THE IMPACT OF RIDE-HAILING SERVICES ON TRAVEL BEHAVIOUR

BY

ANUSHA MISHRA

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Urban Planning in Urban Planning
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Master's Committee:

Associate Professor Bumsoo Lee, Adviser
Assistant Professor Lindsay Maurer Braun

ABSTRACT

The introduction of Ride- hailing Services into our transport systems has been rapidly transforming the way people travel. Ride-hailing services provide multi-modality and fill transit gaps, but they also impact the modal share of other modes such as public transit and car ownership. This study delves into links between ride-hailing services and private vehicles ownership. It also addresses the questions about the impact of ridesharing services on public transit use and the role neighborhood context plays on the link between ride-hailing and car-ownership, using a Path Analysis. The primary database for the research is the NHTS 2017 survey. Data compilation is done to establish a dataset of cities with TNCs operating in them and the duration of operation.

Main findings of the study are as follows:

First, the relationship between public transit and ridesharing is found to be statistically insignificant while a descriptive analysis shows that ride sharing services complement public transit more especially in small towns. Second, ride-hailing has a significant and comparatively large impact on car ownership. Due to the bidirectional nature of the model, we were able to study the reverse relationship as well. The model did not show car-ownership having a significant impact on frequency of rideshare use. Finally, through a moderation estimation that urban form does play a significant role in impacting the role of rideshare on car ownership. The length of duration since the introduction of TNCs in a city plays an important role on car ownership. The longer TNCs have been around, the smaller the value of car ownership is. Denser Urban forms deepen this relationship while sprawled neighborhoods weaken the correlation.

Based on this research a few areas have been identified as areas with critical data deficiency which are needed to understand and properly manage the ever-changing travel behavior. These areas include the links between city types, public transit and rideshare.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Advances in technology are rapidly transforming the global transportation system. Technology assisted services like Intelligent Transportation Systems, Smart Cards, Wayfinding and Navigation Apps, and GPS can impact travel behavior in multiple ways, such as removing uncertainties, increasing mode choice or improving accessibility.

These advancements have given rise to shared mobility services such as ride-hailing. They are also known as Mobility Service Providers (MSP), Transportation Network Companies (TNCs), ride-hailing services or ride-sourcing. The evolving market for shared mobility services has expanded exponentially since the introduction of smart phones. Since 2009, the rise of peer-to-peer ridesharing services such as Lyft and Uber have redefined the transportation system, globally. Not only did they account to 10 % of all work-related trips made in 2016[1], they have transformed the structure of the US modal split to an extent that ride-hailing can be considered its own category. Although initially concentrated in dense markets such as cities and university campuses, these services have a growing presence in the U.S. market (Alemi et all, 2018). As per the 2017 NHTS statistics 9.81% of Americans use ride-hailing at least once a month (Conway et all, 2018). SFCTA 2017's report showed that in San Francisco, the total share of ride-hailing trips can exceed 15% of all trips on a typical weekday and account for 20% of all VMT within city limits[2]. By 2035, the number of ride-railing users in the United States is expected to amount to 82.7m[3].

---

[1] Lyft Revenue and Usage Statistics (2019). (2019, April 29). Retrieved from https://www.businessofapps.com/data/lyft-statistics/

[2] San Francisco County Transportation Authority (SFCTA). 2017. "TNCs Today: A Profile of San Francisco Transportation Network Company Activity". available at http://www.sfcta.org/sites/default/files/content/Planning/TNCs/TNCs_Today_112917.pdf (last accessed on January 22, 2018).

[3] Ride Hailing - United States: Statista Market Forecast. (n.d.). Retrieved from https://www.statista.com/outlook/368/109/ride hailing/united-states#market-users

With the rapid rise of ride-hailing, several challenges have arisen. The primary challenge is the lack of data. Presently, no policy exists which mandates TNCs to share their data with local or regional planning authorities. As most TNCs refuse to volunteer information, data availability is reduced to either primary collection or state/federal surveys. The lack of data cripples a planner's ability to make informed decisions such as those regarding public transit investments or transportation infrastructure (Clewlow et al.,2017). This research aims to reduce some of the data shortage present today.

Another key challenge that has emerged with ride-hailing is the increase in multi-modality. Although ride-hail fills mobility gaps, there has been research (Clewlow et al.,2017) that shows ride-hailing replaces transit usage, active transportation trips and, to a lesser extent, personal vehicle trips. This changing modal split needs to be addressed while making transportation decisions. Improper transit or infrastructure investments can be a wasteful use of tax dollars.

Ride-hailing is sometimes referred to as disruptive transportation. The business model of TNCs has 'disrupted' the existing Taxicab Industry. Drivers have migrated to TNCs as their competitively lower process attract more customers. Taxicabs also have high overheads costs due to compliance regulations. Since TNCs take up to 25 % of driver's fares while traditional taxi services take 8%, there have been global protests by drivers for better wages and other labor issues. Extensive research has been done regarding this topic.

