Effects of Built Environment and Weather on Demands for Transportation Network Company Trips

Md Sami Hasnine, Ph.D., Corresponding author
Assistant Professor
Department of Civil and Environmental Engineering
Howard University
2300 Sixth Street, NW #1026
Washington, DC, USA, 20059
Email: mdsami.hasnine@howard.edu

Jason Hawkins, Ph.D.
Postdoctoral Researcher
Cockrell School of Engineering
The University of Texas at Austin
Email: jason.hawkins@mail.utoronto.ca

Khandker Nurul Habib, Ph.D., PEng.
Percy Edward Hart Professor in Civil & Mineral Engineering
Associate Director, Data Management Group (DMG)
Department of Civil and Mineral Engineering
University of Toronto
35 St George Street, Toronto, ON, Canada, M5S1A4
Email: khandker.nurulhabib@utoronto.ca

ABSTRACT
This paper investigates the effects of the built environment and weather on the demands for transportation network companies (TNC) in Toronto. The research is based on a historical dataset of Uber trips from September 2016 to September 2018 in Toronto. A wide range of built environments, sociodemographic, and weather data are generated at the dissemination area-level and fused with the monthly aggregated Uber dataset. To provide insight into the underlying factors that affect TNC demand, a series of aggregate demand models are estimated using log-transformed constant elasticity demand functions, with consideration of the seasonal lag effect. To capture the weather effect, an autoregressive moving average model is estimated for the downtown core of Toronto. The model results show that the influence of lagged ridership and seasonal lag effect have a positive correlation with TNC demand. The trip generation and attraction models reveal that TNC trips increase where when the commuting trip duration is longer than 60 minutes. It is found that the number of apartments in a dissemination area is positively correlated with TNC trip generation, while the number of single-detached houses has a negative correlation. The time-series model indicates that temperature and total daily precipitations are positively correlated with TNC demand. Due to the lack of comprehensive data sources on the Uber and Lyft ridership, the policymakers often struggle to make evidence-based policy recommendations to regulate such disruptive technologies. The series of models presented in this study will help us better understand the potential users of transportation network companies (TNC) and the effects of land use, built environment and weather on transportation network company trips.

Keywords: Transportation Network Companies (TNC), Aggregate Demand, Trip generation, Trip Attraction, Time Series Model

1. INTRODUCTION

Transportation Network Companies (TNC), such as Lyft and Uber, have garnered increasing attention among urban dwellers as an affordable and reliable mode of transportation. In the City of New York, there was a 200% increase in TNC pickups from 2015 to 2017 (Gerte et al. 2019). According to a recent study by the City of Toronto, ride-hailing trips comprise 5-8% of downtown Toronto traffic and a 196% growth in ride-hailing trips was observed in just eighteen months (Big Data Innovation Team 2019). Speculation exists that TNC can reduce transit ridership and cause more traffic congestion. Graehler, Mucci, and Erhardt (2019) report a 1.7% decrease in bus ridership and a 1.27% decrease in heavy rail ridership, according to a longitudinal dataset from twenty-two US cities. Though it is unclear whether the increasing rate of TNC market penetration caused this decrease in transit ridership, there could be other complementary factors that caused this decline in transit ridership, such as increased car and bike ownership or unreliable transit service. Municipal and federal governments in different parts of the world are trying to regulate and limit the number of TNCs in operation. Due to the lack of adequate data on Uber and Lyft ridership, policymakers often struggle to make evidence-based policy recommendations to regulate such disruptive technologies.

There have been a few notable attempts to study the factors that influence individuals to choose TNC for different types of trips, such as recreational and commuting (Habib 2019). A series of studies on Uber data in the City of New York show that weather and built environment attributes are significant predictors of Uber demand (Gerte et al. 2019; Gerte, Konduri & Eluru 2018). Another set of studies are conducted on a TNC operator RideAustin, an Austin-based TNC (Lavieri et al. 2018; Dias et al. 2018). Lavieri et al. (2018) find a higher number of TNC trips are generated near the University of Texas Austin, suggesting that trip purposes in these zones are conducive to ride-hailing. A few studies try to capture the relationship between sociodemographic variables and TNC demand. These studies postulate that educated, young, and professional individuals are the main user of TNCs (Spurlock et al. 2019; Alemi et al. 2019; Hall, Palsson & Price 2018; Alemi et al. 2018).

TNC serve all travel markets (commuting and non-commuting) in Toronto, but the main trip attractor in downtown Toronto is recreation, particularly late-night trips on Fridays and Saturdays. Since speculation about increasing congestion and reduction in transit ridership is possible, it is deemed necessary that we improve our understanding of TNC demand variation in large metropolitan cities like Toronto based on empirical evidence. Micro-level studies capture the individual-level choice dimension, whereas macro-level studies capture the region-wide phenomenon in a specific period of time (Chen, Varley & Chen, 2011). Thus, macro-level studies are more effective in the context of TNC to better understand how TNC demand is influenced by built environment and weather. Compared to the previous research efforts that mainly looked at the aggregate demand in the US cities only, our approach exploits a comprehensive multiyear dataset capturing the surge in TNC demand in the City of Toronto.

This paper uses historical data of Uber ridership in Toronto from September 2016 to September 2018 and estimates a set of aggregate demand models to investigate the effect of sociodemographic, weather and firmographic attributes on the Uber ridership demand. The proposed modelling framework also allows capturing the seasonal lag effect of TNC demand, something that lacks in the past research. The paper contributes to the existing literature by using the entire population of Uber trips for a longer time period (two years versus a few weeks to months in other studies) for a megacity like the City of Toronto. Access to these data allows us to examine seasonal variation in trip rates and perform other more detailed analyses of trips, which is not possible with smaller datasets.

The rest of this paper is outlined as follows. The next section highlights a review of the recent studies on transportation network companies in various parts of the world. The subsequent sections describe data fusion methods. Then the modelling framework and results of the empirical models are discussed elaborately. The conclusions section connects the key findings across the different sections of this paper and provides some policy recommendations.

2. LITERATURE REVIEW

Despite the incipient stage of TNC adoption, a range of studies examine the demand for these services. As most TNC data are proprietary, research is often based on survey data. However, in some cases, researchers have been able to gain access to trip-level data from TNC service providers. We classify the literature according to the type of data used as this tends to determine the methods of analysis.

Studies Based on Survey Data

In the case that only survey data is available, analysis tends to focus on the attributes of TNC users. The most extensive such analysis is summarized in a series of papers by Alemi et al (2019, 2018a, 2018b). They use an attitudinal survey administered in California to determine the preferences of Millennials and Generation X. The probability of a person using TNC is then modelled as a latent utility based on their stated preference. Most surveys find that potential users are young, highly educated, and professionals (Spurlock et al. 2019; Alemi et al. 2019; Hall, Palsson & Price 2018; Alemi et al. 2018).

Lavieri and Bhat (2019a) administer a survey in the Dallas-Fort Worth area. They first ask respondents whether they have experience with TNC and those with experience are asked a series of retrospective trip questions. A generalized heterogenous data model (GHDM) is used to jointly estimate residential location, vehicle availability, TNC experience, and TNC frequency of use. The model includes latent constructs for privacy sensitivity, tech-savviness, and propensity towards variety-seeking or green lifestyles. One pattern that emerges in many of these studies is their use as tools to anticipate autonomous vehicle adoption (Lavieri and Bhat, 2019b). The standard hypothesis is that stated preferences for making TNC and, particularly, shared TNC trips is indicative of a propensity to use autonomous vehicles when they enter the consumer market. Gao et al. (2019) make a similar argument of TNC being a precursor to autonomous vehicles. However, their survey includes a series of stated preference experiments, which tend to provide more statistical control to the researcher (Louviere et al., 2010).

Wang et al. (2018) develop a set of technology acceptance models to examine the adoption of ride-hailing. They define six latent measures: personal innovativeness, perceived ease of use, perceived usefulness, perceived risk, environmental awareness, and behavioural intention. A survey is administered to 426 participants and factor analysis performed for the six latent measures. Personal innovativeness arises as a significant positive influence on actual ride-hailing trip frequency, while perceived risk has a negative correlation with making trips. Perceived ease of use is not found significant in their model. The authors find that perceived ease of use has an indirect effect through perceived usefulness of ride-hailing and suggest that it is only when the service has a perceived benefit that its ease of use influences the probability of making a trip.

