2022

Spatiotemporal Analysis of Taxi and Transportation Network Companies (TNC) Demand in the Wake of COVID-19

Dewan Ashraful Parvez
University of Central Florida

SPATIOTEMPORAL ANALYSIS OF TAXI AND TRANSPORTATION NETWORK COMPANIES (TNC) DEMAND IN THE WAKE OF COVID-19

by

DEWAN ASHRAFUL PARVEZ
B.Sc. Bangladesh University of Engineering and Technology, 2017

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Civil, Environmental and Construction Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Summer Term
2022

Major Professor: Naveen Eluru

ABSTRACT

The objective of the thesis is to understand the factors affecting spatiotemporal ridehailing demand patterns as the COVID-19 pandemic has evolved. Specifically, the current study examines the key contributing factors of weekly ridehailing demand by employing Taxi and Transportation Network Companies (TNC) trip data from January 2019 through December 2020 for New York City. The ridehailing demand is partitioned across four time periods including Morning Peak, Morning Off Peak, Evening Peak and Evening Off Peak to accommodate for the time-of-day specific variations. Drawing on the high-resolution NYC data, the current study developed pooled spatial panel models to accommodate for the spatial and temporal heterogeneity. The thesis employs a recasting approach that enables the estimation of a parsimonious model specification across the four time periods. Two recasted spatial models: 1) Spatial Lag Model and 2) Spatial Error Model are estimated for ridehailing demand across the two services - Taxi and TNC - while considering a comprehensive list of factors including COVID-19 pandemic attributes, sociodemographic characteristics, land use and built environment attributes, transportation infrastructure and weather attributes. The model estimation results are further augmented with a robust policy analysis to predict potential ridehailing demand for future months. The policy exercise also illustrates how the proposed model can be employed by ridehailing companies and transportation agencies to examine ridehailing demand evolution as the pandemic continues.

Keywords: Ridehailing demand, Pandemic, Parsimonious pooled linear regression model, Spatial correlation, Policy exercise.

ACKNOWLEDGMENT

I would like to convey my sincere gratitude and thanks to my honorable supervisor Dr. Naveen Eluru for his excellent supervision and also for his constant support in this thesis.

I would also like to gratefully acknowledge New York City Taxi & Limousine Commission (NYCTLC) for providing access to trip data of Taxi and High Volume For Hired Vehicle for New York City.

I would further like to convey my heartiest gratitude to Dr. Tanmoy and Dr. Shamsunnahar Yasmin for their constant mentoring throughout this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

The ubiquitousness of technological advances including smartphone penetration, smartphone ridehailing app accessibility, seamless integration of payment systems, real-time driver and user reviews have resulted in the rapid growth of ridehailing demand across urban regions in the US. In fact, the advent of Transportation Network Companies (TNC) (such as Uber, Lyft, and Via) has significantly influenced ridehailing demand patterns. Prior to TNC emergence, the average daily trips by taxi (Yellow taxi) in New York City (NYC) ranged between 400,000 and 500,000 for the years 2010-2014 (NYC Taxi and Limousin Commission, 2021). After TNC services (also often referred to as ridesharing and/or ridesourcing services) began operation in 2014, the total number of ridehailing trips have increased. From 2015 through 2019, average daily TNC service demand increased from 60 thousand to 720 thousand while the corresponding taxi demand (Yellow and the newly introduced Green cabs) reduced from 450 thousand to 230 thousand (NYC Taxi and Limousin Commission, 2021). The reader would note that while there is a noticeable reduction in taxi trips, overall ridehailing demand has rapidly increased (see Dey et al., 2021b for details).

The tremendous growth story of ridehailing demand has met with a system shock in the form of Coronavirus Disease 2019 (COVID-19). As of January 2022, COVID-19 has significantly impacted the entire world with reported cases (and fatalities) of 324 million (5.5 million) worsening physical, mental and financial health of billions of individuals (COVID Live - Coronavirus Statistics - Worldometer, 2022). The pandemic and the associated stay-at-home mandates, and social distancing has affected every facet of life. In March 2020, after World Health Organization (WHO) declared COVID-19 as a pandemic and the first COVID-19 case was detected in NYC, a sudden drop in Taxi and TNC demand was observed (Archived: WHO

Timeline - COVID-19, 2021, First Case of COVID-19 in NYC, 2020, NYC Taxi and Limousin Commission, 2021). Being one of the first epicenters of COVID-19 in the US, NYC was significantly affected by COVID-19 associated mandates and shutdowns (New York City COVID-19 Economic Impact Update, 2020, Thompson, 2020). The influence of COVID-19 on transportation will be a matter of research for the foreseeable future. Research dimensions of interest are likely to include: (a) COVID-19 era teleworking patterns and their evolution, (b) transportation mode choice behavior changes, (c) public transportation and ridehailing adoption and (d) tourism travel mode choice preferences.

1.1 Motivation for The Study

In the current thesis, we seek to understand the factors affecting ridehailing demand patterns as the pandemic evolved. Any attempt to study the impact of COVID-19 should account for ridehailing patterns prior to COVID-19. Also, the influence of various factors on ridehailing demand might be moderated due to COVID-19 cases emerging in the region. To elaborate, during pre-COVID, restaurants and employment centers were potentially significant contributors to ridehailing demand. These factors impact ridehailing demand differently across different time periods of the day. For instance, pre-COVID employment centers might have been strong contributors to demand in the AM peak and PM peak periods. On the other hand, during COVID-19, employment centers might have a substantially lower effect (relative to pre-COVID-19) on ridehailing demand. While restaurants in dense locations might have been strong contributors of demand during midday (work lunch) and PM peak and night periods (dinner) in pre-COVID time period, the impact might have significantly altered during COVID-19. Thus, the demand mechanism for ridehailing might have gone through a significant change due to the COVID-19 pandemic. For effective policy implications, it is of utmost importance to accommodate for such changes (if they exist) in examining ridehailing demand.

Examining such detailed hypothesis requires compiling longitudinal ridehailing demand data with spatial and temporal variabilities. Using high resolution data from NYC, the current study examines the key contributing factors of ridehailing demand by employing TNC and Taxi trip data from January 2019 through December 2020. The detailed spatio-temporal analysis is conducted at the taxi zone[1] level by aggregating TNC (such as Uber, Lyft, and Via) and Taxi (Yellow and Green taxi combined) weekly demand from January 2019 through December 2020. To accommodate for the time-of-day specific variations, the ridehailing demand is considered at the time period resolution as follows: AM peak (6AM–10AM), Midday (10AM–4 PM), PM peak (4 PM-8PM), and Night (8PM-6AM). Separate ridehailing demand models are estimated for Taxi and TNC services by time-of-day while considering an extensive set of independent variables including COVID-19 pandemic attributes, sociodemographic characteristics, land use and built environment attributes, transportation infrastructure and weather attributes. The ridehailing demand models are estimated considering spatial correlations across different taxi zones by employing Spatial Lag or Autoregressive Model (SAR) and Spatial Error Model (SEM). The reader would note that instead of estimating 4 separate time of day models, we undertake a

[1] Taxi zones are spatial zones analogous to the transportation analysis zone that allow aggregate reporting of the pick-up and drop-off locations for riders in NYC (NYC Taxi Zones, 2021). The NYC region is divided into 263 taxi zones. After cleaning the data, 258 taxi zones remained.

recasting exercise to treat the data as a repeated measure by time-of-day enabling us to specify a parsimonious demand model structure.

1.2 Study Methodology and Objective

Toward addressing the aforementioned issues, the current study aims to develop a comprehensive model for TNC and taxi demand by integrating the impact of COVID-19. Spatial Lag Model and Spatial Error Model were implemented to capture the spatial dependency at a finer resolution on NYCTLC's data from 2019 to 2020 for New York City region. Furthermore, performance of the developed model is examined by comparing system level observed demand versus system level predicted demand which includes the projection of 2019, 2020 and first six months of 2021. The projection by proposed model reflects reasonable performance. Finally, application of the developed model is illustrated in this study. Model illustrations provide plots recovery rate in the study area and projection versus observed demand.

1.3 Thesis Structure

The rest of the thesis is organized as follows: Chapter 2 provides a brief review of relevant research and positions the current study. Chapter 3 describes the formulation and estimation procedure of the pooled regression model structures with and without spatial effect. Chapter 4 provides details of data preparation, dependent and independent variables considered in the analysis. Measure of fit for the models, model estimation results are presented in chapter 5. Chapter 6 illustrates the ridehailing demand projection by considering two different hypothetical scenarios and comparing the prediction with the observed trends. Finally, a summary of study findings and conclusions are presented in Chapter 7.

CHAPTER 2: LITERATURE REVIEW

This chapter describes the existing literature on rideshare demand analysis and positions the current context of this study.

2.1 Existing Literature

The revival of ridehailing in urban regions with the emergence of TNC alternatives has received significant attention in the transportation field. Several earlier research efforts focused on operational and quantitative analysis of traditional taxi services (Nie et al., 2021, Wang et al., 2018), TNC evolution and qualitative aspects of TNC adoption (Alemi et al., 2018, Chan and Shaheen, 2011, Clewlow, Regina R. Mishra, 2017, Furuhata et al., 2013, Lahkar et al., 2018, Loa and Nurul Habib, 2021, Sun and Edara, 2015), competition between ridehailing and public transit (Faghih-Imani et al., 2017, Habib, 2019, Komanduri et al., 2018, Lavieri et al., 2018, Rayle et al., 2016), and ridehailing demand analysis (Clewlow, Regina R. Mishra, 2017, Correa et al., 2017, Feigon and Murphy, 2016, Loa and Nurul Habib, 2021, Zhang and Zhang, 2018). As the focus of the current study is on ridehailing demand, for our literature review, we restrict ourselves to the following dimensions of interest: (a) quantitative studies examining ridehailing demand focusing on Taxi and/or TNC (mainly conducted pre COVID) and (b) studies examining the influence of COVID-19 on ridehailing demand.

With regards to the first group of research, several studies analyzed the dynamics of pre-COVID ridehailing demand. Researchers examined data from New York, Austin, and Shenzhen. Using data from New York, Correa et al., 2017 developed spatial regression models for taxi and Uber demand. Gerte et al., 2018 studied Uber demand in NYC employing a panel based random effects model to identify if Uber demand can continue to grow unbounded. The authors concluded that Uber demand has hit its upper bound in heavily populated residential neighborhoods. Dey et al., 2021b studied the evolution of ridehailing demand in New York at the taxi zone level after the introduction of TNC services using a joint negative binomial and multinomial fractional split model. Two studies analyzing RideAustin TNC trip demand employing sophisticated count model frameworks – Geographically Weighted Regression and Multivariate Spatial Count model - concluded that demographic characteristics, built environment and public transit affected TNC demand (Lavieri et al., 2018, Yu and Peng, 2019). Examining data from Shenzhen, China, the authors examined the competition between taxi industry and TNC. The authors surprisingly found that taxi industry in China was successful at surviving the emergence of TNC (Nie et al., 2021).