1.1. Research background

It is evident that ride-hailing services are transformative to transportation systems and are changing travel behavior such as mode choice, car ownership, vehicle shedding, etc. There is a developing body of analysis by transportation researchers which assess these impacts. The lack of data pertaining to the users is limited and as such, studies have been limited to either a national level, which use NHTS data, or to the cities of San Francisco and New York, which have had a few independent surveys conducted. The presented research aims to understand the extent of the impacts of ride-hailing services on travel mobility, in particular, on public transit and car ownership. The following section elaborates on the hypothesis, aims and methodology undertaken of this research.

1.2. Hypothesis

This research is an inquiry into the impacts of ride-hailing Services on Public Transit and Car Ownership in the United States through Structural Equation Modeling (SEM) and path analysis. Due to nation-wide samples, this study chose the 2017 National Household Travel Survey as its primary data source. The main motivations of this research are to understand whether ride-hailing complements or competes with Public Transit, whether it impacts car ownership and what impact neighborhood context has on links between car ownership and ride-hailing use. It accounts for variables such as city size, neighborhood type, number of cars per household, number of drivers per household, mode to work, years since introduction of TNCs, as the research accepts the answer to the primary research question may be context related. The final result of this study finds the relationship between public transit frequency and frequency of ride-hail usage at a household level through a path analysis analyzed through SEM.

The study consists of five chapters. The first chapter briefly introduces the background and aims and methodology of the research. The second chapter reviews relevant literature to the research question and establishes the variables to be taken into consideration for this study. The next chapter expands upon the methodology and rationalizes using SEM for the research design. The fourth chapter consists of the descriptive analysis and the SEM result. The last chapter concludes the results and presents limitations of the research.

1.3. Need for the Study

Transportation networks and mobility are critical for providing access to basic services such as employment, health, education, commerce and recreation. Ride-hailing services are rapidly changing the global transportation system. With prices that can compete with transit fares (Sadowsky, 2017), the use of such services is rapidly rising. A common perception is that such services provide better accessibility, faster service, and point to point convenience. This study questions these perceptions in order to establish policies and regulations that need to be put in place to effectively cultivate this new means of transport. This can only be done by understanding the relationship ride-hailing has with other modes of transport such as public transit and private vehicles. Under this umbrella topic, this study shall determine the impact of ride-hailing services on public transit and car ownership.

1.4. Research Objectives

The research questions identified for the study are:

- *Do ride-hailing services compete with, or complement Public Transit?*
- *Are ride-hailing services replacing private vehicles?*
- *What role does Neighborhood Context play on the links between Car Ownership, Ride hailing and Public Transit?*

1.5. Research Procedure

A research methodology was formed after establishing the scope and limitations of the study.

1. Conduct a thorough study of existing literature. Study similar research papers to understand the different statistical methods used to analyze variables related to the research question. Identify and select variables to be further analyzed.

2. Conduct a descriptive analysis based on the selected variables using the selected dataset to determine trends and relationships. Choose and implement advanced statistical techniques.

3. After choosing the statistical technique, SEM in this case, formulate the model and try different iterations until the most accurate model is formulated.

4. Use the model to find the relationship between the selected variables and determine the impact of ride-hailing services on public transit and car ownership.

CHAPTER 2: LITERATURE REVIEW

There is an emerging body of literature investigating the relationship between ride-hailing services and travel behavior, such as car ownership, mode to work and vehicle miles travelled. The main goal of this chapter is to identify variables which can be further analyzed to assess the impact of ride-hailing services on travel behavior. These are identified through the existing literature and chosen for their relativity to the research goals. At the end of the segment, the results are summarized to show relevant findings.

2.1. Ride-Hailing and travel demand impacts

A large share of the body of literature pertaining to ride-hailing is related to its impact on travel demand, user and driver behavior. Although studies on car-sharing have been conducted as early as the 1990s (Muheim et al, 1999; Klintman, 1998), they largely analyzed examples of contractual sharing of cars, carpooling or car rentals. Cervero's work: City CarShare, First-Year Travel Demand Impacts (2003) is one of the earliest studies based on empirical data from TNCs in North America. The study, conducted 9 months after the introduction of Uber in San Francisco, shows preliminary evidence of ride-hailing trips subsequently gaining popularity and Uber inducing motorized travel. A demographic profile showed that most users did not own cars and active trips were being replaced by ride-hailing trips. Cervero's research showed that ride-hailing was not popular in areas well served by transit.

Further studies expanded on the demographic profile of the user base (Dias et al, 2017; Alemi et al, 2018) who were found to be well-educated, environmentally conscious, older millennials with higher incomes, residing in high-density areas.