A second class of survey-based studies use general travel surveys. These studies have the advantage of being based on reported trips from travel diaries. However, they often rely on smaller samples than purpose-built surveys. Dias et al. (2019) use the Puget Sound household travel survey collected in 2015 and 2017 to examine the rate of TNC adoption. They use a joint binary probit-ordered probit model of survey year and trip frequency, respectively. It is found that demographic characteristics have a diminishing effect on TNC use over time. Wigginton Conway et al. (2018) use a panel of the U.S. National household travel survey between 1995 and 2017 to examine trends in taxi use and its relationship with TNC. They do not find the same trend of increasing equity as Dias et al. A similar analysis is carried out in Toronto by Ozonder and Miller (2019) using a panel of the local travel survey. They confirm the results of Wigginton Conway et al. that the average TNC user has a higher income than taxi users. Habib (2019) uses the same travel survey data for Toronto to examine mode choice competition. His analysis uses a novel form of choice set generation model to capture the factors that influence consideration of TNC as a modal alternative.

Studies Based on Trip Data

A second stream of literature focuses on examines TNC demand using trip-based data. This approach can explicitly capture the aggregate demand for TNC trips. However, the data generally lack demographic information due to privacy concerns. As such, studies generally focus on aggregation demand as a function of built form, weather conditions, and other location data. Gerte et al. perform analysis for the City of New York (2018). Their analysis is based on a dataset obtained through a Freedom of Information Law (FOIL) request by the company FiveThirtyEight for six months in 2014 (April to September) and six months in 2015 (January to June). Pickup locations are aggregated to latitude and longitude in the 2014 data and taxi zones in the 2015 data. The authors aggregate all data to match taxi zones and sum daily to weekly totals, giving a final dataset with 69 zones and 49 weeks. The TNC data are fused with the weather, land use, and sociodemographic variables and a panel model estimated for demand generation. A key finding in this study is the negative correlation between demand over time and percent of land devoted to residential uses in the zone. Gerte et al. postulate a diminishing growth in demand within existing coverage areas, even while overall demand for TNC trips grows across the city (and to other cities). Increasing access to data and more extended time series should allow us to test the continued existence of this pattern.

In a follow-up study, the same authors examine the demand for shared modes (TNC, taxi, bike share, subway) in the City of New York with an expanded dataset (Gerte et al. 2019). They use TNC and taxi data provided by the Taxi and Limousine Commission and combine it with Citi Bike, New York City subway, weather, and permitted event data. The analysis begins with descriptive plots of trips per day by the four shared modes over the period from January 2015 to May 2017. Gerte et al. find that taxi demand fell over the study period but that the rise in TNC trips can not fully explain the decline. A dynamic linear model is developed for daily TNC demand as a function of seasonal factors and the increasing penetration of the mode (captured through drift and autoregressive coefficients). The model suggests that precipitation and increased subway demand are both positively correlated with TNC demand.

A series of studies have also been conducted in Austin, TX using data made available by the local TNC operator RideAustin (Lavieri et al. 2018; Dias et al. 2018). This company entered the Austin market following the departure of Uber and Lyft in response to disputes over local regulations. In the first study, demand generation and distribution models are estimated based on data from June 2016 to April 2017 (Lavieri et al. 2018). It is estimated that RideAustin represents about one-third of TNC trips made in Austin. The data are aggregated to Traffic Analysis Zones (TAZs), which gives 458 spatial units in the study area. A spatial bivariate count model is estimated for the number of trips generated on a weekday and weekend day, with spatial autocorrelation tested between adjacent TAZ. They find a higher rate of trips near the University of Texas Austin, suggesting both that trip purposes in these zones are conducive to ride-hailing and that the student demographic is more likely to make trips by this mode. A fractional split model is estimated for trip distribution to examine the question of trip attraction. Lavieri et al. (2018) find that trips to University of Texas Austin decrease on weekend days, but retail employment is positively correlated with trip attraction regardless of the day of the week, with the authors suggesting retail employment represents a measure of the attractiveness of a TAZ for out-of-home activities. Interestingly, this study does not find a strong attractiveness of CBD zones, which suggests that ride-hailing is not a popular mode of commuting in Austin.

In addition to trip generation, the RideAustin data are used to explore trip purpose imputation. Dias et al. (2018) fuse these data with parcel-level land use data and use the zoning of each parcel as a proxy for the origin and destination activities. For example, a trip originating in a residential zone and destined for a commercial zone is classified as a shopping trip. Home locations were inferred from the most frequently visited residential zone, using census block groups as the spatial unit of analysis. From these inferred data, the authors estimate a multivariate ordered probit model for trip frequency. This is possible because the RideAustin data contains anonymized user IDs, which are not available in most other TNC datasets. The inferred home locations are combined with American Community Survey data to impute sociodemographic characteristics for each user. Model results suggest that these sociodemographic factors are essential in determining both the frequency of purpose of trips. Higher-income users tend to make more trips to the airport and recreational locations, while lower-income users appear to be more likely to use ride-hailing for shopping and commuting trips.

In some cases, researchers could obtain TNC data through the use of calls to TNC API (Grahn et al., 2020; Shokoohyar et al., 2020; Sun and Ding, 2019). Shokoohyar et al. (2020) focus on accessibility in Philadelphia, while Grahn et al. (2020) focus on substitution between TNC and public buses in Pittsburgh. Sun and Ding provide a case study outside North America using data collected for Didi in Shanghai. They find a strong effect of residential and commercial land use on-demand, as well as a strong correlation between weather and demand. Results suggest that TNC may be more complementary with the metro service than the bus. Their data allows them to distinguish between shared and single-user rides, as well as identify specific drivers. These features are unavailable in most of the North American datasets. In another study, Sun et al. (2018) try to capture the relationship between built environment attributes and road traffic emissions in terms of ride-hailing service DiDi.

Factors affecting trip-making are important to TNC operators. A study by Uber outlines a neural network model for forecasting trip generation following a major event (e.g., sports game or public celebration) (Laptev et al., 2017). Their work is motivated by decreasing wait times for users, by reallocating drivers in anticipation of known major demand spikes. Uber does not suffer from the same data sparsity, having access to anonymized rider and driver data for hundreds of cities. However, the specific focus of the study limits the available data. They estimate a long short-term memory (LSTM) neural network; a common approach applied in machine learning to time series analysis. Such approaches help to overcome the need to frequently retrain the model when applied to a highly heterogeneous time series, such as extreme event forecasting. They introduce an autoencoder, which processes historical data for input into a second stage model that includes new demand data. The proposed technique offers a 2-18% improvement over the existing proprietary univariate model used by Uber.

In a recent paper, Chen et al. (2020) identified that various trip characteristics, sociodemographic attributes, and land use attributes affect individuals' mode choice decisions. The authors estimated two separate Binomial logit models. The first model has two alternatives: metro and taxi, and the second model also has two alternatives: mobibike and taxi. This study reveals that individuals are likely to use taxis for official business trips, which are done during off-peak hours.

Table 1 Comparison of studies by approach and data sources

Study	Approach	Data source	Key variables
Alemi et al., 2019, 2018a, 2018b	Probit/logit regression	Targeted survey in California	• Individual demographics • Attitudinal factors
Grahn et al., 2020	Linear regression	TNC trips for seven months in 2016-2017 in Pittsburgh, PA	
Gerte et al., 2018; Gerte and Konduri, n.d.	Linear regression	TNC trips for six months in 2014 and six months in 2015 in New York, NY	• Weather • Zonal demographics • Built environment
"Fusing Multiple Sources of Data to Understand Ride-Hailing Use," n.d.	Probit regression	TNC trips for in Austin, TX	
Lavieri and Bhat, 2019a, 2019b	Structural equation model	Household travel survey in Dallas-Fort Worth, TX	• Individual demographics

			• Attitudinal factors
Gao et al., 2019	Discrete choice	Stated preference experiment in United States	• Individual demographics • Travel mode attributes
Ozonder and Miller, 2019	Descriptive statistics	Household travel survey in Toronto, ON	• Time of trip • Origin/destination location of trip
Conway et al., 2018	Logit regression	Household travel survey	• Individual demographics
Dias et al., 2019, 2017	Probit regression	Household travel survey in Puget Sound, WA	• Individual demographics
Wang et al., 2018	Technology acceptance model	Targeted survey in China	• Individual demographics
Spurlock et al., 2019	Linear regression	Targeted survey in San Francisco, CA	• Individual demographics
Habib, 2019	Discrete choice	Household travel survey in Toronto, ON	• Individual demographics
Sun and Ding, 2019	Two-level growth model	TNC trips for four months in Shanghai	• Weather • Built environment
Shokoohyar et al., 2020	Spatial lags regression	TNC trips for two months in Philadelphia, PA	• Zonal demographics
Grahn et al., 2019	Linear regression	Household travel survey in Pittsburgh, PA	• Individual demographics
Laptev et al., 2017	LSTM neural network	TNC trips for large sample (several years and cities) in North America	• Unknown

The contributions of this study in comparison to the previous literature are outlined below. It is noted that the objective of these points is not to criticize any previous studies but to indicate how our study complements the previous studies.