The second group of studies focused on the influence of COVID-19 on ridehailing demand. A number of studies, using survey data from Australia, India and China found that preference for ridehailing services during the pandemic were substantially lowered (Beck and Hensher, 2020, Bhaduri et al., 2020, Tan and Ma, 2021). Using ridehailing data from Chicago, the first study focusing on COVID-19 impact on ridehailing concluded that ridehailing demand fell substantially, ridehailing travel distances lengthened and ridehailing travel times reduced (Du and Rakha, 2020). A study using six months of data from Chongqing, developed spatial regression models for daily taxi demand and arrived at similar conclusions (Nian et al., 2020). An analysis of taxi data from Shenzhen during and after the lockdown concluded that taxi demand recovered slower than personal vehicle travel. The government incentives offered post lockdown allowed maintaining pre-pandemic level taxi supply (Nian et al., 2020).

2.2 Current Study Context

The review of literature clearly illustrates the burgeoning literature on ridehailing and emerging literature on COVID-19 impacts on ridehailing. However, there is still scope for improving our understanding of ridehailing demand, especially in the context of COVID-19. Earlier research investigating the influence of COVID-19 did not employ data over a longer time frame i.e., evolving ridehailing behavior with time over the pandemic was rarely incorporated. The data analysis also did not directly consider the varying patterns of COVID cases in the region. Additionally, while ridehailing demand was studied on a daily basis, the influence of time-of-day effects were often neglected. The reader would note that a same independent variable might have different impacts on ridehailing demand for different parts of the day. Finally, while some studies consider longitudinal demand data, ridehailing demand is not always examined at a spatially fine resolution.

The current research addresses these limitations by developing a comprehensive model for taxi and TNC demand. The proposed study employs data for twenty-four months - from January 2019 through December 2020. The data was obtained at the taxi zone resolution. The data was appropriately aggregated by week and time of day (AM peak, Midday, PM peak, and Night periods) to obtain the repeated measure demand variables for taxi and TNC. The reader would note that yellow and green taxi trips were added to obtain taxi demand. TNC demand was generated as a sum of Uber, Lyft, Via and Juno services. With these preliminary operations, we end up with 8 dependent variables (4 each for taxi and TNC). Instead of estimating a 4-dimensional multivariate model for taxi or TNC, we resort to a recasting approach that allows us to consider a single pooled model for all 4 time periods. Specifically, the data is organized as repeated records of trip demand by time period for each ridehailing service. The proposed recasting approach of using pooled univariate model as opposed to multivariate model has been adopted successfully in Bhowmik et al. 2019 (Bhowmik et al., 2019). The recasting approach is appealing for several reasons including: (i) it is computationally less burdensome, (ii) a single trip demand model for each ridehailing service can be estimated instead of estimating separate models for separate time periods, and (iii) it can easily accommodate for random effects and correlation structures without leading to an explosion in the number of parameters (Bhowmik et al., 2019). The demand data was augmented with a comprehensive set of independent variables including COVID-19 case information over time, sociodemographic characteristics, land use and built environment characteristics, transportation infrastructure and weather attributes. The model results are augmented with a robust policy analysis to predict potential ridehailing demand for future time periods. The policy exercise also illustrates how the proposed model can be employed by ridehailing companies and transportation agencies to examine ridehailing demand evolution as the pandemic continues.

2.3 Summary

This chapter provides a detailed discussion of existing literature related to ridehailing demand focusing on pre and post COVID-19 era. Further, the current study context is also described in this chapter. The next chapter will describe econometric methodology used in this thesis.

CHAPTER 3: METHODOLOGY

The previous chapter presented a detailed discussion of the different methodologies used in earlier research for ridehailing demand in pre and post COVID-19 era. In this chapter, we provide mathematical details of the model frameworks employed in our study. The chapter starts with the formulation for the pooled linear regression model without spatial effect and then provides details of the spatial panel regression models subsequently.

3.1 Pooled Linear Regression Model

The major focus of this study is to analyze the weekly ridehailing demand for Taxi service and TNC services by different time periods. The demand variables are continuous in nature, and hence, a linear regression technique is employed for analyzing these continuous variables. However, the records are associated with multi-level repeated measures - spatially (at taxi zone level) and temporally (at week level). In taking into account possible correlations of the repeated measures, in the current research effort, we employ spatial panel regression model (please see Elhorst (Elhorst, 2003) for details on the methodology).

Let i $(= 1, 2, 3 \dots, N = 258)$ be an index to represent the taxi zone, $t(= 1,2,3 \dots, T = 4)$ be an index to represent the various time periods, $k(= 1, 2, 3 \dots \dots \dots, K = 106)$ be an index to represent weeks and d be an index to represent the different ridehailing services (Taxi and TNC). Let $y_{d_{itk}}$ be the natural logarithm of weekly ridehailing demand in taxi zone i at time period t and week k for the ridehailing service d. In the current study context, separate models are estimated for Taxi and TNC services and hence d is omitted in the following equations for simplicity. Then the formulation of the pooled linear regression model considering spatial effects without the spatial dependency can be written as:

$$y_{itk} = \beta' X_{itk} + \varepsilon_{itk} + \delta_i \quad (1)$$

where X_{itk} is a vector of attributes that influence ridehailing demand (taxi or TNC) and β' is the corresponding coefficients to be estimated (including a scalar constant). ε_{itk} is independently and identically distributed error term with zero mean and variance σ^2. The δ_i represents time-invariant unobserved spatial effect for taxi zone i . We can specify δ_i by using either as a fixed effect or random effect. However, a fixed effect model is not suitable in the presence of time-invariant exogenous variables (such as sociodemographic and land use attributes) (Faghih-Imani and Eluru, 2016).Therefore, in the current study, δ_i is specified as a random effect.

3.2 Spatial Panel Regression Model

In accommodating for spatial effects, several models have been considered in existing literature including Spatial Lag or Autoregressive Model (SAR), Spatial Error Model (SEM), and Geographically Weighted Regression Model (GWRM). In the current study, SAR and SEM modeling techniques are adopted for capturing the spatial dependence across different taxi zones. The main difference between these two modeling structures lies in how the framework accounts for spatial dependence. In SAR model, the spatial interactions are considered through a spatially lagged dependent variable whereas SEM model considers spatial lagged error structure in incorporating spatial correlation (please see Rahman et al., 2021 for details).

The general structure of the SAR (see equation 2) and SEM (see equation 3 and 4) models can be presented as:

$$y_{itk} = \alpha \sum_{j=1}^{N} W_{ij} y_{itk} + \beta' X_{itk} + \varepsilon_{itk} + \delta_i \tag{1}$$

$$y_{itk} = \beta' X_{itk} + \delta_i + \vartheta_{itk} \tag{2}$$

$$\vartheta_{itk} = \gamma \sum_{j=1}^{N} W_{ij} \vartheta_{jtk} + \varepsilon_{itk} \tag{3}$$

where, α is a spatial autoregressive coefficient; γ is a spatial autocorrelation coefficient, ϑ_{itk} is a spatial autocorrelated error term in taxi zone i, at time period t and week k; and W_{ij} is an element between taxi zones i and j in the spatial weight matrix W where i is the taxi zone of interest and j is any other taxi zone other than i.

In our analysis, building on earlier literature, several functional forms of W_{ij} matrix is adopted including adjacent zones, inverse of distance square from the taxi zone of interest to other taxi zones, inverse of distance from the taxi zone of interest to other taxi zones, or different threshold values of distance (such as unit within 500m, 1mile, 5 miles). In the current study context, inverse of distance is found to provide the best data fit and hence is considered in the final model specification. Please note that, the diagonal of weight matrix is set to zero to prevent the use of y_{itk} to the model itself. Also, a row-normalized form of the W matrix is employed as the spatial weight matrix for enhancing stability in estimation (Elhorst, 2003). The models are estimated in MATLAB using the routines provided by Elhorst, 2014a and Elhorst, 2003. All the parameters are estimated using the maximum likelihood approaches (see Elhorst, 2014b for details on likelihood functions).

3.3 Summary

The main objective of this study is to develop a spatial panel regression model to capture spatial dependency across different taxi zones. This chapter presented a detailed discussion of the econometric methodology employed in this thesis. The next chapter will present a detailed description of the dataset used for our analysis.

CHAPTER 4: DATA PREPARATION

The previous chapter discussed the econometric framework employed in this thesis. This chapter will provide a detailed discussion of the ridehailing dataset employed in our study. The following section will describe procedure and steps for data preparation along with the dependent and independent variables used for the analysis.

4.1 Data Preparation

The data for the current study is sourced from the NYC Taxi and Limousine Commission (TLC) data warehouse which provides spatially aggregated trip data from all ridehailing companies (including taxi, Uber, Lyft, Juno and Via) for public use (TLC Trip Record Data - TLC, 2021). The trip data for the 258 Taxi zones of NYC are extracted from January 2019 to December 2020.

The major focus of this study is to examine weekly ridehailing trip data at taxi zone level across different times of the day for taxi and TNC services separately. To better understand the impact of COVID-19 on ridehailing services, the trip data is separated by time of day including AM peak (6AM–10AM), Midday (10AM–4 PM), PM peak (4 PM-8PM), and Night (8PM-6AM). Finally, in generating the dependent variables, the daily raw trip data are aggregated by time of day and by week for different ridehailing services for each taxi zone. The trip data is considered for 24 consecutive months (or 106 consecutive weeks) from January 2019 through December 2020. During this period, COVID-19 weekly cases was recorded to be as high as 6000 weekly new cases in NYC. Over the duration of the pandemic considered in our analysis (March 2020 through December 2020), there was an 83% drop in weekly Taxi trips and a 53% drop in weekly TNC trips in NYC. Figure 4.1 (a) and 1 (b) illustrate these weekly trends in ridehailing trip behavior and COVID-19 cases for NYC.

The independent variables considered in this study can broadly be categorized as: 1) Pandemic attributes (COVID-19 Map - Johns Hopkins Coronavirus Resource Center, 2021), 2) Land use and built environment attributes (NYC Open Data, 2021), 3) Transportation infrastructure attributes (NYC Open Data, 2021, NYS GIS Clearinghouse - NYS GIS Program Office - NYS Tax Parcels, 2021), 4) Sociodemographic attributes (Explore Census Data, 2021), and 5) Weather attributes (Daily Summaries Station Details | National Climatic Data Center, 2021). Pandemic attributes, obtained from John Hopkins University (COVID-19 Map - Johns Hopkins Coronavirus Resource Center, 2021), included COVID-19 cases from 2 weeks before, COVID-19 cases from 3 weeks before and percent difference between COVID-19 cases of last week and average COVID-19 cases of last three weeks. Land use and built environment attributes, compiled using data from NYC Open Data (NYC Open Data, 2021) and Department of Finance Tax Map Office (NYS GIS Clearinghouse - NYS GIS Program Office - NYS Tax Parcels, 2021) included transit score, number of restaurants, proportion of parking area, land use mix, airport, distance from Times Square[2] and points of interest. Transportation infrastructure attributes, sourced from NYC Open Data (NYC Open Data, 2021) included street length density, bike share stations and number of bus stops and subway stations. Sociodemographic attributes, drawn from

[2] Times Square represents an iconic destination in New York representing the center of the business district. Several earlier papers modeling bikeshare and ridesourcing have considered Times Square as a significant point of interest (see Dey et al., 2021a; Kumar Dey et al., 2021, Liu et al., 2021, Espín Noboa et al., 2016 and Dimitriou et al., 2016).

