeply starting from the second week of March 2020. This is potentially associated with the declaration of national emergency in response to COVID-19 on March 13th 2020 and the implementation of state-wise stay-at-home orders starting from March 19th 2020. Travel Volume starts to rebound in the first week of April and reaches an average of 0.97 in the first week of June. From June to October, it is relatively stable and remains near an average of 0.95. Travel Volume decreases again from October to December 2020. Due to data availability, we only analyzed Travel Volume from January to April in 2021. It is relatively stable before the third week of February. However, the fourth week of February experiences a sharp increase, from an average of 0.89 to 1.08. This abrupt increase might be associated with the nationwide distribution of COVID-19 vaccines.

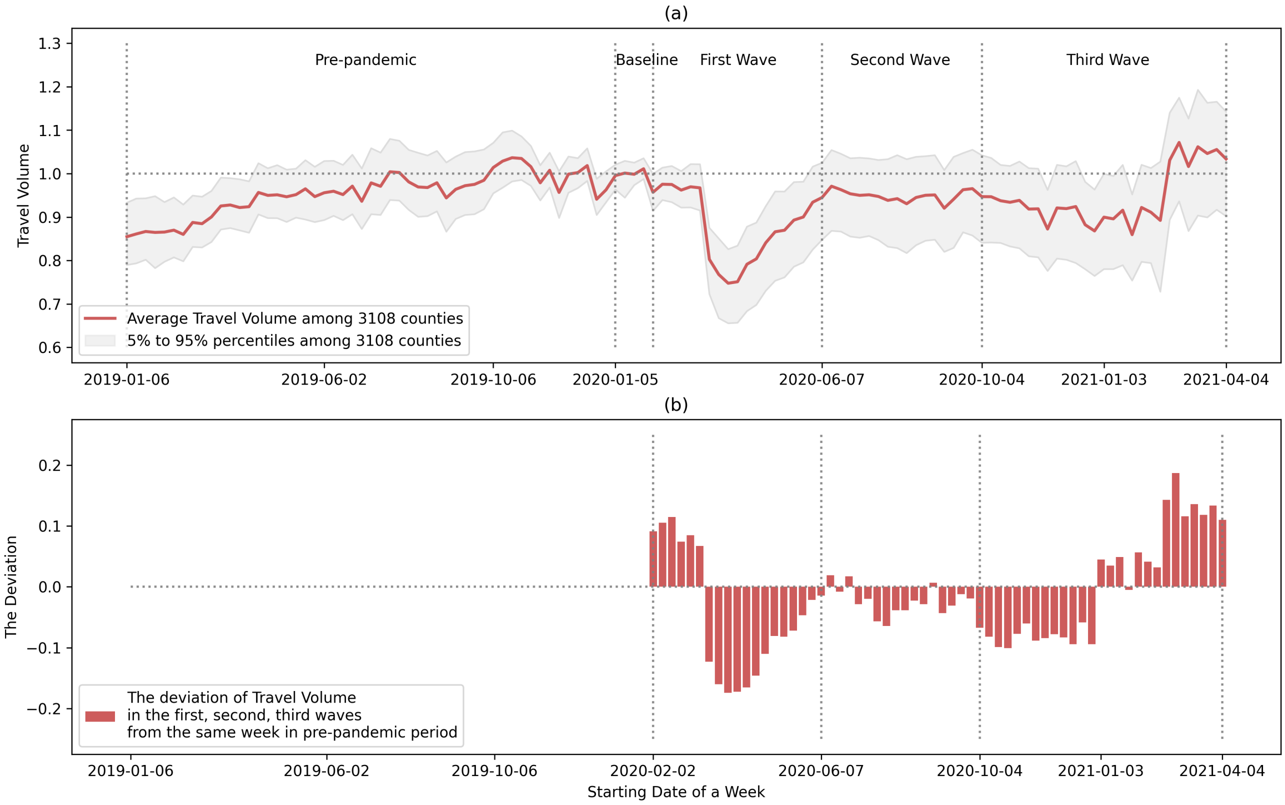


Figure 8: (a) The temporal change of Travel Volume from the week of January 6th 2019 to April 4th 2021. The red line represents the average of all the 3108 counties. The grey buffer represents 5% to 95% percentiles. (b) The deviation of Travel Volume each week during first three waves from the same week in the pre-pandemic period.

During the initial three waves, we calculated the deviation of Travel Volume from the corresponding week in the pre-pandemic period to eliminate the effect of seasonal and holiday patterns (Figure 8.b) A positive deviation indicates greater mobility compared with the pre-pandemic level (Figure 8.b). During the first wave, the deviations are initially positive and abruptly switch to negative values. It indicates that people initially travel more than the pre-pandemic level but abruptly reduce traveling. In addition, the deviations exhibit large variability, suggesting a notable distinct travel pattern compared with the pre-pandemic period. In contrast, during the second wave, the deviations are relatively small and stable, showing a similar pattern as the pre-pandemic period. In terms of the third wave, the deviations are initially negative with a magnitude larger than the second wave but smaller than the first wave. It turns positive afterwards, indicating larger Travel Volume than the pre-pandemic level.

We took one snapshot of each period from the pre-pandemic period to the third wave to explore the spatial distribution of Travel Volume (Figure 9). Overall, the middle part of the

US exhibits larger Travel Volume than the west and east coasts. The standard deviation is the smallest in the week of March 24th 2019 and highest in March 21st 2021 (Table 7). Furthermore, we calculated the Global Moran's I statistics and found significant spatial dependence in Travel Volume every week (Table 7). The positive z-scores suggest spatial clustering. If a county has a large Travel Volume, its neighboring counties are likely to have a large Travel Volume as well. The results suggest a potential need for controlling spatial dependence when fitting regression models.

Table 7: Spatial dependence analysis of Travel Volume in the four selected weeks.

Starting date of a week	**March 24th 2019**	**March 22nd 2020**	**August 16th 2020**	**March 21st 2021**
Mean	0.93	0.77	0.94	1.06
Standard deviation	0.04	0.05	0.07	0.09
Moran's I statistics	0.34***	0.60***	0.52***	0.63***
Z-score of Moran's I statistics	19.93	64.35	66.58	84.13