Limitations of the previous studies

- Table 1 reveals that many past studies relied on household-level travel surveys (1995–2017) where in most of the cases, Uber and Lyft trips are not even reported (Conway et al., 2018; Ozonder and Miller, 2019). Authors extracted those individuals' records where individuals reported taxi as a mode choice.
- Conventional household level travel surveys (e.g., National Household Level Travel Surveys-NHTS 2017, Transportation Tomorrow Survey 2016) do not have a separate survey component containing TNC related questions (Habib 2019; Grahn et al., 2020; Grahn et al., 2019). The only new addition in NHTS 2017 was the frequency of TNC trips in the last 30 days. Authors from the previous studies extracted those individuals' records where individuals reported TNC as a mode choice. In addition, NHTS 2017 (data collection period: March 2016 to May 2017) is a much older dataset to capture today's TNC demand. A similar comment applies to a few other studies which used the California Millennials Dataset (data collected in Fall 2015) (Alemi et al., 2019, 2018a, 2018b).
- Another set of studies relied on the stated-preference survey for automated vehicles, which did not have any revealed preference part for TNC (Gao et al., 2019).
- A few studies in New York indeed have historical data, which is very old compared to 2021. (six months in 2014 and six months in 2015) (Gerte et al., 2018). Also, the TNC demand in 2014 and TNC demand in 2021 are a little different.

How this study complements the previous studies:

- It is clear from the above discussion that none of the previous studies had access to a disaggregate trip-level dataset that we used in this study. Also, this study uses all Uber trips exclusively in Toronto between the period of September 2016 to September 2018. This study is different than many other studies which used household-level travel survey where TNC trips are often under-represented.
- Therefore, our work contributes to the TNC trip data literature. This stream of research has advantages over survey-based analysis because it uses revealed preference data rather than relying on retrospective reporting or the stated preferences of respondents. Many previous studies analyzed taxi trips and made policy recommendations for TNC services based on those analyses. We extend these previous works by estimating spatially detailed aggregate demand models of trip generation for each of 2017 and 2018 and aggregate attraction models based on actual TNC trips in a megacity (generally, less well represented in the literature).

- Another contribution of this study is using a multi-year dataset that allows us to better capture the seasonality and the evolution of TNC demand relative to past studies, which have relied on only a few months of data (Shokoohyar et al., 2020; Sun and Ding, 2019).
- Further, we estimated the time-series models of TNC demand. These models show the emerging pattern of TNC adoption as a travel mode that should be considered in transportation demand analysis and modelling.
- Due to the lack of comprehensive data sources on the Uber and Lyft ridership, the policymakers often struggle to make evidence-based policy recommendations to regulate such emerging modes. The series of models presented in this study will help us better understand the potential users of transportation network companies (TNC) and the effects of sociodemographic, weather, and built environment on transportation network company trips.

3. DATA DESCRIPTION

For this study, multiyear individual-level trip records are obtained from Uber through a partnership with the City of Toronto. This study uses the trips between the period of September 2016 to September 2018. In this study, records include the trip origin and destination locations given as the nearest intersection. Detailed start and end times are provided for records before April 2017, while start times for trips occurring between April 2017 and September 2018 are aggregated to the nearest hour. The detailed spatial resolution of the data means it is possible to perform more detailed analysis than is possible with most other TNC datasets. The individual-level data provided by Uber lack sociodemographic information, which makes it challenging to estimate conventional disaggregate demand models. Thus, data is analyzed at a dissemination area (DA) level, and all trips are aggregated based on the trip generation and trip attraction. According to Statistics Canada, dissemination area (DA) is the smallest geographical unit for which all census data are distributed (Statistics Canada. 2016). Most of the cases DA is smaller than Traffic Analysis Zone (TAZ).

This study focuses on 2017 and 2018 Uber trips to provide comparisons between years since there are only four months of data available for 2016. The full dataset (2016-2018) is used in section 5.5 for the autoregressive moving average (ARMA) model. We used only a portion of the data for the monthly models because we needed consistent coverage for all years of these models. We only had data for January-September for the 2018 data. That’s why in terms of trip generation and attraction model, Uber demand for September 2018 is used as the dependent variable. This is the final month (September 2018) for which data are available for 2018. To compare the two models, we used the Uber demand for September 2017 as a dependent variable for the trip generation and attraction model in 2017. In sections 5.1-5.4, models are estimated for total trip generation and attraction in September of each year, with data from earlier in the year included as lagged variables.

Figure 1 reveals the trip generation propensities of Uber trips in the year 2017 and 2018. The number of TNC trips in January 2017 was roughly 2 million, rising to 3 million by the end of that year (a 50% increase). Demand grew much faster in 2018, with the same percentage increase in demand occurring by the end of the first nine months. The aggregated datasets are fused with land use, built environment and sociodemographic data by aggregating trips to DA. Sociodemographic attributes are then incorporated at the DA-level based on 2016 census data provided by Statistics Canada (Statistics Canada. 2016). Land use attributes are generated from a combination of census data (i.e., total population and DA land area), and firmographic data derived from enhanced points of interests (EPOI) files provided by DMTI Spatial Inc. (2019). The EPOI data lists all establishments in the study region, including their coordinates and industry designation by North American Industry Classification System (NAICS) code. These coordinates are used to characterize the types of establishments located in each DA. These statistics are generated to use as explanatory variables in the trip attraction models. A summary of the data fusion method is illustrated in **Figure 2**, including the model variables derived from each dataset.