United States Census Bureau (Explore Census Data, 2021) included employment density and low-income indicator. Finally, weather attributes, obtained from National Oceanic and Atmospheric Administration (NOAA) (Daily Summaries Station Details | National Climatic Data Center, 2021), included snow depth and precipitation. In order to capture the effect of pandemic on ridehailing demand, several interaction terms of the exogenous variables with pandemic effect (binary indicator March 2020 to December 2020) were considered. The land use and built environment attributes, transportation infrastructure and sociodemographic attributes are computed at taxi zone level while the weather variables are generated specific to the week for which the ridehailing demand is computed. Table 4.1 presents the descriptive statistics of the dependent variables and exogenous variables in the first and second row panels, respectively.

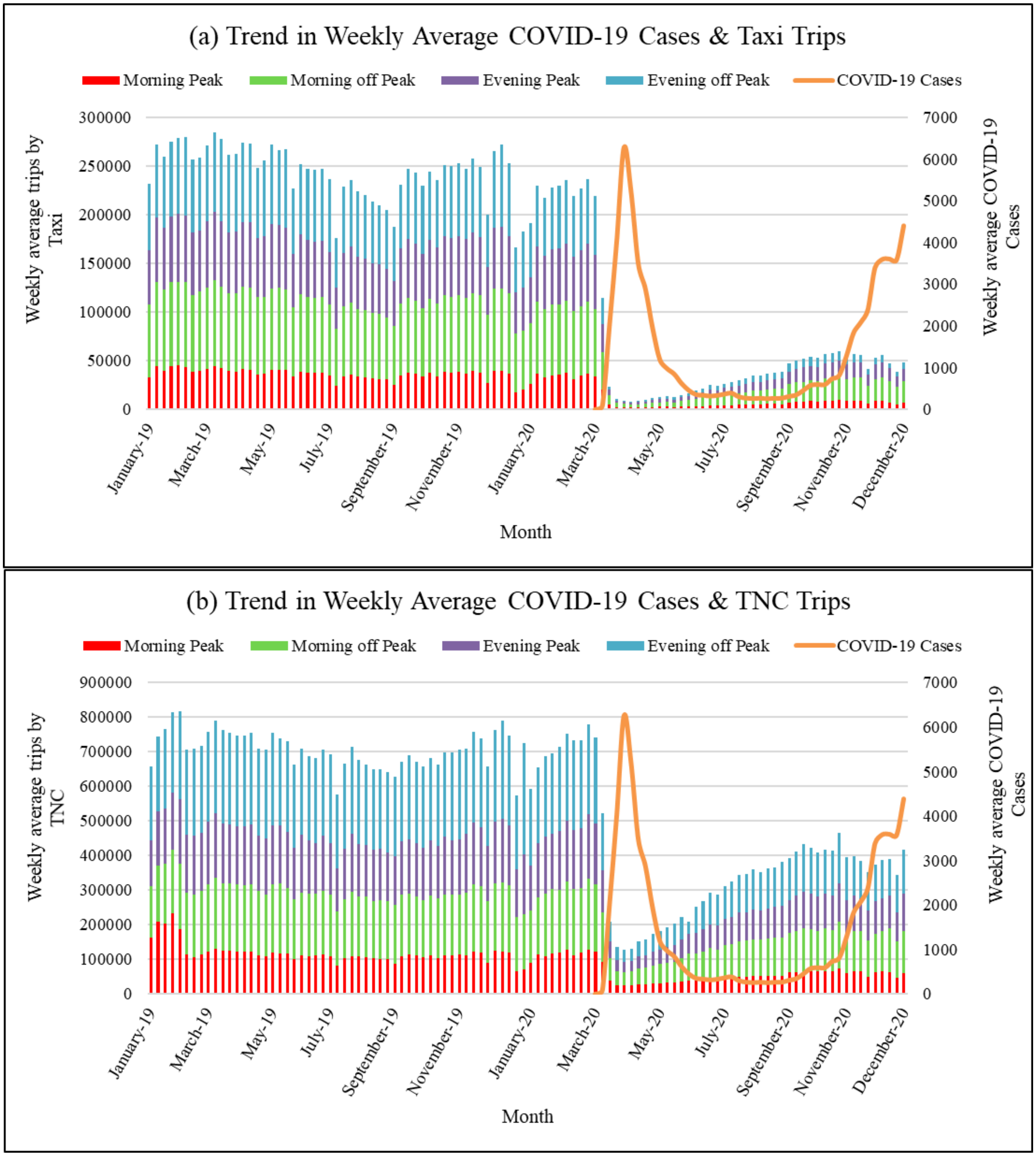


Figure 4.1 Trend in weekly average COVID-19 cases & ridehailing demand: (a) Taxi; (b) TNC

Table 4.1 Descriptive statistics of variables

Variables	Variable Descriptions	Descriptive Statistics		
		Minimum	Maximum	Mean
DEPENDENT VARIABLES				
Number of taxi trips during AM peak period	Ln (Number of weekly taxi trips during AM peak period (6AM – 10 AM))	0.000	9.694	3.952
Number of taxi trips during Midday period	Ln (Number of weekly taxi trips at Midday period (10AM – 4 PM))	0.000	10.417	4.570
Number of taxi trips during PM peak period	Ln (Number of weekly taxi trips at PM peak period (4 PM – 8 PM))	0.000	10.085	4.008
Number of taxi trips during Night period	Ln (Number of weekly taxi trips at Night period (8 PM – 6 AM))	0.000	10.238	4.018
Number of TNC trips during AM peak period	Ln (Number of weekly TNC trips at AM peak period (6AM – 10 AM))	0.000	10.110	7.083
Number of TNC trips during Midday period	Ln (Number of weekly TNC trips at Midday period (10AM – 4 PM))	0.000	10.444	7.624
Number of TNC trips during PM peak period	Ln (Number of weekly TNC trips at PM peak period (4 PM – 8 PM))	0.000	10.208	7.429
Number of TNC trips during Night period	Ln (Number of weekly TNC trips at Night period (8 PM – 6 AM))	0.000	10.849	7.689
INDEPENDENT VARIABLES				
COVID-19 Pandemic Attributes				
COVID-19 Cases (2-week lag)	Total COVID-19 cases in each Taxi Zone 2 weeks before the week for which ridehailing trips are considered/100 population	0.000	0.688	0.043
COVID-19 Cases (3-week lag)	Total COVID-19 cases in each Taxi Zone in 3 weeks before the week for which ridehailing trips are considered/100 population	0.000	0.688	0.040
Percent Difference in COVID-19 cases between last week and 3 weeks moving average	(COVID-19 cases in each taxi zone a week before the week for which ridehailing trips are considered - average COVID-19 cases in each taxi zone recorded over 3 weeks before the week for which ridehailing trips are considered) % for 3 weeks	-45.479	187.568	4.346
Land Use and Built Environment Attributes				
Number of Restaurants	Total number of restaurants in each Taxi Zone/100	0.000	6.810	1.141
Proportion of Parking area	Parking area in each Taxi Zone/total area in each taxi zone	0.000	0.281	0.023
Land use mix	Land use mix = $[\frac{-\sum_k (P_k(lnP_k))}{lnN}]$, where k is the category of land-use, p is the proportion of the developed land area for specific land-use, N is the number of land-use categories	0.000	0.965	0.533

Variables	Variable Descriptions	Descriptive Statistics		
		Minimum	Maximum	Mean
Airport indicator	Presence of airport in Taxi Zone (Dummy Variable)	0.000	1.000	0.008
Distance from Times Square	Distance from Times Square to each Taxi Zone (length/10,000) in meter	0.008	3.636	1.297
Points of Interest	Total number of points of interest in each Taxi Zone/100	0.030	3.250	0.762
Transportation Infrastructures				
Transit Score	Transit Score (a measure of serviceability of public transit) in each Taxi Zone/100*	0.000	1.000	0.891
Street Length Density	Total Street Length in each Taxi Zone/Total area of each Taxi zone	0.003	0.103	0.052
Bike Share Stations	Total number of bike share station in each taxi zone	0.000	27.000	2.302
Bus Stops and Subway Stations	Total number of bus stops and subway stations in each taxi zone/100	0.000	0.610	0.151
Sociodemographic Attributes				
Employment Density	Total number of Employment in each Taxi Zone per 1000 acre	0.001	0.713	0.155
Low Income Indicator	Taxi Zone with median income under $50 thousand USD (25th percentile) (Dummy Variable)	0.000	1.000	0.217
Weather Attributes				
Snow Depth	Snow depth in each Taxi Zone (in)	0.000	5.457	0.084
Precipitation	Precipitation in each Taxi Zone (in)	0.000	0.570	0.134

4.2 Summary

This chapter provides details of data preparation procedures including a detailed description of the dependent and the independent variables. The next chapter will present the empirical analysis, model estimates and discuss the results in detail.

CHAPTER 5: EMPIRICAL ANALYSIS

This chapter provides model estimation results of the proposed simple linear regression, pooled linear regression and spatial panel regression models including the parameter estimates for the best model structure.

5.1 Model Specification and Overall Measure of Fit

The empirical analysis involves estimation of eight different models – four models for Taxi service and four models for TNC service. For each ridehailing service, the models estimated are: 1) Simple Linear Regression Model (LR), 2) Pooled Linear Regression Model (PLR), 2) Spatial Lag Pooled Linear Regression Model (SLPM) (weight matrix: 1/distance) and 4) Spatial Error Pooled Linear Regression Model (SEPM) (weight matrix: 1/distance). The reader would note that the simple linear regression (LR) model system includes 4 separate models by time of day. Subsequently, a parsimonious pooled linear regression (PLR) model system is estimated that reduces the number of parameters significantly without any measurable loss in data fit. The PLR model serves as the base model for estimating the advanced spatial models.

Before discussing the estimation results, the estimated models are compared using the Bayesian Information Criteria (BIC) for identifying the best model. The BIC results for the models are summarized in Table 5.1.

Table 5.1 Measures of fit: Log-likelihood (LL) and Bayesian Information Criteria (BIC)

Model	Taxi			TNC		
	Number of Parameters	LL	BIC	Number of Parameters	LL	BIC
Simple	88	-179275.0	358993.5	88	-142958.8	286360.9
Pooled	70	-179281.2	358915.1	60	-142967.9	286238.2
SLPM	67	-88750.2	178278.7	55	-27342.2	55322.5
SEPM	60	-88774.0	178244.2	55	-22869.5	46091.8

From the Table, we can observe that for Taxi and TNC services, spatial error pooled linear models (SEPM) regression models have the lowest BIC values. Thus, in the current study context, we can conclude that SEPM model is the best model structure in accommodating the spatial correlations across taxi zones for Taxi and TNC ridehailing services. In the following sections, only SEPM model results are discussed for the sake of brevity.