*** denotes $p < 0.001$.

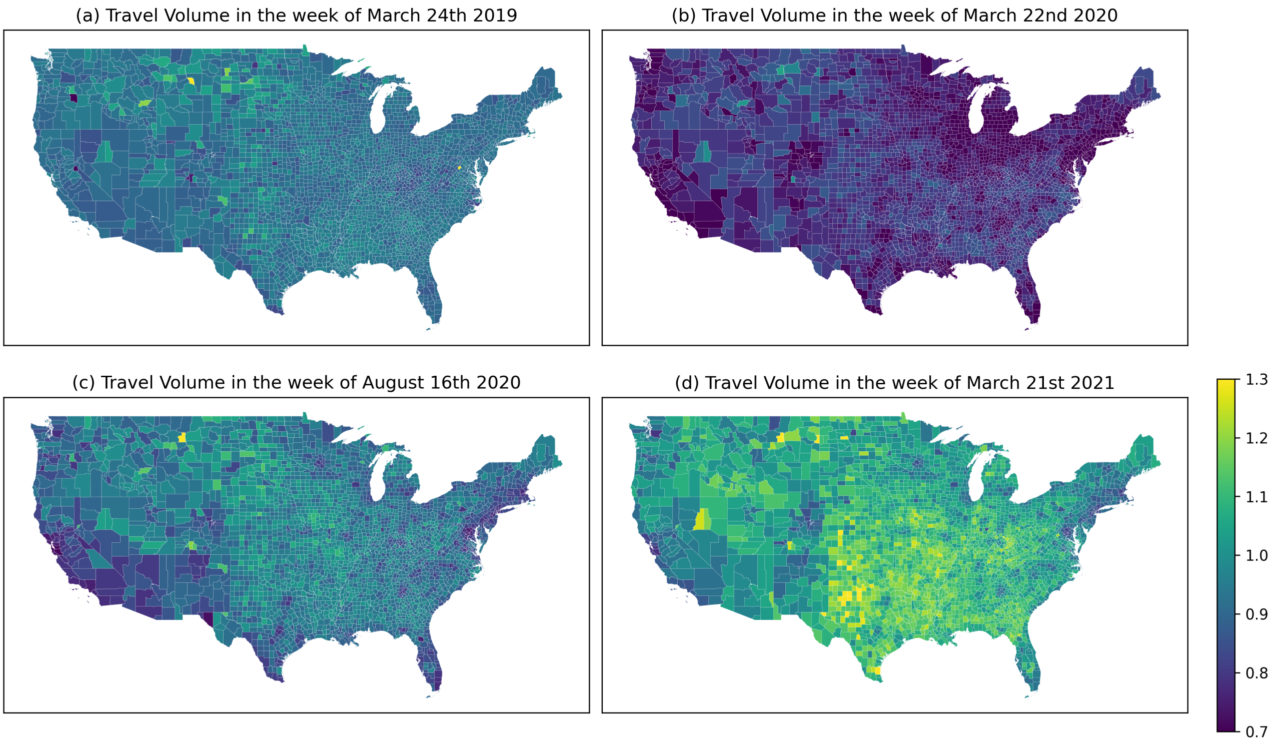


Figure 9: The spatial distribution of Travel Volume in the week of (a) March 24th 2019, (b) March 22nd 2020, (c) August 16th 2020, (d) March 21st 2021.

To have an overview of the correlation before fitting regression models, we calculated pairwise correlation coefficients between Travel Volume and each socioeconomic indicator per week (Figure 10). Travel Volume displays correlations with all three socioeconomic indicators and the correlations change over time. Linear regression models were fitted to have a further understanding of the relationship.

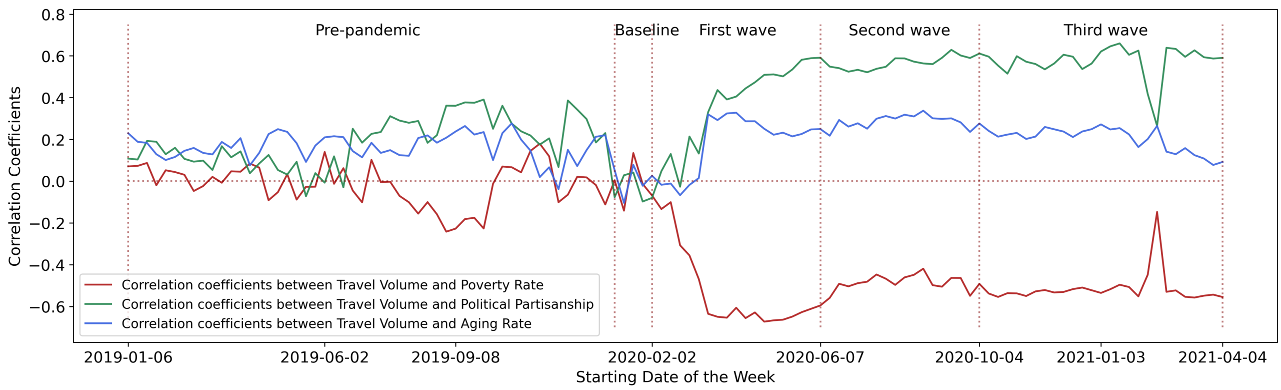


Figure 10: The pairwise correlation coefficients between Travel Volume and each socioeconomic indicator every week from the pre-pandemic period to the third wave.

6.1.2 The Result of Residence Time

Residence Time is from the third wave to the fifth wave, and the post-pandemic period due to data availability (Figure 11.a). It is always larger than 1, suggesting that people had been spending more time at places of residence compared to the baseline. The average Residence Time lies between 1.04 and 1.12 during the third wave, 1.02 to 1.05 during the fourth wave, 1.02 to 1.11 during the fifth wave, and 1.01 to 1.05 during the post-pandemic period. Spikes can be observed during holidays, such as Thanksgiving and Christmas.

To mitigate the influence of seasonal and holiday patterns, we calculated the deviation of Travel Volume each week in the fifth wave from the third wave, and the deviation of the post-pandemic period from the fourth wave. This analysis was performed because the third and fifth waves correspond to the same weeks of a year, as do the fourth wave and the post-pandemic period. The deviations are mostly negative (Figure 11.b). The trend suggests that people gradually spent less time at places of residence when transitioning from the third wave to the post-pandemic period.

Figure 12 shows the pairwise correlation coefficients between Residence Time and each socioeconomic indicator per week. In general, Residence Time shows a correlation with all the socioeconomic indicators, and the correlations change over time. The changes are relatively less in the correlation with Aging Rate. We further examined the relationship in depth using linear regression models.

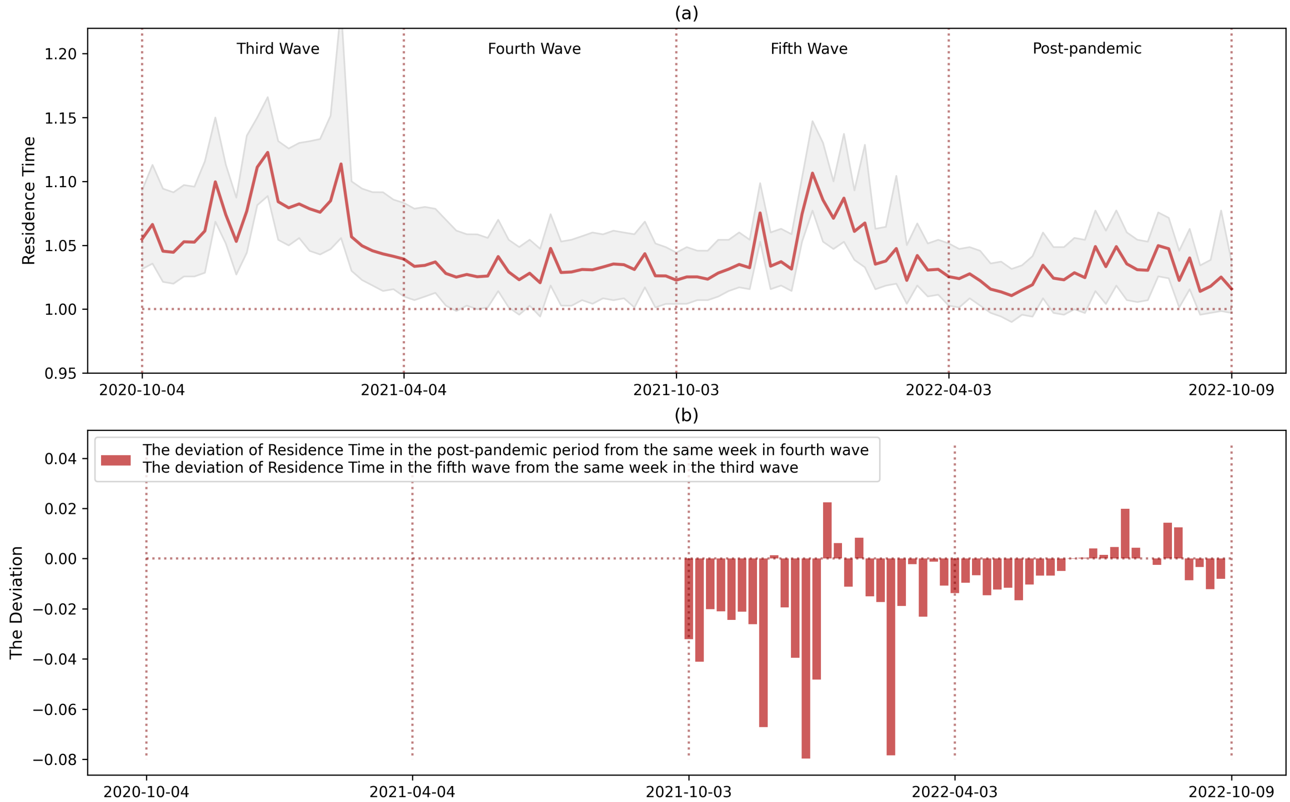


Figure 11: (a) The temporal change of Residence Time from the third wave to the post-pandemic period. (b) The deviation of Travel Volume each week in the fifth wave from the third wave, the deviation of the post-pandemic period from the fourth wave.