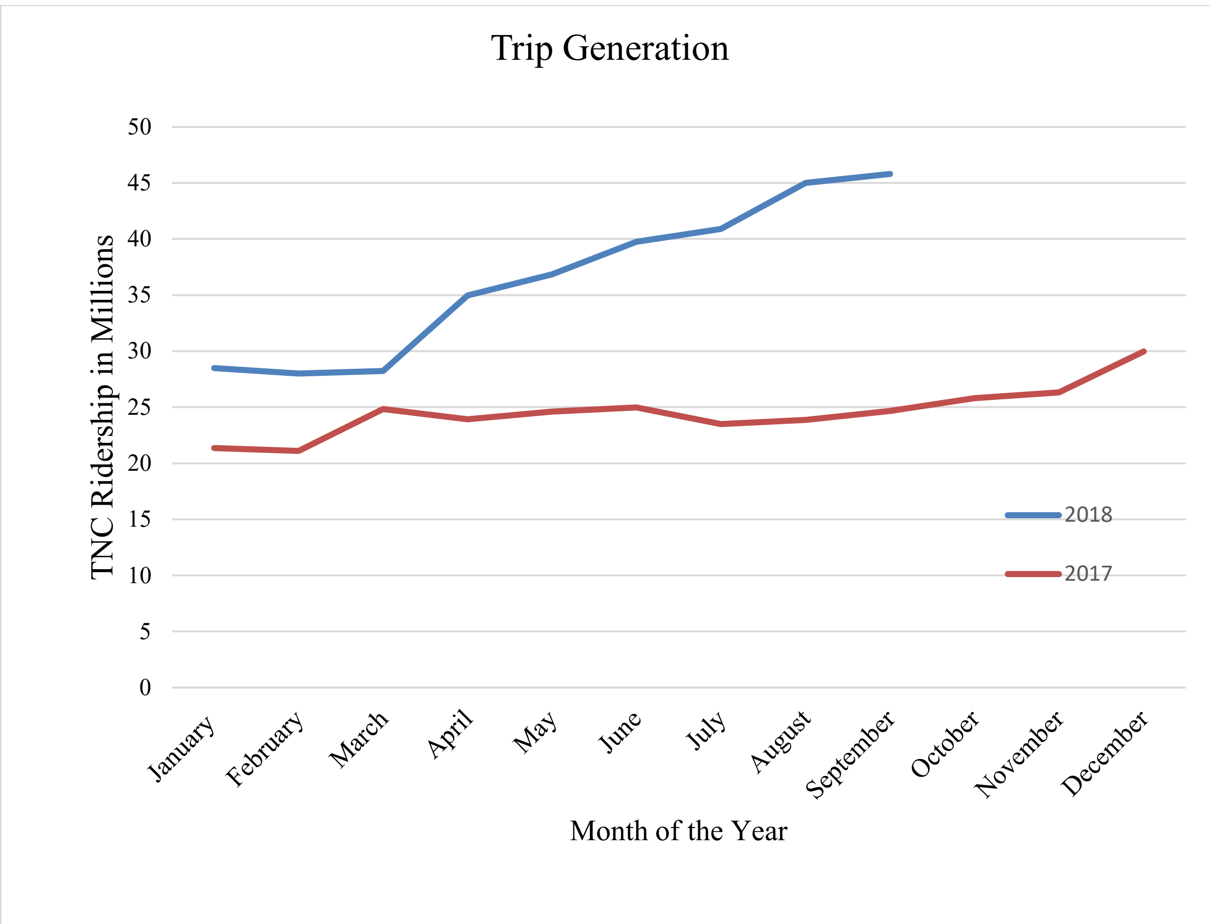

Figure 1 Trip Generation in 2018 and 2017

Table 1 shows the summary statistics of the key variables that are used in the empirical modelling. The City of Toronto is highly urbanized, as evident by more households living in apartments in tall buildings (greater than five storeys) than single-family detached dwellings in the average DA. However, despite the high density and accessibility of Toronto, many workers have commutes that exceed an hour. There is a high diversity in variable values across DA. TNC demand is highly variable between DA, with the standard deviation being a multiple of the average across Toronto. Comparing generation and attraction, average TNC attractions in 2018 are noticeably lower than their respective generation totals. Past research suggests that TNCs are often used in a single direction of travel tour (Big Data Innovation Team 2019). The difference between generation and attraction represents a stronger concentration of trip generation in specific DA than attraction. For the time-series model, we also include variables derived from other sources that vary daily. These data include Toronto Transit Commission (TTC) trips (all modes), the mean temperature each day in the study period, and bikeshare trips with origins in downtown Toronto (Historical weather data, 2019). Monthly Uber trip data is aggregated in dissemination area-level. The mean value in Table 1 represents the average number of trips generated from all the DAs in a given month.

The use of the terms "built environment" and "land use" in reference to features of a city that affect transportation demand is not consistent in the literature. In this paper, we rely on a combination of the definition of built environment by Sallis et al. (2012) and land use by Ewing and Cervero (2001). Land use is used to describe variables that result directly from zoning and land use regulations. These variables represent the local composition of residential and non-residential land use. Building environment variables are defined more broadly as infrastructure and urban form factors affecting the travel patterns of individuals. These factors manifest as travel times and mode choice variables that are directly influenced by the built environment - e.g., walking mode share is influenced by network connectivity, accessibility due to land use mix, and the quality and safety of pedestrian infrastructure (Sallis et al., 2012).

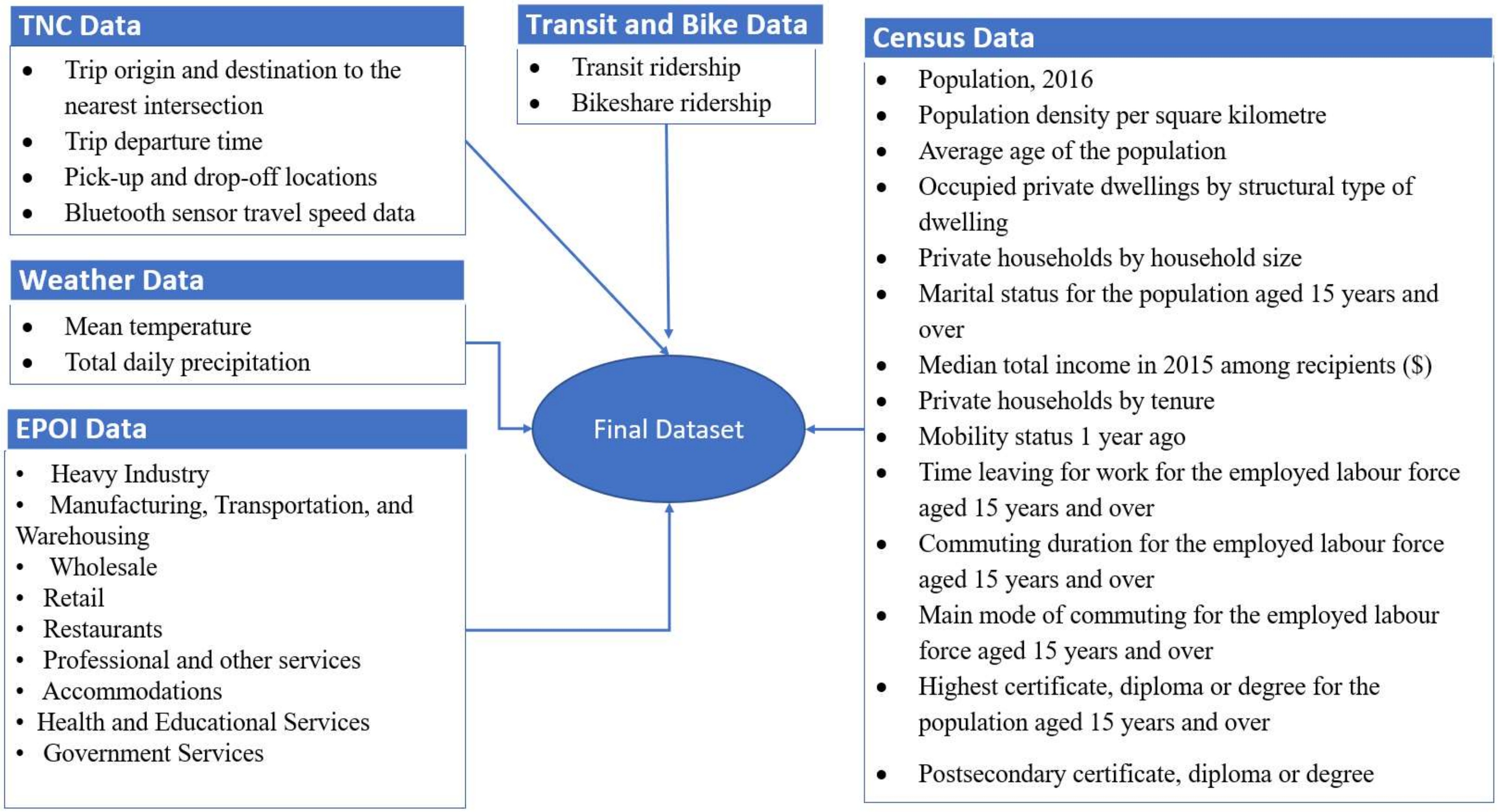

Figure 2 Data Fusion Method

TABLE 1 Descriptive Statistics of Key Variables (N = 3634 DA)
*** Count of households in DA**
**** Count of establishments in DA**

Trip Generation	Variable type (Land Use/Built Environment/ other)	Mean		Standard Deviation	
		Census and EPOI data 2016			
The average age of the population (count)	Demographic	41.29		5.50	
Married or living common-law couple*	Demographic	322.27		270.90	
Two-person household*	Demographic	89.96		105.36	
Private households by tenure: Owner*	Demographic	161.77		157.42	
Apartment in a building with ≥5 storeys*	Land Use	128.83		329.90	
Single-detached house*	Land Use	75.97		75.97	
Mobility status: (non)-movers*	Demographic	620.35		515.38	
Average commuting duration 15-29 minutes*	Built Environment	96.56		107.43	
Average commuting duration 45 to 59 minutes*	Built Environment	49.13		51.88	
Average commuting duration > 60 minutes*	Built Environment	54.92		61.91	
Leave for work between 12 p.m. and 4:59 a.m.*	Demographic	51.44		57.08	
Leave for work between 5 a.m. and 5:59 a.m.*	Demographic	15.14		18.35	
		Uber Data 2017		**Uber Data 2018**	
Total Generation in September		634,007		1,168,327	
		Mean	Standard Deviation	Mean	Standard Deviation
Lagged demand spring (number of trips)	---	656.72	2079.94	1968.13	6016.95
Lagged demand early summer (number of trips)	---	675.71	2126.23	1009.68	3060.30
Lagged demand late summer (number of trips)	---	655.26	2157.27	1233.65	3801.54
Trip Attraction		Mean		Standard Deviation	
		Census and EPOI Data 2016			
Private households by tenure: Renter*	Demographic	138.51		229.35	
Apartment in a building with <5 storeys*	Land Use	44.77		65.66	
Apartment in a building with ≥5 storeys*	Land Use	75.99		329.90	
Single-detached house*	Land Use	75.99		75.17	
Semi-detached house*	Land Use	19.61		33.10	
Average commuting duration 15-29 minutes*	Built Environment	96.55		107.43	
Average commuting duration 45 to 59 minutes*	Built Environment	49.11		51.88	
Average commuting duration > 60 minutes*	Built Environment	54.89		61.91	
Leave for work between 12 p.m. and 4:59 a.m.*	Demographic	51.41		57.08	
Commuting mode: Walk*	Built Environment	28.97		92.62	
Retail Stores**	Land Use	5.90		33.32	
Government services**	Land Use	1.33		0.45	
		Uber Data 2017		**Uber Data 2018**	