5.2 Model Estimation Results

The estimation results of the SEPM models for Taxi and TNC services are presented in the second and third column panels of Table 5.2, respectively. In models with pooled data structure, one of the dimensions (time periods) must be considered the base for every exogenous variable. In our study, Night period is considered as the base in estimating the main effect that applies to all time periods and deviations are estimated for other three time periods. For main effect, a positive (negative) coefficient corresponds to increase (decrease) in ridehailing demand. If the deviation term is significant for any variable, it highlights that the variable effect significantly differs for that time period. For example, the base effect for number of restaurants is 0.411 and the deviation terms for AM peak, Midday and PM peak periods are -0.201, -0.199, -0.100, respectively. Hence, the net effects across different time periods can be computed as summation of parameter in the main effect and the parameter in the respective time period. The net effects of this variable across the time periods are 0.210 (AM peak), 0.212 (Midday), 0.311 (PM peak) and 0.411 (Night), respectively. These results highlight the distinct impact of number of restaurants across four-time periods. A positive (negative) net effect corresponds to increase (decrease) in ridehailing demand in a time period. The reader would note that for some exogeneous variables, we did not find any significant impact for the night time (for example: parking area and precipitation rate). For such instances, the influence of the variable for other time periods will actually be the deviation itself. In the following sections, the effects of variables are discussed by variable groups. The results for Taxi and TNC models are discussed together.

Table 5.2 provides parameter estimates of the time of the day models for weekly taxi and TNC trips considered in the study. A positive (negative) value of the parameter in Table 5.2 indicates propensity for higher (lower) duration.

5.2.1 Pandemic Attributes

Several pandemic related attributes were considered in our model including weekly COVID-19 transmission rate with 2- and 3-week lag, 3-week moving average of COVID-19 cases, percent difference in COVID-19 cases between last week and the 3-week moving average and an indicator variable to represent whether cases increased relative to the 3-week moving average. As expected, outbreak of pandemic contributes towards a significant reduction in ridehailing demand. For the Taxi demand, the 2-week lag variable offers the same impact across all time periods (as indicated by no significant deviation parameters). However, 2-week lag variable offers a varying impact across time periods for TNC demand. Specifically, a slightly lower impact on TNC demand for AM peak, Midday and PM peak periods was observed relative to Night period. In addition to the 2-week lag variable, the 3-week lag variable also affects TNC demand (no impact on Taxi demand). The variable indicates a time period invariant negative impact across all time periods.

It is important to note that several of the deviations for 2-week and 3-week lags of COVID-19 cases are insignificant indicating across some time periods the impacts are similar. These results support our hypothesis that some exogenous variables may have similar impacts for all time periods and thus employing the recasting approach enhances model parsimony. Percentage difference in COVID-19 cases in the last week and 3-week moving average also show negative impacts on ridehailing demand, which implies that ridehailing demand is likely to decrease if the COVID-19 cases are increasing relative to the 3-week average (Bhaduri et al., 2020, Nian et al., 2020). The reader would note that since pandemic related attributes might be correlated, we tested for the presence of multicollinearity by estimating the Variance Inflation Factor (VIF) (see Bhowmik et al., 2021 for details). From the results, we found that the VIF values are close to 1 indicating an absence of significant correlation.

We also considered indicator variables that represent the impact of increased awareness and experience with COVID on ridehailing demand in our analysis. Specifically, we examined the pandemic effect after various months (such as September 2020, October 2020 and so on). In our analysis, we found a positive impact on the ridehailing demand post September 2020. The coefficients across time of day indicate that Taxi demand is likely to be greater at daytime while TNC demand was found to be similar across midday and PM peak.

5.2.2 <u>Land Use and Built Environment Attributes</u>

Several land use and built environment attributes are found to be significant indicators of ridehailing demand. With regards to number of restaurants, the main effects in both ridehailing systems are positive indicating that the ridehailing demand is likely to be higher in the zones with higher number of restaurants. The variable has negative effects for both systems in AM peak, Midday and PM peak periods. However, the net effects of the variable across all time periods are still positive which implies that there is higher demand of ridehailing services in zones with higher restaurants (see similar results in Dey et al., 2021b), but the demand is likely to be less in daytime relative to Night period. As expected, the effect of number of restaurants after the pandemic outbreak has negative impact indicating lower eating-out activities in these zones. The interaction terms have positive coefficients across different time periods, but the net effects are negative implying overall lower ridehailing demand for the zones with higher restaurants during pandemic. Availability of higher parking area in taxi zones are less likely to attract taxi services during Midday. For TNC, parking area is found to be significant only for AM peak with a negative coefficient. These results are perhaps indicating that these zones are rather attractive for higher level of personal vehicle activities (Sabouri et al., 2020).

The effect of land use mix is positive for both services. It is interesting to note that a higher level heterogenous land use mix generates more taxi demand in midday period relative to other time periods. On the other hand, midday period generates less TNC trips in zones with higher land use mix relative to other time periods. Such contrasting effects support our hypothesis that the demand mechanisms in different time periods are different for different ridehailing services (similar findings Li et al., 2021).

As expected, airport indicator has a positive impact on both ridehailing demands. After the outbreak of COVID-19 in NYC in March 2020, number of flights have declined internationally and domestically. Such reduced air travel activities resulted in lower level of travel to and from airports which is reflected in our models. The interaction of airport and pandemic outbreak show negative associations with ridehailing demand with heterogenous effects across different time periods. The taxi demand is likely to decrease with increasing distance from Time Square. The effects are more negative during daytime relative to night period. For TNC, the distance of Time square is significant for midday and PM peak with negative coefficients. On the other hand, points of interest are found to have positive effects on ridehailing demand for Taxi. In the case of TNC, points of interest contribute to increased demand only in the night time. The results are perhaps indicating that higher level of recreational activities in the zones with higher tourist attractions can contribute to a higher level of taxi demand as they are likely to be more visible in such locations (Wang and Noland, 2021, Zhang et al., 2020)

Table 5.2 Pooled spatial panel model results

Ridehailing Services	TAXI				TNC			
Time-of-The-Day	**Main Effect**	**Deviation for AM Peak**	**Deviation for Midday**	**Deviation for PM Peak**	**Main Effect**	**Deviation for AM Peak**	**Deviation for Midday**	**Deviation for PM Peak**
Variable Name	**Estimate (t stat)**	**Estimate (t stat)**	**Estimate (t stat)**	**Estimate (t stat)**	**Estimate (t stat)**	**Estimate (t stat)**	**Estimate (t stat)**	**Estimate (t stat)**
Constant	1.642 (2.854)	-0.325 (-12.643)	--	-0.555 (-10.799)	1.513 (4.332)	-0.723 (-30.866)	0.096 (3.368)	--
Pandemic Attributes								
COVID-19 Cases (2-week lag)	-2.000 (-45.414)	--	--	--	-1.997 (-22.861)	0.600 (7.681)	0.357 (4.599)	0.231 (2.985)
COVID-19 Cases (3-week lag)	--	--	--	--	-0.333 (-4.220)	--	--	--
Percent Difference between last week and 3 weeks moving average	-0.0003 (-2.371)	--	--	--	-0.0013 (-14.187)	0.0012 (6.341)	--	--
Effect Since September'20	0.043 (1.992)	0.303 (10.128)	0.432 (14.371)	0.367 (11.536)	0.213 (23.794)	-0.065 (-3.632)	--	--
Land Use and Built Environment Attributes								
Number of Restaurants	0.411 (4.225)	-0.201 (-31.457)	-0.199 (-30.942)	-0.100 (-15.479)	0.321 (3.697)	-0.165 (-49.420)	-0.148 (-45.159)	-0.094 (-30.252)
Number of Restaurants*Pandemic Period	-0.148 (-24.551)	0.067 (7.933)	0.124 (14.716)	0.075 (8.716)	-0.075 (-23.534)	0.068 (15.046)	0.073 (16.110)	0.060 (13.226)
Proportion of Parking Area	--	--	-0.793 (-6.982)	--	--	-0.974 (-16.203)	--	--
Land Use Mix	1.061 (3.298)	--	0.230 (10.948)	-0.085 (-4.302)	3.418 (11.583)	0.212 (17.343)	-0.133 (-11.452)	-0.245 (-22.562)
Airport Indicator	7.317 (8.290)	-0.802 (-17.227)	--	--	5.426 (6.622)	--	--	--
Airport Indicator*Pandemic Period	-1.808 (-45.056)	-0.607 (-8.335)	--	0.240 (4.127)	-1.408 (-82.851)	-0.873 (-30.567)	--	--

Ridehailing Services	TAXI				TNC			
Time-of-The-Day	Main Effect	Deviation for AM Peak	Deviation for Midday	Deviation for PM Peak	Main Effect	Deviation for AM Peak	Deviation for Midday	Deviation for PM Peak
Times Square Distance	-1.034 (-7.986)	0.234 (15.743)	0.216 (17.947)	0.106 (6.753)	--	--	-0.072 (-8.145)	-0.083 (-11.505)
Points of Interest	0.477 (2.897)	-0.206 (-21.897)	-0.062 (-6.325)	--	--	--	--	0.086 (21.621)
Transportation Infrastructures								
Transit Score	1.916 (3.647)	--	0.059 (2.258)	0.409 (9.897)	3.058 (7.108)	-0.234 (-9.636)	-0.101 (-4.196)	-0.061 (-3.900)
Transit Score*Pandemic Period	-1.087 (-71.517)	--	--	-0.163 (-5.894)	-0.602 (-39.195)	-0.042 (-1.879)	0.132 (6.499)	0.113 (5.550)
Street Length Density	--	--	1.753 (5.413)	0.619 (1.821)	20.145 (4.764)	5.572 (31.649)	1.529 (8.689)	--
Number of Bike Share Stations	0.042 (2.498)	0.004 (2.502)	0.007 (4.739)	0.010 (7.206)	-0.030 (-1.997)	0.004 (6.205)	--	--
Number of Bus Stops & Subway Stations	3.281 (3.531)	1.573 (26.967)	0.848 (14.188)	0.650 (12.691)	4.669 (6.599)	0.831 (33.792)	0.500 (20.635)	--
Sociodemographic Attributes								
Employment Density	5.162 (7.277)	--	0.166 (2.775)	--	--	--	--	--
Employment Density*Pandemic Period	-1.259 (-20.733)	0.192 (2.520)	0.309 (3.574)	0.533 (6.513)	-0.104 (-4.640)	-0.166 (-4.471)	--	--
Low Income Indicator	-0.798 (-4.815)	--	-0.145 (-11.514)	-0.148 (-11.824)	-0.793 (-5.434)	-0.068 (-8.710)	-0.029 (-3.766)	-0.029 (-3.764)
Weather Attributes								
Snow Depth	--	0.079 (6.232)	0.083 (6.554)	0.060 (4.752)	0.072 (17.459)	--	--	--
Precipitation	--	--	--	--	0.105 (4.651)	--	--	--
Spatial Autoregressive Coefficient	0.638 (268.305)	--	--	--	0.747 (398.807)	--	--	--

5.2.3 Transportation Infrastructure

The parameters associated with the transit score (a measure of serviceability of public transit) have a net positive coefficient on ridehailing demand indicating higher ridehailing demand in zones with higher transit score. The results are perhaps indicating that the ridehailing services and public transport share the service regions. As expected, the net effect of ridehailing demand is found to be lower after pandemic outbreak in zones with higher transit scores. Higher street density is likely to generate more taxi trips during midday and PM peak periods (see Correa et al., 2017 for similar results). On the other hand, street density has an overall net positive impact in TNC demand model. Number of bike share stations has opposing effects in the Taxi and TNC demand models. The zones with higher bike share stations are likely to generate an overall net higher taxi demand, but TNC demand is likely to be lower in these zones. The negative effect of TNC in the zones with higher bike share stations is perhaps indicating the competing nature of different shared mobility systems. With regards to number of bus stops and subway stations, the zones with higher transit facilities are likely to contribute towards higher net ridehailing demand. Such results are perhaps indicative of the first-and-last mile connections through ridehailing for transit facilities (Dey et al., 2021b, Sadowsky and Nelson, 2017).