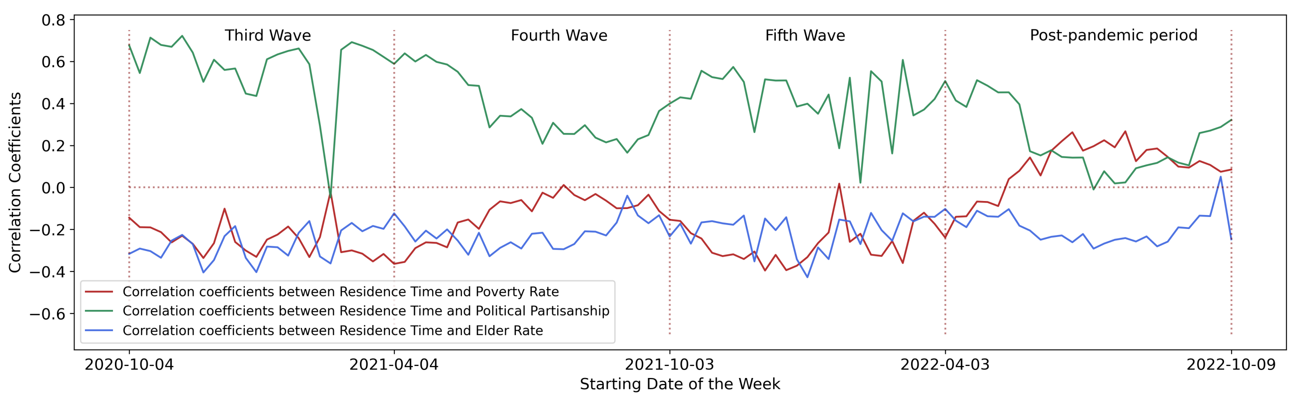


Figure 12: The pairwise correlation coefficients between Residence Time and each socioeconomic indicator every week from the third wave to the post-pandemic period.

6.2 Linear Regression Models

6.2.1 Model Diagnostics

In addition to examining the spatial dependence in the mobility indicators (Section 6.1.1), we also fitted the OLS regression model per week and tested whether the residuals of each model are spatially dependent. Note that we only used the main mobility indicator, Travel Volume, to perform model diagnostics due to large missing data in the alternative mobility indicator, Residence Time. As a result, residuals of OLS exhibit statistically significant spatial dependence every week. Furthermore, we computed Lagrange Multiplier (LM) tests and Robust LM tests, which diagnose the appropriateness of spatial regression with respect to spatial lag and spatial error dependence [Anselin and Griffith, 1988; Baltagi and Yang, 2013]. The tests suggest that OLS residuals exhibit spatial lag and spatial error dependence.

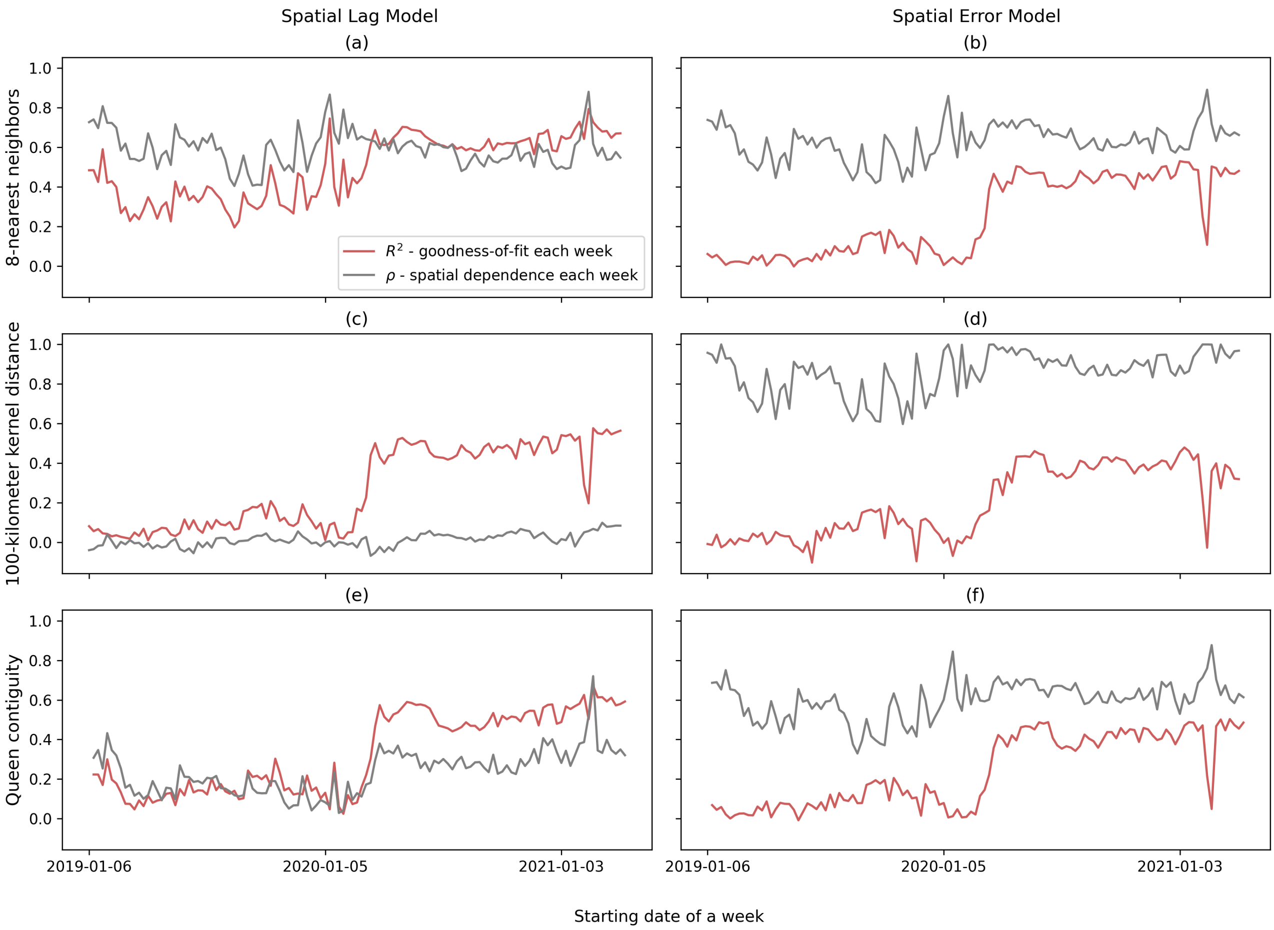


Figure 13: The goodness-of-fit R^2 and spatial dependence ρ of: (a) SLM based on 8-nearest neighbors spatial weight matrix; (b) SEM based on 8-nearest neighbors; (c) SLM based on 100-km kernel distance; (d) SEM based on 100-km kernel distance; (e) SLM based on the queen contiguity; (f) SEM based on the queen contiguity spatial weight matrix.