		Mean	Standard Deviation	Mean	Standard Deviation
Lagged demand winter (number of trips)	---	574.73	2119.80	756.18	2704.63
Lagged demand spring (number of trips)	---	641.21	2311.47	744.94	3006.76
Lagged demand summer (number of trips)	---	667.98	2436.23	971.56	3372.89
		Other Data Sources			
		Mean		Standard Deviation	
Toronto Transit Commission (TTC) average weekday ridership	Travel Variables	1685706.90		68037.39	
Mean temperature (°C)	Weather Variable	8.69		9.94	
Bikeshare trips (number of trips)	Travel Variables	2034.95		2175.15	

Figure 3 to **Figure 6** show the TNC demand variation in 2017 and 2018. TNC trips are concentrated in the central business district (lower center), Toronto Pearson International Airport (upper left), and Downsview Aiport/York University (upper center). The same spatial pattern is observed for both generation and attraction. Comparing plots for 2017 and 2018, the spatial distribution of trips is quite similar, but totals increase across the city. The lowest trip density is found in Scarborough at the eastern extreme of the city.

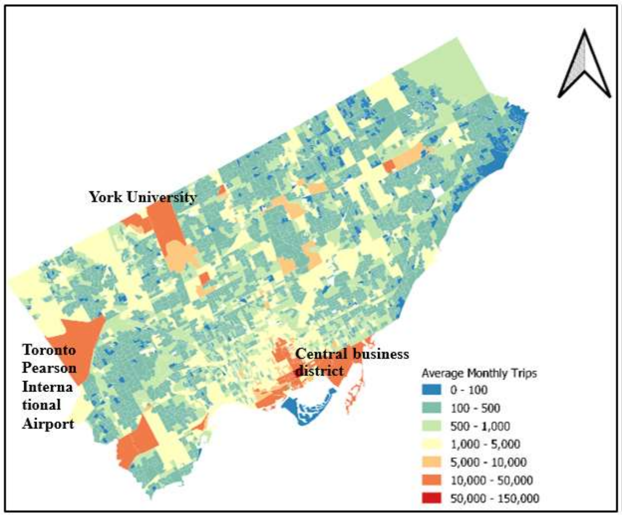

Figure 3 Trip Generation in 2018

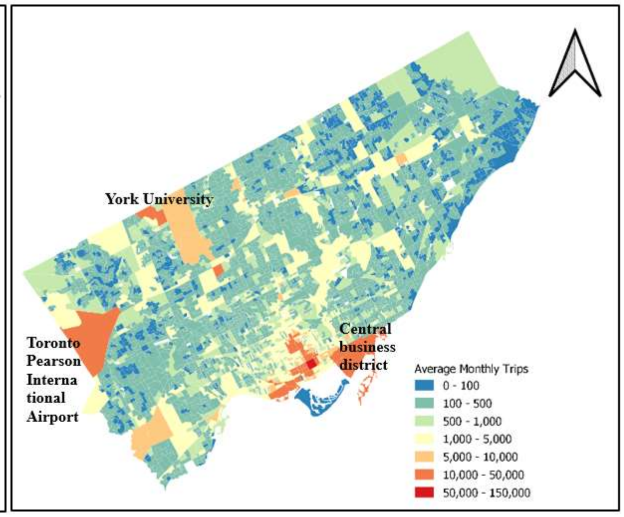

Figure 4 Trip Generation in 2017

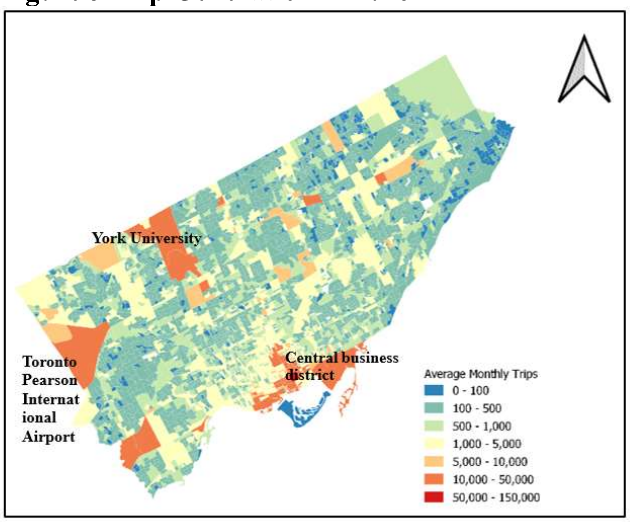

Figure 5 Trip Attraction in 2018

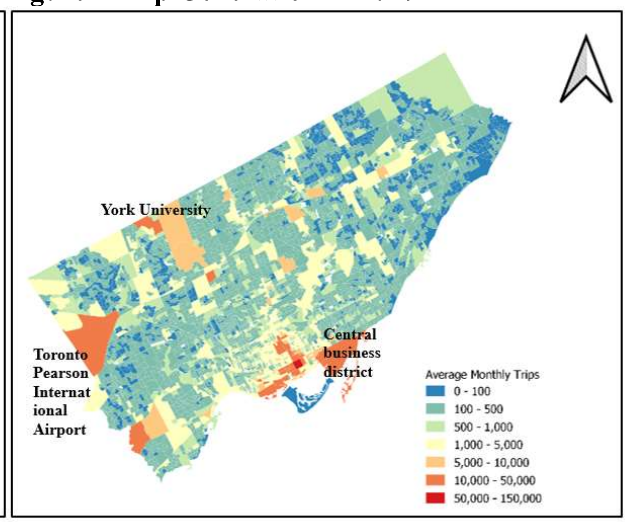

Figure 6 Trip Attraction in 2017

4. MODEL FORMULATIONS

4.1 Constant Elasticity of Demand Model

The multiplicative constant elasticity of demand model takes the following form (Greene 2005, Nerlove 1963):

$$D = {X_1}^{\beta_1}{X_2}^{\beta_2} \dots {X_n}^{\beta_n} \exp(\alpha_0 + \alpha_1 Z_1 + \alpha_2 Z_2 + \cdots \alpha_m Z_m + \varepsilon) \quad (1)$$

$$D = \prod_{i=1}^{n} {X_i}^{\beta_i} \prod_{j=1}^{m} (\beta_0 + \beta_j Z_j) \quad (2)$$

where
D = Total demand
α_0 = Constant
X_i = Continuous variable
β_i = Estimated parameter for continuous variable $\forall i \in n$

Z_j= Dummy or categorical variable (log-transformation is not required) $\forall j \in m$
α_j=Estimated parameter for dummy or categorical variable
ε = random error

Taking log-transformations of the continuous variables in equation (2), the demand function takes the the following form:

$$\ln(D) = \alpha_0 + \sum_{i=1}^{n} \beta_i \ln(X_i) + \sum_{j=1}^{m} \alpha_j Z_j + \varepsilon \quad (3)$$

This functional form was chosen for the following reasons:

1. A log-transformed constant elasticity formulation will ensure that the predicted demand value is non-negative. This property is critical in this study, since we are modelling Uber trip demand.
2. The parameter values are equivalent to the elasticity for continuous variables (as shown in equation (4)).
3. This study assumed a multiplicative error-term, since a few studies found that the multiplicative error model performs better than the additive error model (Iyaniwura et al., 2019).

For the continuous variables in Equation (3), taking the partial derivative, the elasticity of the corresponding variable can be computed as:

$$E_i = \frac{\partial ln(D)}{\partial X_i}\frac{X_i}{D} = \frac{\beta_i D}{X_i}\frac{X_i}{D} = \beta_i \quad (4)$$

4.2 Distributed Lag Model

After carefully analyzing the historical data, it is assumed that the seasonal factors of past ridership may impact the future ridership of Uber. Therefore, distributed lag models are estimated for both trip generation and attraction. A distributed lag model accounts for the effect of repeated use of TNCs. Besides, a distributed lag model captures the seasonal effect on TNC demand. If D_{t-1} is a lagged dependent variable and γ is the corresponding parameter, using a constant elasticity demand function, the demand at time t can be computed using the following equation:

$$\ln(D_t) = \alpha_0 + \gamma \ln(D_{t-1}) + \sum_{i=1}^{n} \beta_i \ln(X_i) + \sum_{j=1}^{m} \beta_j Z_j + \varepsilon \quad (5)$$

According to Owen and Philips (1987), the relationships between demand at time t (D_t) and the lagged demand (D_{t-1}) can be written:

$$\ln(D_t) - \ln(D_{t-1}) = \eta[\ln(D_t) - \ln(D_{t-1})] \quad (6)$$

In equation (6) η is the speed of adjustment, which should be in between zero to one (Brainard & Tobin 1968). The equality is true if η is one, meaning the adjustment is immediate. If η is close to zero, then no adjustment is observed (Brainard & Tobin 1968; Owen and Philips 1987).