5.2.4 Sociodemographic Attributes

With regards to sociodemographic attributes, taxi demand is positively associated with employment density and the magnitude of impact is found to be higher in midday period relative to other periods. However, after pandemic, the variable is found to be negatively associated with taxi demand. While the effect of employment density is not found to be significant in the TNC demand model, the interactions of employment density and pandemic effect is found to be significant with an overall negative impact across time periods (and a pronounced reduction for AM peak period). These results are perhaps indicating lower level of commuting activities in these zones during pandemic. As expected, taxi zones in the low-income category are likely to generate less ridehailing demand (similar finding in Dong and Guerra, 2020).

5.2.5 Weather Attributes

During adverse weather, people tend to prefer a door-to-door service and are likely to be reluctant to drive. The results in Table 5.2 are in line with such expectations. Higher level of snow and precipitation depth is likely to generate higher ridehailing demand. For TNC, the effect of snow depth and precipitation is similar across all time periods while taxi demand is relatively higher during daytime. This increase in ridehailing demand in adverse weather is also confirmed by previous studies (Liu et al., 2020).

5.2.6 Spatial Dependency and Correlation

The significant spatial autoregressive coefficient in both models provides evidence of the presence of significant spatial correlation for taxi zone level ridehailing demand. The correlation was captured as inverse of distance for both models indicating that the correlation is higher for proximal zones and reduces as distance increases between the zones.

5.3 Summary

In this chapter, model selection steps based on data fit measures and model estimation results are presented. Additionally, estimation results are discussed in detail to understand the factors affecting ridehailing demand in different time of day. The next chapter will present ridehailing demand projection by the proposed model in this thesis.

CHAPTER 6: RIDEHAILING DEMAND PROJECTION

In this chapter, we provide an illustration of the applicability of the proposed model using conditions for multiple hypothetical scenarios.

6.1 Hypothetical Scenarios

The objective of the proposed research is to provide a mechanism for ridehailing companies and transportation agencies to evaluate changes in demand in response to the evolving case numbers. Toward illustrating this strength, we consider two COVID-19 scenarios in Figure 6.1. An optimistic scenario is considered where COVID-19 cases are expected to stay same without a surge. On the other hand, in a pessimistic scenario, a potential spike in COVID-19 transmission is considered in the middle of August 2021. The spike is assumed to be 60% higher relative to the previous week (2^{nd} week of August) and around 2.5 times more compared to the last week of July. Under these two scenarios, we generate the estimates of ridehailing demand for taxi and TNC services. Figure 6.2 (a to d) provides a temporal summary illustrating how the weekly demand varies across the months of July, August and September. Further in Figure 6.3 and 6.4, we plot the taxi zone level recovery rate of ridehailing demand measured as the ratio of predicted demand for 2^{nd} week of September 2021 and observed demand for the corresponding week from 2019. A value closer to 1 indicates a nearly complete recovery. Several interesting observations can be made from these figures. First, our model represents ridehailing behavior well. As should be expected, there is a slight uptick in ridehailing demand for the optimistic scenario while there is a drop in ridehailing demand with increase in cases as per the pessimistic scenario. Overall, the temporal trends highlight how the proposed model is representing ridehailing demand patterns reasonably. Second, there is a clear preference for TNC services over Taxi services in terms of the recovery patterns. Across both scenarios, the number of taxi zones recovering by 50% for TNC are nearly double the number of zones recovering by 50% for taxi. In the optimistic case, TNC demand recovers more than 50% in 52.70% taxi zones whereas only 28% taxi zones recovered more than 50% in case of taxi demand. On the other hand, in pessimistic case TNC demand is recovered more than 50% demand in 30.62% taxi zones compared to taxi demand which recovers to 50% in around 17.44% taxi zones only. The pandemic might result in further cannibalization of taxi services in New York. Finally, the spatial trends in Figures 4 and 5 highlight the unequal demand recovery patterns for ridehailing in NYC. The recovery rate for taxis is faster in Staten Island and northern Brooklyn while it is lower in other parts of the city. TNC demand appears to be recovering better in various parts of the city including Manhattan, Brooklyn and parts of Queens. Overall, the policy application illustrates the flexibility offered by the proposed model in examining temporal and spatial demand trends for ridehailing services.

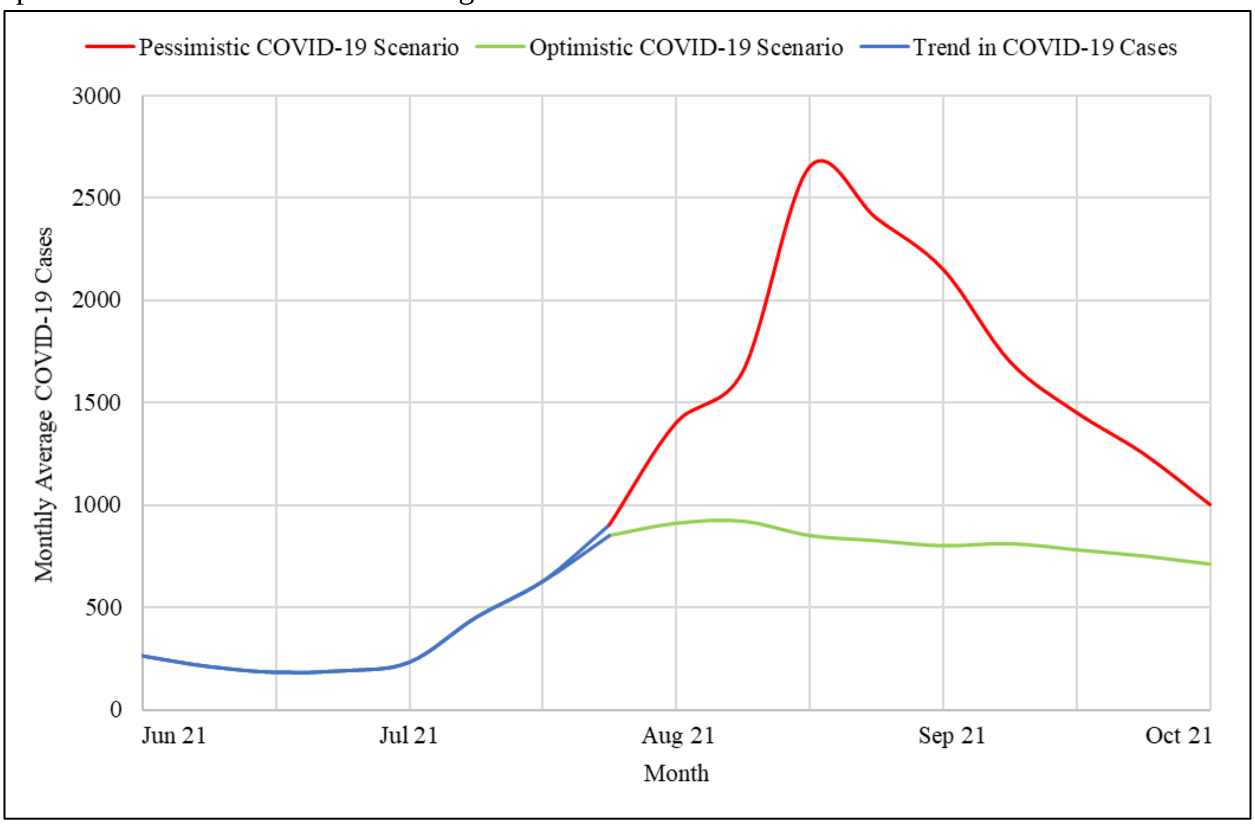


Figure 6.1 Optimistic and Pessimistic scenarios between August 2021 and September 2021

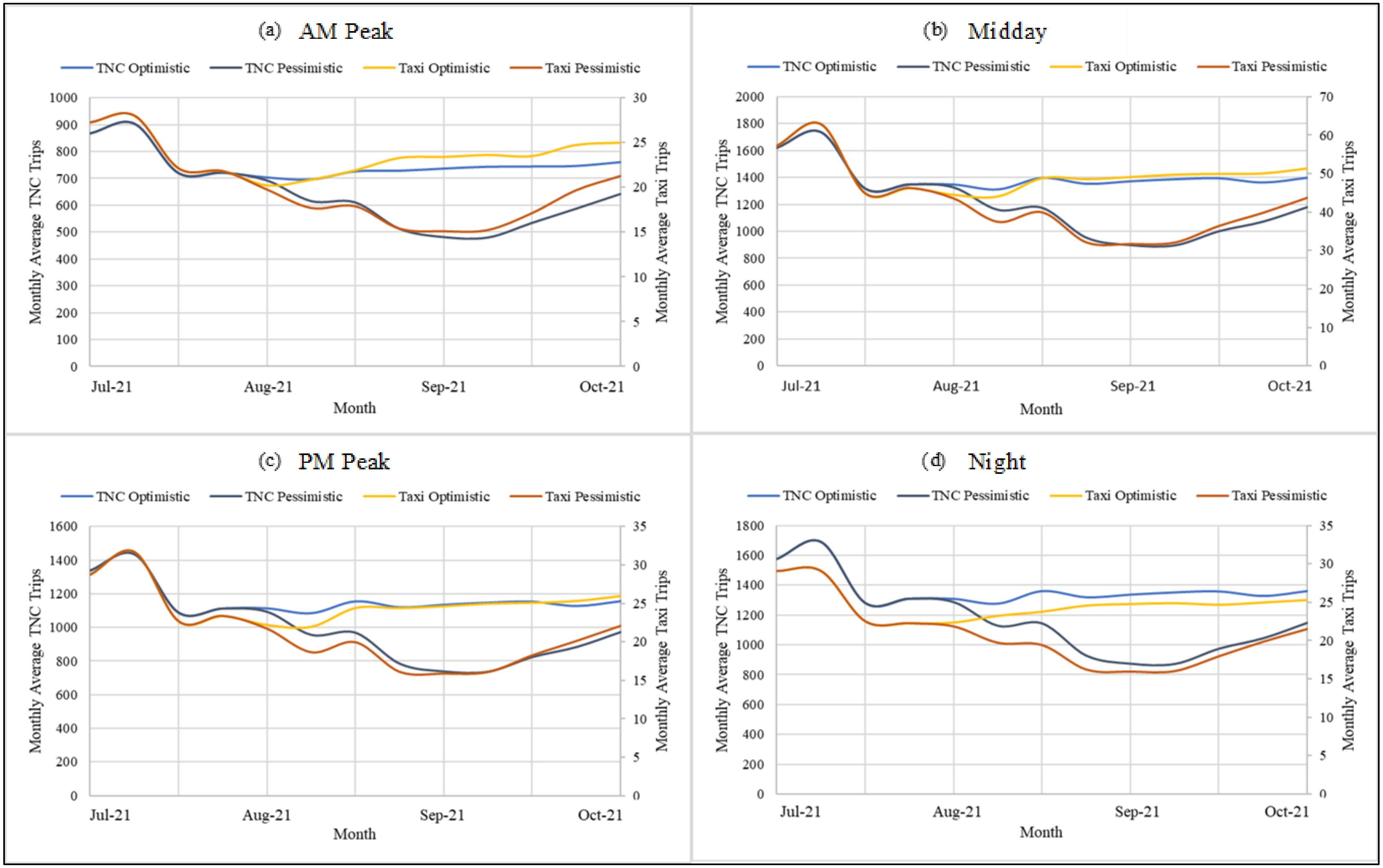


Figure 6.2 Temporal variation in ridehailing demand: (a) AM peak; (b) Midday; (c) PM peak; (d) Night