All the analysis results indicate the necessity of spatial regression models. Thereafter, we implemented spatial lag model (SLM) and spatial error model (SEM), based on seven different spatial weight matrices (Subsection 5.2.4). Goodness-of-fit R^2 and spatial dependence coefficient ρ are utilized to select the better model. We only visualized the results of three weight matrices in Figure 13, as the three have the best performance among each type of spatial weight matrix. In general, SLM has better goodness-of-fit than SEM. Therefore, we chose SLM over SEM. In addition, the result of SLM based on the 8-nearest neighbors spatial weight matrix has higher spatial dependence coefficients than the 100-km kernel distance and the queen contiguity spatial weight matrices. Consequently, we selected SLM based on the 8-nearest neighbors weight matrix and analyzed the results in detail.

6.2.2 The Result of Travel Volume

We fitted SLM every week to analyze the evolution of the relationship between human mobility and socioeconomic indicators. The dependent variable is Travel Volume, which changes weekly. The independent variables are Poverty Rate, Political Partisanship, and Aging Rate, all of which are constant over time. The goodness-of-fit R^2 changes over the study period. It is better during the pandemic period with R^2 near 0.7 (Figure 14.b). The spatial dependence coefficients fluctuate between 0.4 and 0.9. It suggests spatial clustering. When a county has a large Travel Volume, its neighboring counties also tend to have a large Travel Volume. Figure 14 illustrates the time-varying regression coefficients of the three independent variables. The regression coefficients exhibited frequent fluctuation. Furthermore, there is a salient difference between the pre-pandemic and pandemic periods, especially for the coefficients of Poverty Rate and Political Partisanship.

The relationship between Poverty Rate and Travel Volume experiences a significant change moving from the pre-pandemic to the pandemic period. During 52 weeks of the pre-pandemic period, the regression coefficients are significant and negative in 32 weeks, fluctuating between -0.05 and 0. The Trave Volume decreases by as large as 0.05 unit with Poverty Rate increasing by 1 unit. Besides, the relationship is not statistically significant in 15 weeks ($p > 0.05$). During 66 weeks of the pandemic period, the regression coefficients sharply flip over the x-axis. The coefficients are always positive and significant, ranging from 0.05 to 0.19, except for the week of February 14th 2021. It suggests a reversed and stronger relationship between Poverty Rate and Travel Volume compared with the pre-pandemic period. Instead of decreasing by 0.5 unit,

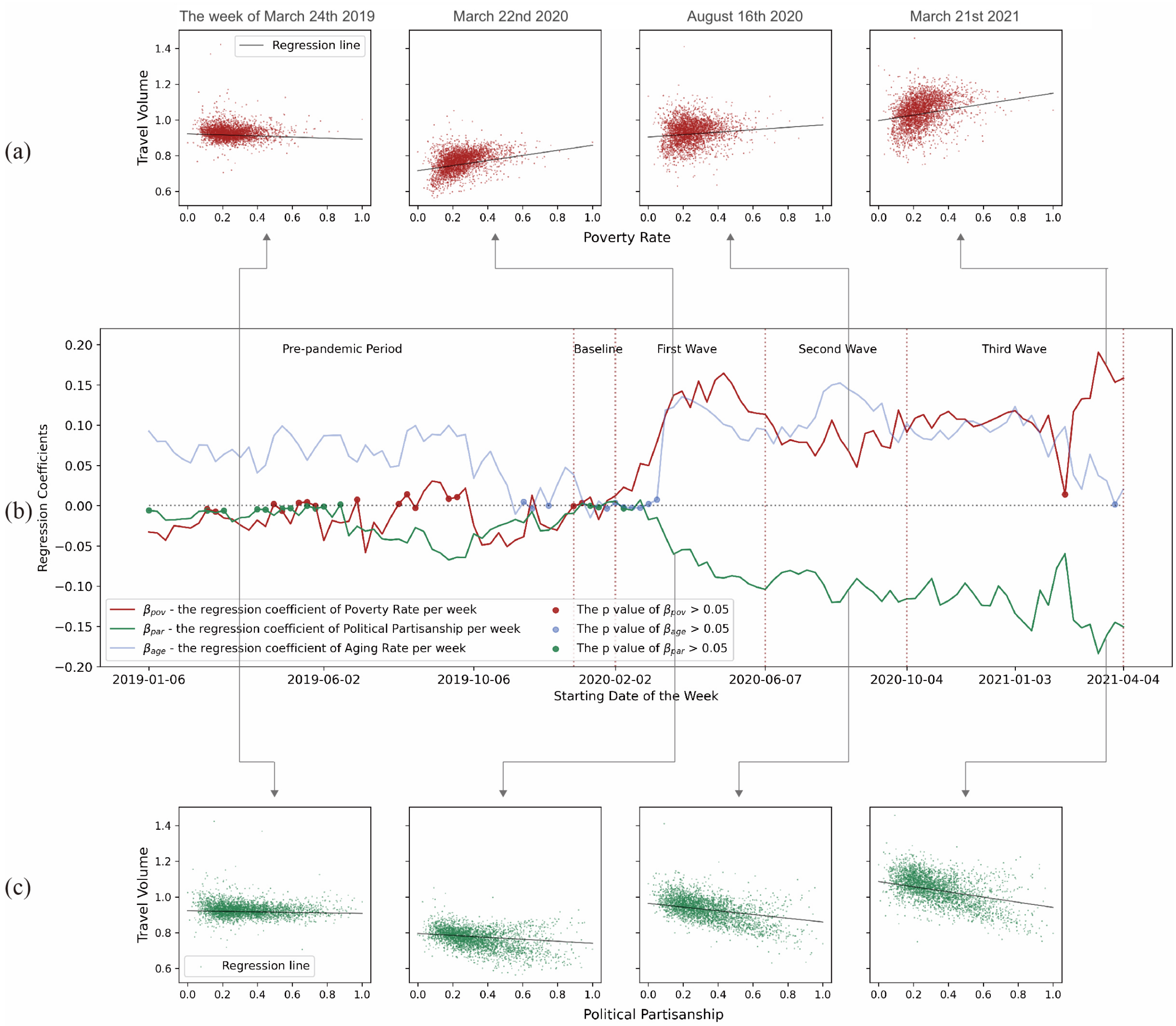


Figure 14: The evolution of the relationship between Travel Volume and socioeconomic indicators. (a) Four snapshots of the relationship between Poverty Rate and Travel Volume in the week of March 24th 2019, March 22nd 2020, August 16th 2020, and March 21st 2021. (b) The time varying regression coefficients β of Poverty Rate, Political Partisanship, and Aging Rate every week. (c) Four snapshots of the relationship between Political Partisanship and Travel Volume.

Travel Volume increases by as large as 0.19 unit with Poverty Rate increasing by 1 unit. In addition, the relationship exhibits different characteristics from the first to the third wave. It becomes relatively weaker during the second wave but stronger again during the third wave (Figure 14.a).

The relationship between Political Partisanship and Travel Volume also exhibits dramatic changes. During the pre-pandemic period, the coefficients are always negative, fluctuating between -0.07 and 0. Furthermore, 10 out of 52 weeks show statistical insignificance of the relationship. It indicates that the increase of 1 unit in Political Partisanship is associated with as large as the decrease of 0.07 unit or as small as no change in Travel Volume. During the pandemic period, the relationship is still negative but becomes stronger. The coefficients lie between -0.01 and -0.11 during the first wave, -0.08 to 0.12 during the second wave, -0.05 to -0.19 during the third wave. Political Partisanship shows an increasingly stronger relationship with human mobility during the pandemic period (Figure 14.c).