4.3 Time-series Model
For the time-series model, we adopted the autoregressive moving average (ARMA) *(j,d,k)* model, which harnesses the power of autoregressive (lagged dependent variable) terms *(j)* and moving average terms *(k)*. '*d'* means the number of times the data are differenced. The general form of the ARMA model can be written as follows:

$$D = \alpha_0 + \sum_{i=1}^{n} \beta_i \mathrm{X_i} + \sum_{i=1}^{j} \Phi_i D_{t-i} + \sum_{i=1}^{k} \theta_i \varepsilon_{t-i} + \varepsilon \qquad (7)$$

In equation (7)
D = total demand
D_{t-i} = lagged demand
α_0 = constant
X_i = Explanatory variable
β = Estimated parameter associated with explanatory variable
j = The order of lagged demand
k = The order of moving average
Φ=Parameter for lagged demand
ε_{t-i}=Parameter for moving average component

Random-walk and non-stationary issues of time-series data are well studied in the past literature (Chen, Varley & Chen, 2011). Random walk means that the future value of the dependent variable is a random step away from the current value. After a standard statistical test for the random walk, it is found that the time series data used for this study is not a random walk (Doornik & Hendry 2006). Besides, the Dickey-Fuller test is conducted to determine whether the dataset is non-stationary (Greene 2005). It is found that the time-series data used in this study is stationary. Therefore, no data transformation is required in this case, such as differencing and detrending. Differencing and detrending are commonly applied methods to ensure stationarity of a time-series dataset. Detrending is needed if the time-series dataset exhibits any specific trend. After estimating a regression model, we need to estimate the residuals. Eventually, we model the residuals if it shows a stationary pattern. When a variable is not stationary, another common method is using differenced variables (also known as "differencing"). A differenced variable can be generated by taking the difference between the original demand and the lagged demand.

5. EMPIRICAL MODEL

5.1 Monthly DA-level Trip Generation Model 2017 and 2018
Monthly total Uber trips are aggregated at the DA-level for the year 2018. A total of 3,634 DAs are used for empirical modelling. Trip generation in September is used as the dependent variable because this is the final month for which data are available for 2018. A wide range of sociodemographic, land use and built environment data are used as explanatory variables. The previous months of data available for 2017 and 2018 are used to form seasonal lag variables to account for seasonal variation in Uber demand and unobserved spatial variation between DA. **Table 2** shows the parameter estimation results of the trip generation model for September 2018 using equation 5. It is found that the R-squared value is 0.985, which means 98.5% of the variation in trip generation can be explained by the explanatory variables included in this model.

Thus, the model's goodness-of-fit is relatively high.

All parameters are statistically significant at a 95% confidence interval. The model results show that the late-summer lagged demand (August) is positively correlated with the trip generation in September. The early-summer (May) lagged demand is also positively correlated with the trip generation in September. However, the coefficient of the lag-demand in late-summer (0.7) is higher than in the early summer. This finding is intuitive since naturally a user's mode choice habit from the past month is more likely to affect the user's current month's mode choice than is their mode choice habit from three to four months earlier. In terms of the elasticity, lagged demand of late-summer indicates that a 1% increase in the demand in the previous month can increase 0.7% more TNC ridership in the current month.

The average age of the population is positively correlated with TNC demand, indicating that, with an increasing number of older individuals in a DA, an increasing number of the TNC trips will be generated in that DA. Types of dwelling unit are added as explanatory variables in the trip generation model. It is found that if a specific DA has a higher number of single-detached houses, a reduced number of TNC trips will be generated from that DA. As expected, the coefficient for the 'apartment in a building with more than five storeys' is positive, indicating that, the higher number of apartments in the origin DA, the more TNC trips will be generated. It is found that the number of two-person households is negatively correlated with the number of generated trips. Many married or common-in-law couples fall within the category of the two-person household. This group of people may have a car in the household, or they may use public transit for typical commuting, which reduces their reliance on TNCs.

Trip departure time and commuting duration are two critical factors for TNC demand generation. The model results reveal that DA with high numbers of individuals who leave work between midnight and 4:59 am tend to generate more Uber trips. In Toronto, the subway closes at 1:30 am, which forces workers to choose TNC late at night. The results demonstrate a negative correlation between average commuting duration of 15-29 minutes in DA-level and TNC trip generation. In contrast, the results demonstrate a positive correlation between average commuting duration greater than 60 minutes in DA-level and TNC trip generation. This finding suggests that Uber is more prevalent in those DAs where average commuting duration is more than 60 minutes. In other words, individuals are less likely to use Uber for short-haul trips.

TABLE 2 Trip Generation Model Results

Trip Generation Model Results for 2018 Demand			
Number of Observations		3634	
R-squared		0.985	
Variable		Parameter	t-stat
Constant		0.137	5.540
Seasonality	Lagged demand late summer	0.700	58.370
	Lagged demand early summer	0.295	24.640
The average age of the population		0.017	2.620
Single-detached house		-0.004	-2.450
Apartment in a building with ≥5 storeys		0.004	2.780
Two-person household		-0.017	-2.490
Leave for work between 12 p.m. and 4:59 a.m.		0.010	2.920
Average commuting duration 15-29 minutes		-0.022	-4.270
Average commuting duration > 60 minutes		0.013	4.120
Trip Generation Model Results for 2017 Demand			
Number of Observations		3642	
R-squared		0.979	
Variable		Parameter	t-stat
Constant		0.034	0.970
Seasonality	Lagged demand late summer	0.600	49.180
	Lagged demand early summer	0.317	22.220
	Lagged demand spring	0.081	7.960
The average age of the population		0.027	1.82
Single-detached house		-0.004	-2.04
Mobility status: non-movers		-0.010	-1.380
Average commuting duration > 60 minutes		0.016	3.770
Leave for work between 5 a.m. and 5:59 a.m.		-0.004	-1.660

In terms of the trip generation model for 2017, a total of 3,642 DAs are retained for the empirical modelling after aggregating the monthly Uber trips to a DA-level. To match with the 2018 trip generation model, trip generation in September is used as the dependent variable. **Table 2** depicts the parameter estimation results of the trip generation model for September 2017. The model's goodness-of-fit is relatively high with an R-squared value of 0.979. This also indicates that 97.9% of the variation in trip generation can be explained by the explanatory variables included in this model.

The model result shows that all lag-demand parameters are positively correlated with the trip generation in September. The coefficient of the lagged demand of late summer (August) is higher than the coefficient of the lagged demand of early summer (May). In terms of the elasticity, lagged demand of late summer indicates that a 1% increase in the demand in the previous month can increase 0.6% more TNC ridership in the current month. Despite the statistical significance, lagged demand in spring (April) has less influence on the TNC trip generation than the summer season.

The average age of the population is a significant predictor of Uber trip generation. The positive correlation of the age variable indicates that if a specific area has a higher number of older individuals, there is more likely to be an increase in Uber trip generation. Single-detached house is negatively correlated with TNC trip generation. Most of the single-detached houses in Toronto are situated in suburban areas where many households own at least one car. Therefore, Uber demand is significantly lower in these suburbs. The model results demonstrate a positive correlation between commuting trips, where trip duration is more than 60 minutes. This finding echoes the earlier trip generation model for 2018. This finding essentially suggests that individuals tend to use Uber for long duration trips than short duration trips. The model also reveals that individuals who leave for work between 5 am and 5:59 am negatively affect the TNC trip generation. This finding is intuitive since most of the commuters leave home for work in the early morning and they are more likely to use a regular mode (e.g., automobile, transit, bike) than Uber.

5.2 Monthly DA-level Trip Attraction Model 2017 and 2018

Uber trips are aggregated for the destination DA for the year of 2018. A total of 3,649 DAs trip records are used as the dependent variable. Similar to the trip generation model, a wide range of sociodemographic, land use, and built environment data are fused with the aggregate TNC demand. Besides, seasonal lagged demand is incorporated into the model. **Table 3** shows the parameter estimation results of the trip attraction model for September 2018 using equation 5. The R-squared value is 0.972, which indicates a reasonably good fit. This also indicates that 97.2% of the variation in trip attraction can be explained by the explanatory variables included in this model. All parameters used in this model are statistically significant at a 95% confidence interval.