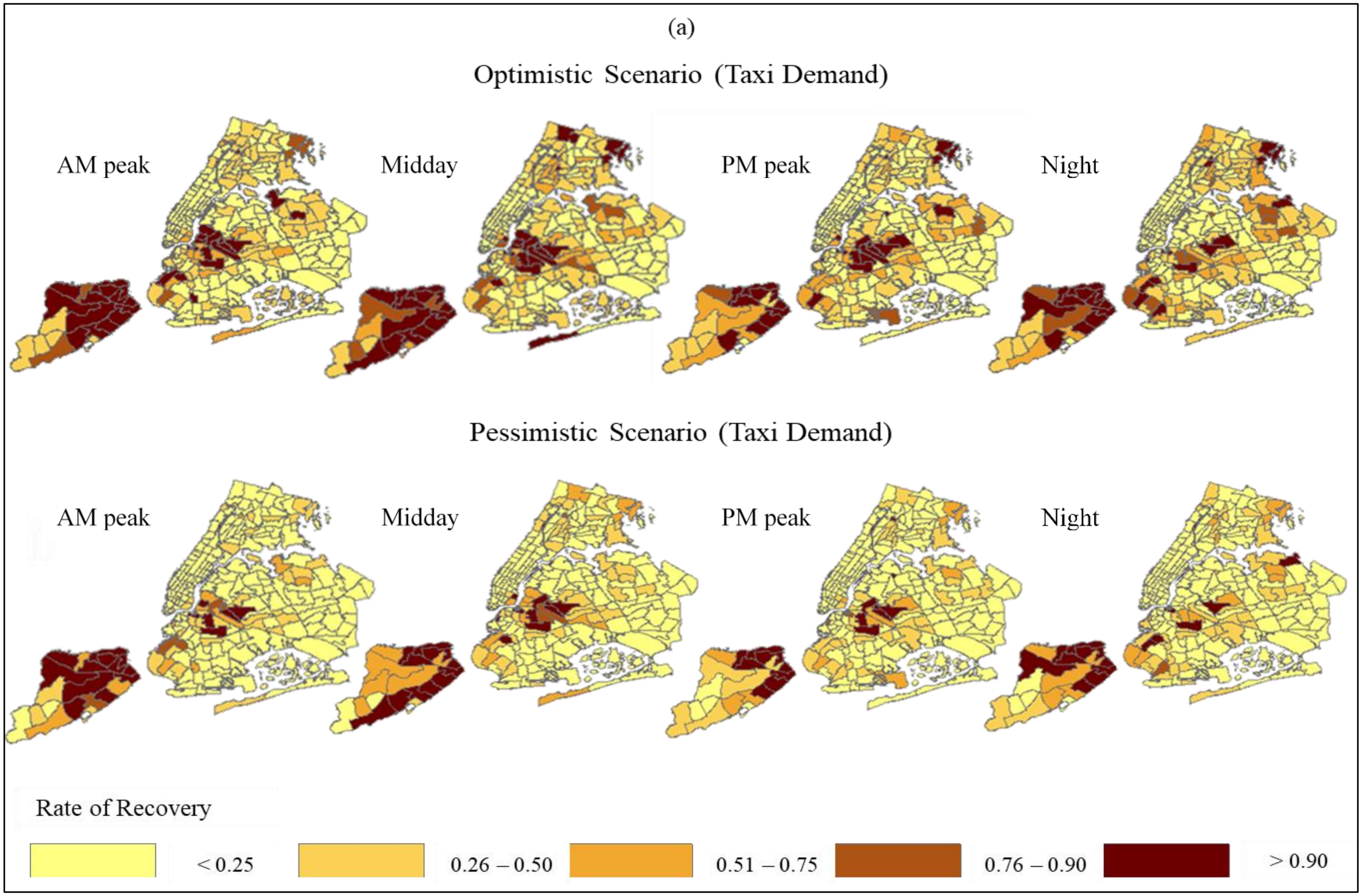


Figure 6.3 Spatial variation in Taxi demand in optimistic and pessimistic scenarios

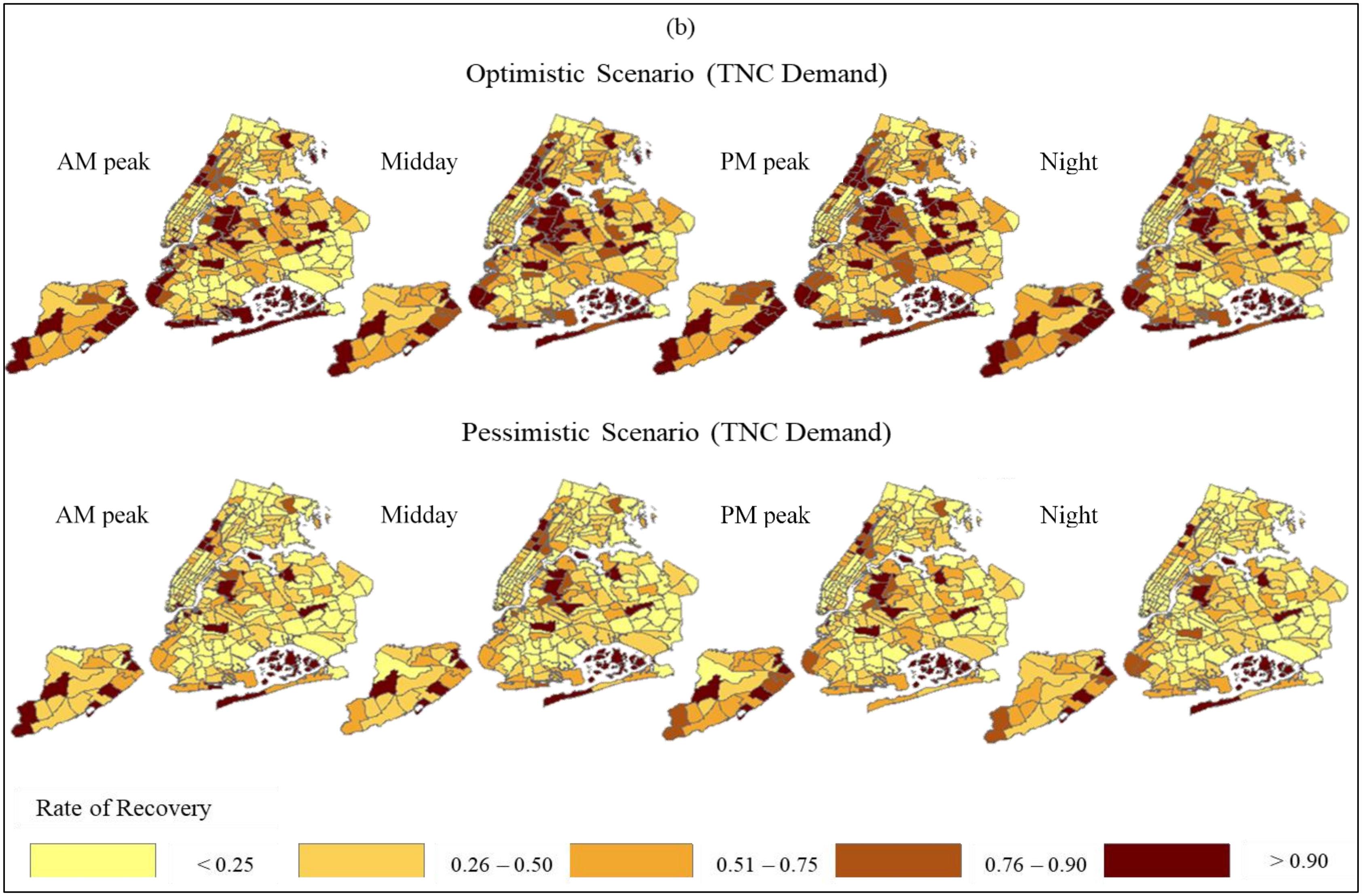
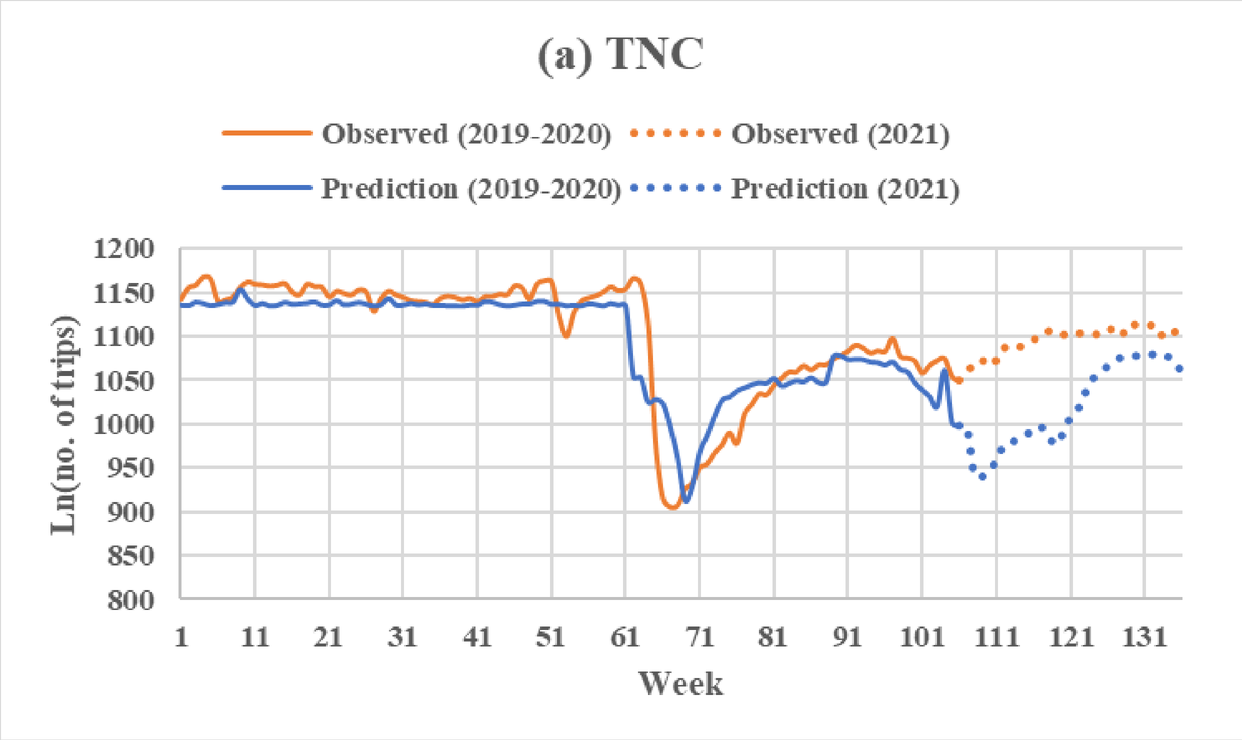