Unlike Poverty Rate and Political Partisanship, Aging Rate shows only small changes in the relationship with Travel Volume. The regression coefficients before October of the pre-pandemic period are significant and remain close to 0.06. It suggests that the increase of Aging Rate in 1 unit is associated with the increase of 0.06 unit in Travel Volume. The relationship weakens from October to December 2019. During the pandemic periods, the association is always significant except for the week of March 28th 2021. It becomes slightly stronger in the beginning and then drops to the pre-pandemic level again during both the first and the second waves. Besides, the third wave observes a weaker association during the late stage.

6.2.3 The Result of Residence Time

To further analyze the evolution of the relationship between human mobility and socioeconomic indicators, we utilized the alternative mobility indicator, Residence Time. It covers the third wave to the post-pandemic period. We fitted SLM per week based on the 8-nearest neighbors weight matrix, taking Residence Time as the dependent variable. The model each week has high goodness-of-fit with $R^2 > 0.9$ (Figure 15.b). In addition, ρ remains near 0, which appropriately accounts for the spatial dependence as only 300 counties have complete data of Residence Time, and these counties spread sparsely in space.

Figure 15.a illustrates the time-varying correlation coefficients. Compared with the main mobility indicator (Travel Volume), Residence Time exhibits the opposite relationship with the

three socioeconomic indicators. In general, higher Poverty Rate and Aging Rate are associated with smaller Residence Time. On the contrary, higher Political Partisanship is associated with larger Residence Time. Furthermore, we observe an increasingly weak association from the third wave to the post-pandemic period. The regression coefficients of Political Partisanship reduce from 0.1 to near 0.01. In addition, the coefficients are statistically insignificant in only 2 of the 83 weeks from the third to the fifth wave. However, the coefficients are statistically insignificant in 13 out of 21 weeks during the post-pandemic period. Regarding Poverty Rate, the coefficients move near 0 as well. The coefficients of Aging Rate show relatively less change.

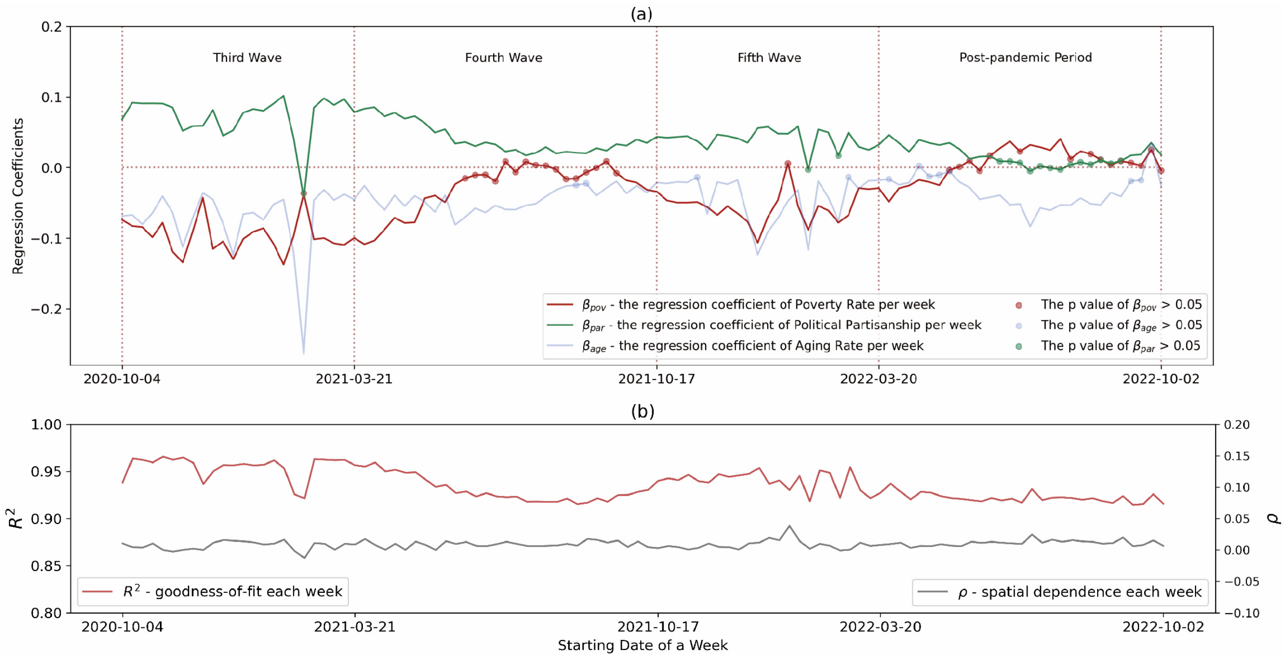


Figure 15: The evolution of the relationship between Residence Time and socioeconomic indicators. (a) The time varying regression coefficients β of Poverty Rate, Political Partisanship, and Aging Rate every week. (b) The goodness-of-fit R^2 and spatial dependence coefficient ρ.

.

Chapter 7

Discussion

7.1 Temporal Changes in Human Mobility

To answer Research Question 1, **how does COVID-19 shape human mobility?** We leveraged two mobility indicators to model aggregated human mobility and analyzed their temporal changes. Due to data availability, the two mobility indicators do not cover the same period. The main mobility indicator, Travel Volume, is from January 2019 to April 2021, including the pre-pandemic period to the third wave. The alternative mobility indicator, Residence Time, is from October 2020 to October 2022, covering the third wave to the post-pandemic period.

During the pre-pandemic period, the temporary change in Travel Volume exhibits seasonal, and holiday patterns, which are in line with existing literature [Kraemer et al., 2020; Dobler et al., 2021]. People generally travel more in summer and less in winter. Holidays observe a decrease in mobility and the magnitude of the decrease varies (Figure 8.a). For example, people reduce traveling larger in the week of Thanksgiving than in the week of Martin Luther King Jr. Day. In addition to these patterns, Travel Volume during the pandemic is significantly influenced by events related to COVID-19. For instance, during the very early stage of the first wave, people travel more than the pre-pandemic level. However, they abruptly reduced traveling and travel significantly less than the pre-pandemic level (Figure 8.b) with the declaration of national emergency by President Donald Trump, and the implementation of stay-at-home orders and social distancing measures by many states in March 2020. Even though people start to increase traveling very soon, it does not recover to the pre-pandemic level. It is not until the last week of February 2021 that the Travel Volume surpasses the pre-pandemic level (Figure 8.b).

Besides, the increase is sharp and reached around 1.25 times of the pre-pandemic level. This might be related to the availability of three COVID-19 vaccines. Late February 2021 marks a significant expansion of the COVID-19 vaccine distribution across the US.

Travel Volume also exhibits different characteristics from the first to third waves. The first wave observes a V-shape change. People reduce traveling sharply during the early period and increase traveling very soon. However, the Travel Volume does not recover to the pre-pandemic level. The change in Travel Volume in the second wave is rather small. Even though it is still lower than the pre-pandemic level, the change is stable and almost follows a horizontal line. The third wave experiences a gentle V-shape change. People reduce traveling moderately in the beginning but increase traveling significantly in the later stage, which then surpasses the pre-pandemic level.

To have a further look at the temporal change of human mobility during the third wave to the post-pandemic period, we analyzed the alternative mobility indicator, Residence Time. To eliminate the seasonal and holiday patterns, we make comparisons between the third wave (October 2020 to April 2021) and the fifth wave (October 2021 to April 2022), the fourth wave (April to October 2021), and the post-pandemic period (April to October 2022), because they cover the same months of a year. In general, people spend less time at places of residence in the fifth wave than in the third wave, in the post-pandemic than in the fourth wave (Figure 11.b). In short, there is a decreasing trend of time spent at places of residence. It indicates that people are increasingly more mobile when moving to the post-pandemic time.