The model result shows that the influence of lagged ridership and seasonal TNC ridership are statistically significant. The model result shows that the summer (May) lagged demand is positively correlated with the trip attraction in September. In terms of elasticity, this finding indicates that a 1% increase in summer demand can increase Uber trip attraction by 0.66%. The winter (January) lagged demand is also positively correlated with the trip attraction in September. However, the effect of summer demand is slightly higher than the effect of winter demand on Uber trip attraction. The negative sign for single-detached house indicates that if certain areas have government services, it is less likely to generate Uber trips destined to those areas. Single-detached house is a significant predictor for Uber trip attraction. The model results indicate that Uber trip attraction decreases with an increasing number of a single-detached house. The model also indicates that apartment buildings with less than five storeys are positively correlated with Uber trip attraction. These results are expected because most individuals who live in downtown tend to live in apartments. Therefore, the Uber trip attraction is higher in such areas than the suburbs where most individuals live in detached houses or semi-detached houses. If the household tenure category is renting, this attracts more Uber trips. These findings are intuitive since individuals who rent a house/apartment are less likely to own an automobile than individuals who permanently live in a house/apartment that they own.

TABLE 3 Trip Attraction Model Results

Trip Attraction Model Results for 2018 Demand			
Number of Observations		3649	
R-squared		0.972	
Variable		Parameter	t-stat
Constant		0.485	17.600
Seasonality	Lagged demand summer	0.676	45.840
	Lagged demand winter	0.287	20.090
Government Services		-0.020	-1.680
Single-detached house		-0.003	-1.880
Apartment in a building with <5 storeys		0.003	1.780
Private households by tenure: renter		0.021	6.540
Average commuting duration 15-29 minutes		-0.031	-5.220
Average commuting duration > 60 minutes		0.030	7.850
Trip Attraction Model Results for 2017 Demand			
Number of Observations		3648	
R-squared		0.949	
Variable		Parameter	t-stat
Constant		0.012	0.340
Seasonality	Lagged demand summer	0.533	43.870
	Lagged demand spring	0.277	11.180
	Lagged demand winter	0.162	6.910
Semi-detached house		-0.0003	-2.070
Leave for work between 12 p.m. and 4:59 a.m.		0.007	1.290
Average commuting duration 15-29 minutes		0.028	3.160
Retail Stores		0.017	3.480
Commuting mode: Walk		-0.004	-0.990

For the year of 2017, all Uber trips are aggregated based on the destination DA. A total of 3,648 DA are used for empirical modelling. **Table 3** shows the estimated parameters of the trip attraction model for September 2017. The R-squared value is 0.949, which shows that 94.9% of the variation in trip attraction can be explained by the explanatory variables included in this model. In terms of seasonality, summer, spring, and winter lagged demand are included as explanatory variables. Overall, the model displays a positive correlation with all lagged demand variables.

The summer (May) lagged demand is more influential than the spring (March) and winter (January). The negative sign of the parameter for walk mode indicates that the number of commuters who walk to the destination increases, Uber trip attraction will be decreased. The trip attraction model for the year 2017 indicates that TNC demand increases if there are more retail stores in a dissemination area. The model results indicate that the dissemination areas where workers leave for work between midnight and 4:59 am tend to attract more TNC trips. This finding is expected since there are no subway services in the City of Toronto after 1:30 am. Even though bus services are operated along the subway corridor in late-night, many individuals choose Uber as a convenient alternative in such cases. Similar to the other models discussed above, it is found that an increase of semi-detached houses in a DA decreases the TNC demand.

5.3 Autoregressive Moving Average Model for Downtown Toronto

Table 4 shows the parameter estimation results of the ARMA model. Even though conventional trip generation model can account for the seasonal effect, its forecasting performance is often poor. In this study, the 24-months Uber ridership data (September 2016 to September 2018) is aggregated for the downtown core, which is also known as planning district-1. A wide range of time-varying attributes such as mean temperature, total daily precipitations, bike-sharing demand, and transit trips are included in the model.

The model results indicate that TNC demand increases with increasing temperature. It is also found that the total daily precipitation is a significant predictor of Uber ridership. The total daily precipitation has a positive sign indicating that with increasing rain, individuals are more like to use Uber. Aggregate bike-sharing trip count has a negative correlation with Uber trip attraction. A substantial portion of downtown Toronto roads have separated bike lanes, and during rush hour, bike share can be a faster and more affordable mode for commuters. Therefore, the negative correlation is expected between the bike share demand and the Uber demand. The model results also indicate that average weekday transit trips are negatively correlated with Uber trip attraction. In a high-frequency transit corridor, individuals rely more on transit. Therefore, this finding is intuitive.

TABLE 4 ARMA (1,0,1) Model Results

Number of Observations	754	
Sum of squares	335.14	
Variable	Parameter	t-stat
Constant	4.103	13.44
Mean temperature (°C)	0.030	16.9
Total daily precipitations (mm)	0.013	3.83
Bike share trips	-0.169	-18.35
Average weekday transit trips	-0.087	-5.64
TNC ridership at t-1 (AR-1)	0.272	8.8
Moving average component	0.607	10.86

6. CONCLUSIONS

This paper put forward several empirical findings around the theme of the temporal and spatial demand generation of transportation network companies. This study made use of a multiyear dataset, which consists of Uber trips from September 2016 to September 2018. Using multiplicative constant elasticity demand functions, a series of aggregate demand models are estimated to investigate the determinants of the TNC demand in the City of Toronto, a study that has not been done before in this region. Besides, to capture the effect of weather, bike-sharing, and transit trips on Uber demand generation, a time-series model-ARMA is also estimated. This analysis has added important empirical implications to the under-discussed issue of aggregate demand modelling of TNC trips, and this finding can be transferrable to other megacities in North America.

It is found that the goodness-of-fit of all trip generation and attraction models are over 0.95, which represents a very good fit. All models show that lag-demand is a significant predictor of Uber demand. The model result shows that summer lagged demand (May) is positively correlated with trip generation in September. The average age of the population is positively correlated with TNC demand, indicating that, with an increasing number of older individuals in a DA, an increasing number of TNC trips will be generated in that DA. It is found that the number of single-detached houses is negatively correlated with Uber demand, whereas the number of 'apartment in a building with more than five storeys' is positively correlated. The household tenure category is a significant predictor of Uber trip attraction. Households that rent their dwelling are more likely to make Uber trips. These findings are intuitive since individuals who rent a house/apartment are less likely to own an automobile than individuals who permanently live in a house/apartment which they own.

The time-series model reveals that total daily precipitation has a positive sign indicating that with increasing rain, individuals are more likely to use Uber. The model results indicate that TNC demand increases with increasing temperature. As expected, aggregate bike-sharing trip count has a negative correlation with Uber trip attraction. The model results demonstrate that average weekday transit trips are negatively correlated with Uber trip attraction. In a high-frequency transit corridor, individuals rely less on Uber.

One critical assumption made in this study is that statistics obtained from the Census profile are relatively stable over the study period of 2016 to 2018. It may be useful to examine how these statistics have changed over the study period and include this variation in the models. The finding of significant parameters for overnight commuting trips and long commutes (> 60 minutes) suggests that Uber is filling gaps in transit. We perceive two possible policy directions for the City of Toronto based on these findings and trends in other cities. The city could accept this complementary relationship and incorporate patterns of Uber use into their planning of bus routing and scheduling. Alternatively, the city could consider the adoption of other on-demand transportation options, perhaps with a larger passenger capacity to address environmental concerns associated with operating many small vehicles.

One possible extension of this study will be panel data regression analysis. Since we have daily Uber demand for each dissemination area (DA), a panel data regression analysis could provide more insight. Also, a more comprehensive travel survey should be conducted among Uber and

Lyft users and drivers to understand their sociodemographic profile. Due to privacy reasons, the dataset presented in this study did not include any personal (i.e., age, gender, income) and household level (i.e., number of vehicles, total household income) attributes.

ACKNOWLEDGMENT
The study was partially funded by an NSERC Discovery grant and Percy Edward Hart professorship awarded to the third author. The study was possible because ride-hailing companies shared trip records that were made available for this research through a research contract with the City of Toronto. However, authors are solely responsible for all errors and interpretations.