Figure 6.4 Spatial variation in TNC demand in optimistic and pessimistic scenarios

6.2 Comparison with Observed Trends

The spatio-temporal high resolution Taxi and TNC data for months in 2021 are not yet available to conduct a comparison of our model performance to observed data. Hence, we conduct a comparison at the aggregate level i.e., system level observed demand versus system level predicted demand across all taxi zones. The model prediction and observed demand are plotted in Figure 6.5. The project includes 2019, 2020 and first six months of 2021 (about 25 weeks). The prediction exercise reflects reasonable performance of the developed model considering the significant system shock around week 60. The reader would note that the model predictions for 2021 are underpredicted in early months of 2021 (week 107 and later). This is expected because, the model does not consider vaccination rates in our analysis. Vaccinations began late December 2020 and resulted in a large population being vaccinated by February (COVID-19 Vaccine Distribution: The Process, 2021). In the absence of the vaccination data in the model, the model only responds to lower cases. Thus, the model is seen to catch up to the demand through the months of May and June 2021. As high resolution spatio-temporal data become available for 2021, the model estimated can be augmented with vaccination data to further improve the prediction.



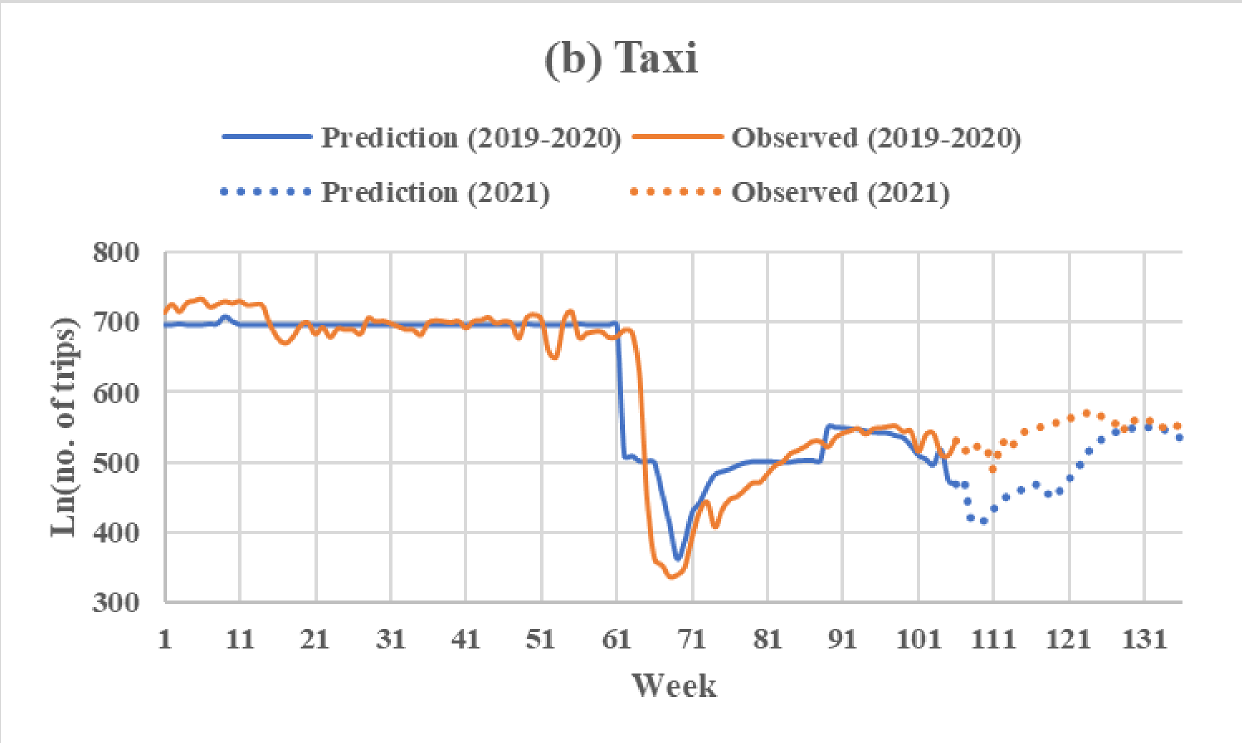


Figure 6.5 Weekly demand prediction of ridehailing services: (a) TNC demand, (b) Taxi demand

6.3 Summary

Results of ridehailing demand projection with hypothetical scenarios and comparison with observed trends are presented in this chapter. The results illustrate a reasonably satisfactory fit from the proposed model. The next chapter will conclude this study by presenting overall findings and limitations of this research.

CHAPTER 7: CONCLUSION

The tremendous growth story of ridehailing demand has met with a system shock in the form of Coronavirus Disease 2019 (COVID-19). In the current research effort, we seek to understand the factors affecting ridehailing demand patterns as the pandemic evolved. Specifically, using high resolution NYC data, we examine weekly ridehailing trip data at taxi zone level for New York city for taxi and TNC services separately. The trip data is categorized across different time periods including AM peak, Midday, PM peak and Night to better understand how the impact of various factors have altered due to COVID-19. Instead of estimating a 4-dimensional multivariate model (for 4 time periods) for taxi or TNC, we resort to a recasting approach that allows us to consider a single pooled model for all 4 time periods (built on our earlier work Bhowmik et al., 2019). Specifically, the data is organized as repeated records of trip demand by time period for each ridehailing service. In terms of the exogenous factors, we consider an exhaustive list of independent variables in the study including COVID-19 case information over time, sociodemographic characteristics, land use and built environment characteristics, transportation infrastructure and weather attributes.

The empirical analysis involved the estimation of three different pooled models for Taxi and TNC service separately including the simple pooled linear regression model and two pooled spatial panel models to accommodate for the spatial heterogeneity: Pooled Spatial Lag or Autoregressive Model (SLPM) and Pooled Spatial Error Model (SEPM). For the spatial models, we adopted several functions of weight matrix and inverse of distance (from the taxi zone of interest to other taxi zones) is found to provide the best data fit and hence is considered in the final model specification. The comparison exercise based on the Bayesian Information Criterion (BIC) clearly highlights the improved performance of the SEPM model over other models for both Taxi and TNC ridehailing demand. The model estimates clearly highlight the impact of COVID-19 cases on ridehailing demand across both services. The model also recovered several important associations with other independent variables including number of restaurants, land use mix, employment density and effect of weather on ridehailing demand. The model results are further augmented with a robust policy analysis to predict potential ridehailing demand for future time periods in response to COVID-19. Specifically, we consider two COVID-19 scenarios: 1) an optimistic scenario where COVID-19 cases are expected to stay same without a surge and 2) a pessimistic scenario where a potential spike in COVID-19 transmission is considered in the middle of August 2021. The results for all the scenarios follow expected trends with the pessimistic scenarios showing lower demand and optimistic scenarios indicating relatively higher or similar demand. Further, we forecast the taxi zone level recovery rate of ridehailing demand in the mid of September 2021 by taking the ratio of predicted demand to the observed demand for the corresponding week from 2019. Interestingly, across both scenarios, the number of taxi zones recovering by 50% for TNC are nearly double the number of zones recovering to 50% for taxi. Overall, the policy application illustrates the flexibility offered by the proposed model in examining temporal and spatial demand trends for ridehailing services. We also compare the model performance to observed data for 2019, 2020 and six months of 2021. The results illustrate reasonable performance of our proposed model (in the absence of consideration of vaccinations).

Our proposed study is not without limitations. We have considered two separate models for Taxi and TNC ridehailing demand. However, it is quite possible to have common unobserved factors influencing these two services and exploring such a correlation might be an interesting avenue for future research. The proposed model can be updated with full 2021 data once it becomes available along with vaccination rates in the region. Further, we did not consider the day of the working week (weekday/weekend) while estimating the model. In future, it will be interesting to explore such variation in ridehailing demand across the day of the week.

LIST OF REFERENCES

Alemi, F., Circella, G., Mokhtarian, P., Handy, S., 2018. On-Demand Ride Services in California: Investigating the Factors Affecting the Frequency of Use of Uber/Lyft.

Archived: WHO Timeline - COVID-19 [WWW Document], 2021. URL https://www.who.int/news/item/27-04-2020-who-timeline---covid-19 (accessed 7.14.21).

Beck, M.J., Hensher, D.A., 2020. Insights into the impact of COVID-19 on household travel and activities in Australia – The early days under restrictions. Transp. Policy 96, 76–93. https://doi.org/10.1016/J.TRANPOL.2020.07.001

Bhaduri, E., Manoj, B.S., Wadud, Z., Goswami, A.K., Choudhury, C.F., 2020. Modelling the effects of COVID-19 on travel mode choice behaviour in India. Transp. Res. Interdiscip. Perspect. 8, 100273. https://doi.org/10.1016/J.TRIP.2020.100273

Bhowmik, T., Tirtha, S.D., Iraganaboina, N.C., Eluru, N., 2021. A comprehensive analysis of COVID-19 transmission and mortality rates at the county level in the United States considering socio-demographics, health indicators, mobility trends and health care infrastructure attributes. PLoS One 16, 1–15. https://doi.org/10.1371/journal.pone.0249133

Bhowmik, T., Yasmin, S., Eluru, N., 2019. Do we need multivariate modeling approaches to model crash frequency by crash types? A panel mixed approach to modeling crash frequency by crash types. Anal. Methods Accid. Res. 24, 100107. https://doi.org/10.1016/J.AMAR.2019.100107

Chan, N.D., Shaheen, S.A., 2011. Ridesharing in North America: Past, Present, and Future. http://dx.doi.org/10.1080/01441647.2011.621557 32, 93–112. https://doi.org/10.1080/01441647.2011.621557

Clewlow, Regina R. Mishra, G.S., 2017. Disruptive Transportation: The Adoption, Utilization, and Impacts of Ride-Hailing in the United States. Genome 44, 401–412.

Correa, D., Xie, K., Ozbay, K., 2017. Exploring the taxi and Uber demands in New York City: An empirical analysis and spatial modeling. Transp. Res. Board 96th Annu. Meet.

COVID-19 Map - Johns Hopkins Coronavirus Resource Center [WWW Document], 2021. URL https://coronavirus.jhu.edu/map.html (accessed 7.13.21).

COVID-19 Vaccine Distribution: The Process | HHS.gov [WWW Document], 2021. URL https://www.hhs.gov/coronavirus/covid-19-vaccines/distribution/index.html (accessed 12.4.21).

COVID Live - Coronavirus Statistics - Worldometer [WWW Document], 2022. URL https://www.worldometers.info/coronavirus/ (accessed 2.7.22).