7.2 The Evolution of the Relationship

We further examined the evolution of the relationship between human mobility and socioeconomic indicators using spatial regression models. In the model, the dependent variable is a mobility indicator of a week (either Travel Volume or Residence Time), and the independent variables are Poverty Rate, Political Partisanship, and Aging Rate. The Policy Stringency is taken as a confounder, considering the influence of travel restriction policies on travel patterns. We fitted the model every week to study the evolution of the relationship and answer Research Question 2, **how does the relationship between human mobility and socioeconomic indicators evolve over the course of the pandemic?**

Poverty Rate. The relationship between Travel Volume and Poverty Rate exhibits a significant change (Figure 14). During the pre-pandemic period, people living in counties with higher Poverty Rate generally travel less. However, the relationship reverses with the outbreak of the pandemic. These people start traveling more. In addition, the relationship is stronger than in the pre-pandemic period. Even though the governments encouraged social distancing and staying at home to curb the spread of COVID-19, populations with worse economic status either travel more or do not reduce travel as much as others do. Our findings are in line with existing literature that lower-income populations had difficulties in mobility adaptation during the pandemic [Circella, 2020]. The poorer population consists of a higher percentage of workers in essential industries [Tai et al., 2021]. They might not have the option to work from home. Our finding not only highlights that reducing traveling is a luxury [Huang et al., 2022] but also stresses the inequalities brought by the pandemic.

Political Partisanship. The relationship between Travel Volume and Political Partisanship also changes significantly (Figure 14). During the pre-pandemic period, smaller Political Partisanship indicator values are generally associated with larger Travel Volume. It suggests that people living in Republican-leaning counties tend to travel more. It is worth noting that the association is overall very moderate except for the summer month. With the pandemic outbreak, the relationship abruptly becomes stronger and residents in Republican-leaning counties tend to travel more significantly. This result is in line with existing literature [Grossman et al., 2020]. Furthermore, the relationship also becomes increasingly stronger, moving from the first wave to the third wave. It highlights the role of Political Partisanship in shaping people's opinions and behavior. Specifically, the findings allow us to deepen our understanding of the effect of political affiliation on people's response to the pandemic.

Aging Rate. The change in the relationship between Aging Rate and Travel Volume is relatively smaller compared to that of Poverty Rate and Political Partisanship (Figure 14). During the pre-pandemic period, counties with a larger population over 65 years old generally observe a larger Travel Volume. Even though old people come across more problems in moving and walking freely, they tend to spend more time outdoors during the working week in the summer compared to young adults [Borecka et al., 2021]. Moving to the pandemic period, the relationship becomes slightly stronger. Even though old people are strongly recommended by the government to stay at home due to their higher risk of severe disease from COVID-19, counties with higher Aging Rate, on the contrary, exhibit even larger mobility. It is possible that some old adults still choose to go out to maintain their physical and mental health or they do not perceive the COVID-19 as a risk. It's worth noting that we cannot reach the conclusion of individuals, as the analysis is at county level.

To further explore the evolution of the relationship from the third wave to the post-pandemic period, we utilized the alternative mobility indicator, Residence Time. Compared with Travel Volume, it exhibits a reversed relationship with the socioeconomic indicators (Figure 15). People living in counties with higher Poverty Rate spend less time at places of residence from the third wave to the fifth wave. Moving to the post-pandemic period, they started to spend more time at places of residence. Then the association becomes insignificant, which shows that Poverty Rate no longer explains Residence Time. Furthermore, Republican-leaning counties exhibit shorter time at places of residence from the third to the fifth wave. However, the association becomes weaker and eventually insignificant moving to the post-pandemic period. The association between Aging Rate and Residence Time shows a moderate change. Counties with a larger population over 65 years old still exhibit shorter time at places of Residence during the post-pandemic period.

7.3 Limitations

This thesis has some limitations. First, the interpretations of the results are always at county level, as we used aggregated human mobility data. One should avoid ecological fallacy or aggregation bias, which assumes that an individual has the same attributes as its aggregated group. Coarser aggregation tends to show stronger associations among variables [Hastie et al., 2009]. However, it is at the cost of higher complexity of data collection and demand of computational power.

Second, the alternative mobility indicator, Residence Time, may under-represent certain counties. The source data of the main mobility indicator, Travel Volume, is complete for the 3018 counties in the Contiguous US. It covers 10% of the US population and has also proved to be aligned with the census data. However, in terms of Residence Time, we did not find sufficient information about the representativeness of its source data. In addition, due to large missing data in space and time, we only selected 300 counties. All three socioeconomic indicators of the 300 counties have the same distribution as the 3108 counties, which is proved by a non-parametric hypothesis test (Cramér-von-Mises-Test). However, the potential issue of representativeness in Residence Time should still be noted.

Third, the two mobility indicators do not cover the same period. Travel Volume is available from the pre-pandemic period to the third wave. Residence Time is from the third wave to the post-pandemic period. Therefore, we could not make any comparison between the pre-pandemic and the post-pandemic periods, as they are based on two different mobility indicators.

Fourth, we used the Policy Stringency as a confounder in the regression models. However, some unobserved confounders may also interact with human mobility. Furthermore, we lack knowledge of the time lag in the travel restriction policies. For example, the influence of a policy imposed at time t may have a stronger impact on mobility at time $t + \Delta t$. It is hardly practical to quantify the Δt accurately.

Chapter 8

Conclusion and Future Work

This thesis investigates the temporal change in human mobility and the evolution of the relationship between human mobility and socioeconomic indicators during the pre-pandemic, the pandemic and the post-pandemic time. We find that human mobility exhibits an abrupt decrease in March 2020, when many states started imposing travel restriction policies. Human mobility surpasses the pre-pandemic level in February 2021, when the three COVID-19 vaccinations started being distributed nationwide. Furthermore, the analysis of the time-varying relationship reveals the changing roles of socioeconomic factors in human mobility, which is not depicted in previous work.

Using Travel Volume, we analyzed how the relationship evolves from the pre-pandemic period to the third wave. First, our analysis reveals that people living in poorer counties travel less before the pandemic. However, the travel patterns begin to change with the outbreak of the pandemic. These people start traveling more and the association becomes notably stronger. Second, people living in Republican-leaning counties generally travel more during the pre-pandemic period. The association becomes significantly stronger with the outbreak. It continues to strengthen from the first to the third wave. Third, counties with larger populations over 65 years old exhibit higher mobility before the pandemic. The association not only remains but also strengthens slightly from the first wave to the third wave.

With Residence Time, we analyzed the evolution of the relationship from the third wave to the post-pandemic period. First, residents in pooer counties spend less time at places of residence from the third wave to the fifth wave, whereas the association undergoes a complete reversal and then becomes statistically insignificant during the post-pandemic period. Second, residents in Republican-leaning counties also tend to spend less time at places of residence from the third wave to the fifth wave. However, the association gradually weakens and eventually becomes insignificant as we transition into the post-pandemic period. Third, counties with larger elder populations do not display a notable change in the association with time at places of residence when moving from the third wave to the post-pandemic period.

The analysis of the relationships from the pre-pandemic to the third wave, and from the third wave to the post-pandemic period reveals the changing roles of socioeconomic factors in mobility. It stresses the social inequality brought about by the pandemic. Even though there are no more travel restrictions, the impact of COVID-19 will cast a long shadow into the future. Efforts should still be made to allocate health care and financial resources properly targeting different social groups during the post-pandemic time.

This thesis analyzed human mobility at county level. Future studies may consider finer spatial granularity. In addition, due to data availability, the two mobility indicators do not cover the same period of time, so we cannot make comparisons between the pre-pandemic and post-pandemic periods. Future studies might consider analyzing human mobility during the post-pandemic time and compare it with the pre-pandemic time if more data is available. Furthermore, unobserved confounders might exist and interact with our variables and influence our analysis results, as we only utilized Policy Stringency as a confounding factor in the regression models. Therefore, future studies should investigate other factors related to human mobility.