AUTHOR CONTRIBUTIONS
Dr. Md Sami Hasnine estimated the empirical models and designed the study. Jason Hawkins prepared the dataset for empirical modelling and performed the literature review. Dr. Md Sami Hasnine, Jason Hawkins and Prof. Khandker Nurul Habib contributed to the final version of the manuscript. Prof. Khandker Nurul Habib supervised the project.

REFERENCES

1. Alemi, F., Circella, G., Handy, S., and Mokhtarian, P., 2018a. What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California. Travel Behaviour and Society, 13, pp.88-104.
2. Alemi, F., Circella, G., Mokhtarian, P., Handy, S., 2018b. Exploring the latent constructs behind the use of ridehailing in California. Journal of Choice Modelling 29, 47–62. https://doi.org/10.1016/j.jocm.2018.08.003
3. Alemi, F., Circella, G., Mokhtarian, P., and Handy, S., 2019. What drives the use of ride-hailing in California? Ordered probit models of the usage frequency of Uber and Lyft. Transportation Research Part C: Emerging Technologies, 102, pp.233-248.
4. Big Data Innovation Team. 2019. The Transportation Impacts of Vehicle-for-Hire in the City of Toronto.
5. Brainard, W.C., and Tobin, J., 1968. Econometric Models: Their Problems and Usefulness. Pitfalls in Financial Model Building. American Economic Review, 58(2).
6. Chen, F., Yin, Z., Ye, Y., & Sun, D. J. (2020). Taxi hailing choice behavior and economic benefit analysis of emission reduction based on multi-mode travel big data. *Transport Policy*, 97, 73-84.
7. Chen, C., Varley, D., and Chen, J., 2011. What affects transit ridership? A dynamic analysis involving multiple factors, lags, and asymmetric behaviour. *Urban Studies*, *48*(9), pp.1893-1908.
8. Conway, M., Salon, D., King, D., 2018. Trends in Taxi Use and the Advent of Ridehailing, 1995–2017: Evidence from the US National Household Travel Survey. Urban Science 2, 79. https://doi.org/10.3390/urbansci2030079
9. Doornik, J. A. and Hendry, D. F. (2006) Econometric Modelling: PcGive 12, 2nd edn. London: Timberlake Consulting Ltd
10. DMTI Spatial Inc., 2019. CanMap GIS Data for GIS Mapping Software - DMTI Spatial. [Online]. Available: https://www.dmtispatial.com/canmap/ . [Accessed: 26-Jul-2019].

11. Dias, F.F., Lavieri, P.S., Garikapati, V.M., Astroza, S., Pendyala, R.M., Bhat, C.R., 2017. A behavioral choice model of the use of car-sharing and ride-sourcing services. Transportation 44, 1307–1323.
12. Dias, F.F., Lavieri, P.S., Kim, T., Bhat, C.R. and Pendyala, R.M., 2018. Fusing Multiple Sources of Data to Understand Ride-Hailing Use. Transportation Research Record
13. Dias, F.F., Kim, T., Bhat, C.R., Pendyala, R.M., Lam, W., Pinjari, A.R., Srinivasan, K., Ramadurai, G., 2019. Modeling the Evolution of Ride-Hailing Adoption and Usage: A Case Study of the Puget Sound Region. https://doi.org/10.1017/CBO9781107415324.004
14. Ewing, R., & Cervero, R. (2001). Travel and the built environment: a synthesis. *Transportation research record*, *1780*(1), 87-114.
15. Gao, J., Ranjbari, A., MacKenzie, D., 2019. Would being driven by others affect the value of travel time? Ridehailing as an analogy for automated vehicles. Transportation 46, 2103–2116. https://doi.org/10.1007/s11116-019-10031-9
16. Gerte, R., Konduri, K.C., and Eluru, N., 2018. Is there a limit to adoption of dynamic ridesharing systems? Evidence from analysis of Uber demand data from New York City. Transportation Research Record, 2672(42), pp.127-136.
17. Gerte, R., Konduri, K.C., Ravishanker, N., Mondal, A., and Eluru, N., 2019. Understanding the Relationships between Demand for Shared Ride Modes: Case Study using Open Data from New York City. Transportation Research Record.
18. Grahn, R., Harper, C.D., Hendrickson, C., Qian, Z., Matthews, H.S., 2019. Socioeconomic and usage characteristics of transportation network company (TNC) riders. Transportation. https://doi.org/10.1007/s11116-019-09989-3
19. Grahn, R., Qian, S., Matthews, H.S., Hendrickson, C., 2020. Are travelers substituting between transportation network companies (TNC) and public buses? A case study in Pittsburgh. Transportation. https://doi.org/10.1007/s11116-020-10081-4
20. Greene, W., 2005. Econometric Analysis, 5th Edn, Pearson, India
21. Hall, J.D., Palsson, C., and Price, J., 2018. Is Uber a substitute or complement for public transit?. Journal of Urban Economics, 108, pp.36-50.
22. Habib, K.M.N., 2019. Mode Choice Modelling for Hailable Trips: An Investigation of the Competition of Uber with Other Modes by Using an Integrated Non-Compensatory Model with Probabilistic Choice Set Formation. Presented at the 98th Annual Meeting of Trans. Res. Board, January 13–17.
23. Historical weather data, 2019. Government of Canada. https://climate.weather.gc.ca/historical_data/search_historic_data_e.html
24. Iyaniwura, J. O., Adepoju, A. A., & Adesina, O. A. (2019). Parameter estimation of Cobb-Douglas production function with multiplicative and additive errors using the frequentist and Bayesian approaches. Annals. *Computer Science Series*, 17(1).
25. Lavieri, P.S., Dias, F.F., Juri, N.R., Kuhr, J. and Bhat, C.R., 2018. A model of ridesourcing demand generation and distribution. Transportation Research Record, 2672(46), pp.31-40.
26. Lavieri, P.S., Bhat, C.R., 2019a. Investigating objective and subjective factors influencing the adoption, frequency, and characteristics of ride-hailing trips. Transportation Research Part C: Emerging Technologies 105, 100–125. https://doi.org/10.1016/j.trc.2019.05.037

27. Lavieri, P.S., Bhat, C.R., 2019b. Modeling individuals' willingness to share trips with strangers in an autonomous vehicle future. Transportation Research Part A: Policy and Practice 124, 242–261. https://doi.org/10.1016/j.tra.2019.03.009
28. Laptev, N., Yosinski, J., Li, L.E., and Smyl, S., 2017, June. Time-series extreme event forecasting with neural networks at uber. In International Conference on Machine Learning (Vol. 34, pp. 1-5).
29. Louviere, J.J., Flynn, T.N., Carson, R.T., 2010. Discrete choice experiments are not conjoint analysis. Journal of Choice Modelling 3, 57–72. https://doi.org/10.1016/S1755-5345(13)70014-9
30. Nerlove, M. (1961). *Returns to scale in electricity supply*. Institute for mathematical studies in the social sciences.
31. Owen, A.D., and Phillips, G.D.A., 1987. The characteristics of railway passenger demand. An econometric investigation. Journal of Transport Economics and Policy, pp.231-253.
32. Ozonder, G., Miller, E.J., 2019. Vehicle for hire bylaw review report 3: Taxi time-series analysis.
33. Sallis, J. F., Floyd, M. F., Rodríguez, D. A., & Saelens, B. E. (2012). Role of built environments in physical activity, obesity, and cardiovascular disease. Circulation, 125(5), 729-737.
34. Shokoohyar, S., Sobhani, A., Ramezanpour Nargesi, S.R., 2020. On the determinants of Uber accessibility and its spatial distribution: Evidence from Uber in Philadelphia. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. https://doi.org/10.1002/widm.1362
35. Spurlock, C.A., Sears, J., Wong-Parodi, G., Walker, V., Jin, L., Taylor, M., Duvall, A., Gopal, A., and Todd, A., 2019. Describing the users: Understanding adoption of and interest in shared, electrified, and automated transportation in the San Francisco Bay Area. Transportation Research Part D: Transport and Environment, 71, pp.283-301.
36. Statistics Canada. 2016. Census Profile-2016 Census
37. Sun, D., Ding, X., 2019. Spatiotemporal evolution of ridesourcing markets under the new restriction policy: A case study in Shanghai. Transportation Research Part A: Policy and Practice 130, 227–239. https://doi.org/10.1016/j.tra.2019.09.052
38. Sun et al., 2018. Analyzing spatiotemporal traffic line source emissions based on massive didi online car-hailing service data, Transportation Research Part D.
39. Wang, Y., Wang, S., Wang, J., Wei, J., and Wang, C., 2018. An empirical study of consumers' intention to use ride-sharing services: Using an extended technology acceptance model. Transportation, pp.1-19.