Daily Summaries Station Details: NY CITY CENTRAL PARK, NY US, GHCND:USW00094728 | Climate Data Online (CDO) | National Climatic Data Center (NCDC) [WWW Document], 2021. URL https://www.ncdc.noaa.gov/cdo-web/datasets/GHCND/stations/GHCND:USW00094728/detail (accessed 7.30.21).

Dey, B.K., Tirtha, S.D., Eluru, N., 2021a. Transport Networking Companies Demand and Flow Estimation in New York City. Transp. Res. Rec. J. Transp. Res. Board 2675, 139–153. https://doi.org/10.1177/03611981211000752

Dey, B.K., Tirtha, S.D., Eluru, N., Konduri, K.C., 2021b. Transformation of ridehailing in New York City: A quantitative assessment. Transp. Res. Part C Emerg. Technol. 129, 103235. https://doi.org/10.1016/J.TRC.2021.103235

Dimitriou, L., Christodoulou, C., Kourti, E., Christodoulou, S., 2016. Dynamic monitoring of taxi demand profiles, utilizing location-specific information in large metropolitan areas. Proc. 18th Mediterr. Electrotech. Conf. Intell. Effic. Technol. Serv. Citizen, MELECON 2016 1–

6. https://doi.org/10.1109/MELCON.2016.7495409

Dong, X., Guerra, E., 2020. A Philadelphia ride-share story: An Investigation of ride-share's impact on transit.

Du, J., Rakha, H.A., 2020. COVID-19 Impact on Ride-hailing: The Chicago Case Study. Findings 2018, 1–7. https://doi.org/10.32866/001c.17838

Elhorst, J.P., 2014a. Matlab Software for Spatial Panels. Int. Reg. Sci. Rev. 37, 389–405. https://doi.org/10.1177/0160017612452429

Elhorst, J.P., 2014b. Spatial Panel Data Models 37–93. https://doi.org/10.1007/978-3-642-40340-8_3

Elhorst, J.P., 2003. Specification and Estimation of Spatial Panel Data Models: Int. Reg. Sci. Rev. 26, 244–268. https://doi.org/10.1177/0160017603253791

Espín Noboa, L., Lemmerich, F., Singer, P., Strohmaier, M., 2016. Discovering and Characterizing Mobility Patterns in Urban Spaces 537–542. https://doi.org/10.1145/2872518.2890468

Explore Census Data [WWW Document], n.d. URL https://data.census.gov/cedsci/ (accessed 7.19.21).

Faghih-Imani, A., Anowar, S., Miller, E.J., Eluru, N., 2017. Hail a cab or ride a bike? A travel time comparison of taxi and bicycle-sharing systems in New York City. Transp. Res. Part A Policy Pract. 101, 11–21. https://doi.org/10.1016/J.TRA.2017.05.006

Faghih-Imani, A., Eluru, N., 2016. Incorporating the impact of spatio-temporal interactions on bicycle sharing system demand: A case study of New York CitiBike system. J. Transp. Geogr. 54, 218–227. https://doi.org/10.1016/j.jtrangeo.2016.06.008

Feigon, S., Murphy, C., 2016. TCRP RESEARCH REPORT 188. Shared Mobility and the Transformation of Public Transit. Shar. Mobil. Transform. Public Transit.

First Case of COVID-19 in NYC, First Death Reported in Washington State - Infectious Disease Advisor [WWW Document], 2020. URL https://www.infectiousdiseaseadvisor.com/home/topics/respiratory/first-case-of-covid-19-in-nyc-first-death-reported-in-washington-state/ (accessed 7.14.21).

Furuhata, M., Dessouky, M., Ordóñez, F., Brunet, M.E., Wang, X., Koenig, S., 2013. Ridesharing: The state-of-the-art and future directions. Transp. Res. Part B Methodol. 57, 28–46. https://doi.org/10.1016/J.TRB.2013.08.012

Gerte, R., Konduri, K.C., Eluru, N., 2018. Is There a Limit to Adoption of Dynamic Ridesharing Systems? Evidence from Analysis of Uber Demand Data from New York City: https://doi.org/10.1177/0361198118788462 2672, 127–136. https://doi.org/10.1177/0361198118788462

Habib, K.N., 2019. Mode choice modelling for hailable rides: An investigation of the competition of Uber with other modes by using an integrated non-compensatory choice model with probabilistic choice set formation. Transp. Res. Part A Policy Pract. 129, 205–216. https://doi.org/10.1016/J.TRA.2019.08.014

Komanduri, A., Wafa, Z., Proussaloglou, K., Jacobs, S., 2018. Assessing the Impact of App-Based Ride Share Systems in an Urban Context: Findings from Austin: https://doi.org/10.1177/0361198118796025 2672, 34–46. https://doi.org/10.1177/0361198118796025

Kumar Dey, B., Anowar, S., Eluru, N., 2021. A framework for estimating bikeshare origin destination flows using a multiple discrete continuous system. Transp. Res. Part A Policy Pract. 144, 119–133. https://doi.org/10.1016/J.TRA.2020.12.014

Lahkar, P., Murray-Tuite, P., Hancock, K., 2018. Familiarity and Use of Transportation Network

Company (TNC) Services in Virginia.

Lavieri, P.S., Dias, F.F., Juri, N.R., Kuhr, J., Bhat, C.R., 2018. A Model of Ridesourcing Demand Generation and Distribution: https://doi.org/10.1177/0361198118756628 2672, 31–40. https://doi.org/10.1177/0361198118756628

Li, W., Shalaby, A., Habib, K.N., 2021. Exploring the correlation between ride-hailing and multimodal transit ridership in toronto. Transp. 2021 1–25. https://doi.org/10.1007/S11116-021-10193-5

Liu, Q., Ding, C., Chen, P., 2020. A panel analysis of the effect of the urban environment on the spatiotemporal pattern of taxi demand. Travel Behav. Soc. 18, 29–36. https://doi.org/10.1016/J.TBS.2019.09.003

Liu, Y., Singleton, A., Arribas-bel, D., Chen, M., 2021. Identifying and understanding road-constrained areas of interest (AOIs) through spatiotemporal taxi GPS data: A case study in New York City. Comput. Environ. Urban Syst. 86, 101592. https://doi.org/10.1016/J.COMPENVURBSYS.2020.101592

Loa, P., Nurul Habib, K., 2021. Examining the influence of attitudinal factors on the use of ride-hailing services in Toronto. Transp. Res. Part A Policy Pract. 146, 13–28. https://doi.org/10.1016/J.TRA.2021.02.002

New York City COVID-19 Economic Impact Update - Partnership for New York City [WWW Document], 2020. URL https://pfnyc.org/research/new-york-city-covid-19-economic-impact-update/ (accessed 7.22.21).

Nian, G., Peng, Bozhezi, Sun, D. (Jian), Ma, W., Peng, Bo, Huang, T., 2020. Impact of COVID-19 on Urban Mobility during Post-Epidemic Period in Megacities: From the Perspectives of Taxi Travel and Social Vitality. Sustain. 2020, Vol. 12, Page 7954 12, 7954.

https://doi.org/10.3390/SU12197954

Nie, W.P., Zhao, Z.D., Cai, S.M., Zhou, T., 2021. Understanding the urban mobility community by taxi travel trajectory. Commun. Nonlinear Sci. Numer. Simul. 101, 105863. https://doi.org/10.1016/J.CNSNS.2021.105863

NYC Open Data - [WWW Document], 2021. URL https://opendata.cityofnewyork.us/ (accessed 7.19.21).

NYC Taxi & Limousine Commission. Taxi and Ride hailing Usage in New York City [WWW Document], 2021. URL https://toddwschneider.com/dashboards/nyc-taxi-ridehailing-uber-lyft-data/ (accessed 7.11.21).

NYC Taxi Zones | NYC Open Data [WWW Document], 2021. URL https://data.cityofnewyork.us/Transportation/NYC-Taxi-Zones/d3c5-ddgc (accessed 10.18.21).

NYS GIS Clearinghouse - NYS GIS Program Office - NYS Tax Parcels [WWW Document], 2021. URL http://gis.ny.gov/gisdata/inventories/details.cfm?DSID=1300 (accessed 7.30.21).

Rahman, M., Yasmin, S., Faghih-Imani, A., Eluru, N., 2021. Examining the Bus Ridership Demand: Application of Spatio-Temporal Panel Models. J. Adv. Transp. 2021. https://doi.org/10.1155/2021/8844743

Rayle, L., Dai, D., Chan, N., Cervero, R., Shaheen, S., 2016. Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. Transp. Policy 45, 168–178. https://doi.org/10.1016/J.TRANPOL.2015.10.004

Sabouri, S., Park, K., Smith, A., Tian, G., Ewing, R., 2020. Exploring the influence of built environment on Uber demand. Transp. Res. Part D Transp. Environ. 81, 102296. https://doi.org/10.1016/J.TRD.2020.102296

Sadowsky, N., Nelson, E., 2017. The Impact of Ride-Hailing Services on Public Transportation Use: A Discontinuity Regression Analysis. Econ. Dep. Work. Pap. Ser.

Sun, C., Edara, P., 2015. Is Getting an Uber-Lyft from a Sidecar Different from Hailing a Taxi?: Current Dynamic Ridesharing Controversy. https://doi.org/10.3141/2536-08 2536, 60–66. https://doi.org/10.3141/2536-08

Tan, L., Ma, C., 2021. Choice behavior of commuters' rail transit mode during the COVID-19 pandemic based on logistic model. J. Traffic Transp. Eng. (English Ed. 8, 186–195. https://doi.org/10.1016/J.JTTE.2020.07.002

Thompson, C.N., 2020. COVID-19 Outbreak — New York City, February 29–June 1, 2020. MMWR. Morb. Mortal. Wkly. Rep. 69, 1725–1729. https://doi.org/10.15585/MMWR.MM6946A2

TLC Trip Record Data - TLC [WWW Document], 2021. URL https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page (accessed 7.19.21).

Wang, H.W., Peng, Z.R., Lu, Q.C., Sun, D.J., Bai, C., 2018. Assessing effects of bus service quality on passengers' taxi-hiring behavior. Transport 33, 1030–1044. https://doi.org/10.3846/16484142.2016.1275786

Wang, S., Noland, R.B., 2021. Variation in ride-hailing trips in Chengdu, China. Transp. Res. Part D Transp. Environ. 90, 102596. https://doi.org/10.1016/J.TRD.2020.102596

Yu, H., Peng, Z.R., 2019. Exploring the spatial variation of ridesourcing demand and its relationship to built environment and socioeconomic factors with the geographically weighted Poisson regression. J. Transp. Geogr. 75, 147–163. https://doi.org/10.1016/J.JTRANGEO.2019.01.004

Zhang, B., Chen, S., Ma, Y., Li, T., Tang, K., 2020. Analysis on spatiotemporal urban mobility based on online car-hailing data. J. Transp. Geogr. 82, 102568. https://doi.org/10.1016/J.JTRANGEO.2019.102568

Zhang, Yuanyuan, Zhang, Yuming, 2018. Exploring the Relationship between Ridesharing and Public Transit Use in the United States. Int. J. Environ. Res. Public Heal. 2018, Vol. 15, Page 1763 15, 1763. https://doi.org/10.3390/IJERPH15081763