Bibliography

[Abdullah et al., 2020] Abdullah, M., Dias, C., Muley, D., and Shahin, M. (2020). Exploring the impacts of covid-19 on travel behavior and mode preferences. *Transportation research interdisciplinary perspectives*, 8:100255.

[Aktay et al., 2020] Aktay, A., Bavadekar, S., Cossoul, G., Davis, J., Desfontaines, D., Fabrikant, A., Gabrilovich, E., Gadepalli, K., Gipson, B., Guevara, M., et al. (2020). Google covid-19 community mobility reports: anonymization process description (version 1.1). *arXiv preprint arXiv:2004.04145*.

[Anselin and Bera, 1998] Anselin, L. and Bera, A. K. (1998). Introduction to spatial econometrics. *Handbook of applied economic statistics*, 237(5).

[Anselin and Griffith, 1988] Anselin, L. and Griffith, D. A. (1988). Do spatial effecfs really matter in regression analysis? *Papers in Regional Science*, 65(1):11–34.

[Asgari et al., 2013] Asgari, F., Gauthier, V., and Becker, M. (2013). A survey on human mobility and its applications. *arXiv preprint arXiv:1307.0814*.

[Balcan et al., 2009] Balcan, D., Colizza, V., Gonçalves, B., Hu, H., Ramasco, J. J., and Vespignani, A. (2009). Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the national academy of sciences*, 106(51):21484–21489.

[Baltagi and Yang, 2013] Baltagi, B. H. and Yang, Z. (2013). Heteroskedasticity and non-normality robust lm tests for spatial dependence. *Regional Science and Urban Economics*, 43(5):725–739.

[Bambra, 2022] Bambra, C. (2022). Pandemic inequalities: emerging infectious diseases and health equity. *International Journal for Equity in Health*, 21(1):6.

[Barbalat and Franck, 2022] Barbalat, G. and Franck, N. (2022). Association of republican partisanship with us citizens' mobility during the first period of the covid crisis. *Scientific Reports*, 12(1):8994.

[Barbosa et al., 2018] Barbosa, H., Barthelemy, M., Ghoshal, G., James, C. R., Lenormand, M., Louail, T., Menezes, R., Ramasco, J. J., Simini, F., and Tomasini, M. (2018). Human mobility: Models and applications. *Physics Reports*, 734:1–74.

[Bonaccorsi et al., 2020] Bonaccorsi, G., Pierri, F., Cinelli, M., Flori, A., Galeazzi, A., Porcelli, F., Schmidt, A. L., Valensise, C. M., Scala, A., Quattrociocchi, W., et al. (2020). Economic and social consequences of human mobility restrictions under covid-19. *Proceedings of the National Academy of Sciences*, 117(27):15530–15535.

[Bonotti et al., 2021] Bonotti, M., Zech, S. T., Bonotti, M., and Zech, S. T. (2021). The human, economic, social, and political costs of covid-19. *Recovering civility during COVID-19*, pages 1–36.

[Borecka et al., 2021] Borecka, O., Farrar, M. D., Osman, J. E., Rhodes, L. E., and Webb, A. R. (2021). Older adults who spend more time outdoors in summer and have higher dietary vitamin d than younger adults can present at least as high vitamin d status: a pilot study. *International Journal of Environmental Research and Public Health*, 18(7):3364.

[CDC, 2022] CDC (2022). Covid-19.

[Census, 2022] Census, U. (2022). United states census bureau american community survey.

[Chi and Zhu, 2008] Chi, G. and Zhu, J. (2008). Spatial regression models for demographic analysis. *Population Research and Policy Review*, 27:17–42.

[Chi and Zhu, 2019] Chi, G. and Zhu, J. (2019). *Spatial regression models for the social sciences*. SAGE publications.

[Circella, 2020] Circella, G. (2020). The covid-19 pandemic: What does it mean for mobility? what the temporary vs. longer-term impacts. *SCAG Modeling Task Force. Retrieved February*, 10:2022.

[Clinton et al., 2021] Clinton, J., Cohen, J., Lapinski, J., and Trussler, M. (2021). Partisan pandemic: How partisanship and public health concerns affect individuals' social mobility during covid-19. *Science advances*, 7(2):eabd7204.

[Cucinotta and Vanelli, 2020] Cucinotta, D. and Vanelli, M. (2020). Who declares covid-19 a pandemic. *Acta bio medica: Atenei parmensis*, 91(1):157.

[Dobler et al., 2021] Dobler, G., Vani, J., and Dam, T. T. L. (2021). Patterns of urban foot traffic dynamics. *Computers, Environment and Urban Systems*, 89:101674.

[Doreian, 1980] Doreian, P. (1980). Linear models with spatially distributed data: Spatial disturbances or spatial effects? *Sociological Methods & Research*, 9(1):29–60.

[Dueñas et al., 2021] Dueñas, M., Campi, M., and Olmos, L. E. (2021). Changes in mobility and socioeconomic conditions during the covid-19 outbreak. *Humanities and Social Sciences Communications*, 8(1).

[Eagle and Pentland, 2006] Eagle, N. and Pentland, A. (2006). Reality mining: sensing complex social systems. *Personal and ubiquitous computing*, 10:255–268.

[Emanuel et al., 2022] Emanuel, E. J., Osterholm, M., and Gounder, C. R. (2022). A national strategy for the "new normal" of life with covid. *Jama*, 327(3):211–212.

[Felsenthal, 2020] Felsenthal, M. (2020). Covid-19 to plunge global economy into worst recession since world war ii. *World Bank*, 8:8.

[Frias-Martinez et al., 2012] Frias-Martinez, V., Virseda-Jerez, J., and Frias-Martinez, E. (2012). On the relation between socio-economic status and physical mobility. *Information Technology for Development*, 18(2):91–106.

[Garnier et al., 2021] Garnier, R., Benetka, J. R., Kraemer, J., and Bansal, S. (2021). Socioeconomic disparities in social distancing during the covid-19 pandemic in the united states: observational study. *Journal of medical Internet research*, 23(1):e24591.

[Genest et al., 2006] Genest, C., Quessy, J.-F., and Rémillard, B. (2006). Local efficiency of a cramér–von mises test of independence. *Journal of Multivariate Analysis*, 97(1):274–294.

[Gonzalez et al., 2008] Gonzalez, M. C., Hidalgo, C. A., and Barabasi, A.-L. (2008). Understanding individual human mobility patterns. *nature*, 453(7196):779–782.

[Google, 2022] Google (2022). Community mobility reports.

[Gregorian, 2021] Gregorian, R. (2021). The plot to kidnap michigan's governor.

[Grossman et al., 2020] Grossman, G., Kim, S., Rexer, J. M., and Thirumurthy, H. (2020). Political partisanship influences behavioral responses to governors' recommendations for covid-19 prevention in the united states. *Proceedings of the National Academy of Sciences*, 117(39):24144–24153.

[Hale et al., 2021] Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., et al. (2021). A global panel database of pandemic policies (oxford covid-19 government response tracker). *Nature human behaviour*, 5(4):529–538.

[Hanson, 1982] Hanson, S. (1982). The determinants of daily travel-activity patterns: relative location and sociodemographic factors. *Urban Geography*, 3(3):179–202.

[Hanson and Hanson, 1981] Hanson, S. and Hanson, P. (1981). The travel-activity patterns of urban residents: dimensions and relationships to sociodemographic characteristics. *Economic geography*, 57(4):332–347.

[Hastie et al., 2009] Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer.

[Hosseini and Gittler, 2020] Hosseini, M. S. and Gittler, A. M. (2020). Factors influencing human mobility during the covid-19 pandemic in selected countries of europe and north america. In *2020 IEEE International Conference on Big Data (Big Data)*, pages 4866–4872. IEEE.

[Hsiehchen et al., 2020] Hsiehchen, D., Espinoza, M., and Slovic, P. (2020). Political partisanship and mobility restriction during the covid-19 pandemic. *Public health*, 187:111–114.

[Hu et al., 2022] Hu, S., Xiong, C., Younes, H., Yang, M., Darzi, A., and Jin, Z. C. (2022). Examining spatiotemporal evolution of racial/ethnic disparities in human mobility and covid-19 health outcomes: Evidence from the contiguous united states. *Sustainable Cities and Society*, 76:103506.

[Huang et al., 2022] Huang, X., Lu, J., Gao, S., Wang, S., Liu, Z., and Wei, H. (2022). Staying at home is a privilege: Evidence from fine-grained mobile phone location data in the united states during the covid-19 pandemic. *Annals of the American Association of Geographers*, 112(1):286–305.

[Iio et al., 2021] Iio, K., Guo, X., Kong, X., Rees, K., and Wang, X. B. (2021). Covid-19 and social distancing: Disparities in mobility adaptation between income groups. *Transportation Research Interdisciplinary Perspectives*, 10:100333.

[Iovan et al., 2013] Iovan, C., Olteanu-Raimond, A.-M., Couronné, T., and Smoreda, Z. (2013). Moving and calling: Mobile phone data quality measurements and spatiotemporal uncertainty in human mobility studies. *Geographic information science at the heart of Europe*, pages 247–265.

[Järv et al., 2015] Järv, O., Müürisepp, K., Ahas, R., Derudder, B., and Witlox, F. (2015). Ethnic differences in activity spaces as a characteristic of segregation: A study based on mobile phone usage in tallinn, estonia. *Urban Studies*, 52(14):2680–2698.

[Jung et al., 2017] Jung, K., Garbarino, E., Briley, D. A., and Wynhausen, J. (2017). Blue and red voices: Effects of political ideology on consumers' complaining and disputing behavior. *Journal of consumer research*, 44(3):477–499.

[Kang et al., 2020] Kang, Y., Gao, S., Liang, Y., Li, M., Rao, J., and Kruse, J. (2020). Multiscale dynamic human mobility flow dataset in the us during the covid-19 epidemic. *Scientific data*, 7(1):390.

[Kim and Kwan, 2021] Kim, J. and Kwan, M.-P. (2021). The impact of the covid-19 pandemic on people's mobility: A longitudinal study of the us from march to september of 2020. *Journal of Transport Geography*, 93:103039.

[Kraemer et al., 2020] Kraemer, M. U., Sadilek, A., Zhang, Q., Marchal, N. A., Tuli, G., Cohn, E. L., Hswen, Y., Perkins, T. A., Smith, D. L., Reiner Jr, R. C., et al. (2020). Mapping global variation in human mobility. *Nature Human Behaviour*, 4(8):800–810.

[Lamb et al., 2021] Lamb, M. R., Kandula, S., and Shaman, J. (2021). Differential covid-19 case positivity in new york city neighborhoods: Socioeconomic factors and mobility. *Influenza and Other Respiratory Viruses*, 15(2):209–217.

[Limtanakool et al., 2006] Limtanakool, N., Dijst, M., and Schwanen, T. (2006). The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium-and longer-distance trips. *Journal of transport geography*, 14(5):327–341.

[Long and Ren, 2022] Long, J. A. and Ren, C. (2022). Associations between mobility and socio-economic indicators vary across the timeline of the covid-19 pandemic. *Computers, environment and urban systems*, 91:101710.

[Mallapaty, 2020] Mallapaty, S. (2020). The coronavirus is most deadly if you are old and male. *Nature*, 585(7823):16–17.

[McKenzie and Rapino, 2011] McKenzie, B. and Rapino, M. (2011). *Commuting in the united states: 2009*. US Department of Commerce, Economics and Statistics Administration, US ....

[Moro et al., 2021] Moro, E., Calacci, D., Dong, X., and Pentland, A. (2021). Mobility patterns are associated with experienced income segregation in large us cities. *Nature communications*, 12(1):4633.

[Oliver et al., 2020] Oliver, N., Lepri, B., Sterly, H., Lambiotte, R., Deletaille, S., De Nadai, M., Letouzé, E., Salah, A. A., Benjamins, R., Cattuto, C., et al. (2020). Mobile phone data for informing public health actions across the covid-19 pandemic life cycle.

[Palmer et al., 2013] Palmer, J. R., Espenshade, T. J., Bartumeus, F., Chung, C. Y., Ozgencil, N. E., and Li, K. (2013). New approaches to human mobility: Using mobile phones for demographic research. *Demography*, 50(3):1105–1128.

[Pappalardo et al., 2015] Pappalardo, L., Pedreschi, D., Smoreda, Z., and Giannotti, F. (2015). Using big data to study the link between human mobility and socio-economic development. In *2015 IEEE International Conference on Big Data (Big Data)*, pages 871–878. IEEE.

[SafeGraph, 2021] SafeGraph (2021). Safegraph.

[Silm and Ahas, 2014] Silm, S. and Ahas, R. (2014). Ethnic differences in activity spaces: A study of out-of-home nonemployment activities with mobile phone data. *Annals of the Association of American Geographers*, 104(3):542–559.

[Smith-Clarke et al., 2014] Smith-Clarke, C., Mashhadi, A., and Capra, L. (2014). Poverty on the cheap: Estimating poverty maps using aggregated mobile communication networks. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 511–520.

[Sy et al., 2021] Sy, K. T. L., Martinez, M. E., Rader, B., and White, L. F. (2021). Socioeconomic disparities in subway use and covid-19 outcomes in new york city. *American Journal of Epidemiology*, 190(7):1234–1242.

[Tai et al., 2021] Tai, D. B. G., Shah, A., Doubeni, C. A., Sia, I. G., and Wieland, M. L. (2021). The disproportionate impact of covid-19 on racial and ethnic minorities in the united states. *Clinical infectious diseases*, 72(4):703–706.

[Voss and Chi, 2006] Voss, P. R. and Chi, G. (2006). Highways and population change. *Rural sociology*, 71(1):33–58.

[Walensky, 2021] Walensky, R. P. (2021). Public health determination regarding an exception for unaccompanied noncitizen children from the order suspending the right to introduce certain persons from countries where a quarantinable communicable disease exists order under.

[Xu et al., 2018] Xu, Y., Belyi, A., Bojic, I., and Ratti, C. (2018). Human mobility and socioeconomic status: Analysis of singapore and boston. *Computers, Environment and Urban Systems*, 72:51–67.

[Yum, 2022] Yum, S. (2022). Spatial autocorrelation between human responses and winter storm grayson. *The Geographical Journal*, 188(4):546–558.

[Zhang et al., 2023] Zhang, Y., Cheng, S., Li, Z., and Jiang, W. (2023). Human mobility patterns are associated with experienced partisan segregation in us metropolitan areas. *Scientific Reports*, 13(1):9768.

[Zhao et al., 2008] Zhao, M., Mason, L., and Wang, W. (2008). Empirical study on human mobility for mobile wireless networks. In *MILCOM 2008-2008 IEEE Military Communications Conference*, pages 1–7. IEEE.

[Zhou and Lin, 2008] Zhou, X. and Lin, H. (2008). Spatial weights matrix.

Personal Declaration

I hereby declare that the submitted thesis is the result of my own independent work. All external sources are explicitly acknowledged in the thesis.

Signature:

Tao Peng

Date: September 30, 2023