Alemi's work (2018) used a binary logit model to identify which variables impact the use of ridesharing services among individuals born between 1965-1997 in California. The study incorporated attitudinal attributes of the users to find that young, non-Hispanic individuals were most likely to use ride-hailing. These individuals actively used technology, had 'variety-seeking' attitudes, were likely to make more long-distance business trips and travel frequently by plane (Diaz et al, 2017; Alemi et al, 2018). It was suggested, and widely accepted, that factors such as land-use mix and regional accessibility by car also significantly impacted the use of ride-hail services.
Other common trends studied in this body of literature were ride-hailing and dynamic pricing (Clewlow, 2017; Feng et al, 2017; Korolko et al, 2018; Li et al, 2016) and driver regulations (Beer et all, 2017). The common conclusion from most studies was that dynamic pricing when coupled with advanced matching, results in reduced wait times for both drivers and users and some studies (Li, et al, 2016; Feng et al, 2017) suggested new pricing methods. It was also found that driver-related regulations, primarily fingerprint-based background checks faced opposition from TNCs.

2.2. Do Ride-Hailing Services compete with, or complement Public Transit?

One of the relationships more relevant to our research is that of ride-hailing and public transit. Boisjoly's (2018) multilevel mixed-effect regression approach illustrates that vehicle revenue kilometers (VRK) and car ownership are the main factors impacting public transport ridership. Their study which captures the 2002-2015 data of 25 transit authorities from US and Canada shows that ridesharing, though not statistically significant, has a positive relationship with transit ridership. The study concludes that transit ridership has decreased over time mainly due to vehicle revenue kilometers (VRK) and car ownership. VRK is positively and significantly associated with ridership, a 10% increase in VRK is associated with an 8.27% increase in ridership. Ridesharing does not negatively impact transit ridership but may, along with bicycle-sharing systems, complement it.

This concept of ridesharing complementing transit use is seen in greater detail in Sadowsky's (2017) discontinuity regression analysis. The study uses a discontinuity regression model using variables like probability of mode choice, cost, speed, access and egress time, waiting time, trip distance to understand the relationship between public transit and ridesharing facilities by using monthly public transit ridership data from the Federal Transit Authority. The study speculates that initially, the introduction of the ridesharing service, Uber, complemented public transit by acting as a solution to last-mile connectivity. Subsequent entry of competitors like Lyft resulted in competitive prices which transformed ridesharing services from complementing public transit to competing with them (and each other). This resulted in an over-all reduction of public transit ridership.

Certain recent studies (Clewlow,2017) have proposed that ridesharing facilities complement heavy rail systems but compete with bus services and light rail services. This is subject to the frequency and quality of the various services. Most ridesharing trips replace trips that would otherwise have not been made, or made by walking, biking or public transit therefore, indirectly, ridesharing may contribute to greater VMT miles rather than reducing them.

One behavior that these studies do not capture is 'confidence of having a ride back'. If a user is confident that they have a mode of transport to reach back home, they will more likely be willing to use a combination of public transport and ride share without worrying about factors such as late-night transit service operations, availability and frequency of taxi or ride-hailing facilities, etc. (Sadowsky, 2017). Ridesharing facilities can complement public transport through this factor.

2.3. Do ridesharing services impact private vehicles?

Although there has been a considerable number of empirical studies that tried to prove the impact of ridesharing services on private vehicles (Clewlow, 2017), they have shown mixed results. Clewlow took an empirical approach to the topic by collecting data through an online self-administered travel and residential choice survey which was targeted at neighborhoods identified through the 2011-2013 ACS statistics. It was found that though car ownership among ride-hail users was comparable to non-users, 9% of the sample population disposed a vehicle since they started using ridesharing services. Rodier's (2018) analysis supported this claim but suggested a much smaller value. The total extent of increase was labeled ambiguous due to the various uncertainties involved.

Some studies like Rayle's 2015 research did not give empirical proof tying car ownership reduction with ride-hail use but had findings which suggested it. Rayle's research was on the user behavior of ridesharing services in San Francisco and it compared the travel time for public transit and rise-sharing services as well as an overall comparison between taxis and ridesharing services. The data was based on a survey of 380 users and showed that the services were most used by individuals looking for short wait times and fast point-to-point service without the inconvenience of parking. It suggested that though ridesharing affected use of personal vehicles, it did not play any significant role in ownership of personal vehicles.

Many empirical studies (Rodier, 2018; Kamargianni, et al, 2018) which explored this relationship used similar independent variables like auto ownership, trip generation, destination choice, mode choice, network vehicle travel, and land use.

2.4. Key Findings

The key findings from the literature study conducted are that users of ridesharing services are most likely to be 18-29 years old with high incomes and high educational attainments who are non-Hispanic in ethnicity. These individuals tend to embrace technology, are environmentally conscious and have variety-seeking attitudes.

The presence of ridesharing services and bicycle sharing are associated with higher levels of transit ridership but are not statistically significant. Initially the presence of ridesharing services complemented transit ridership as it acted as a last-mile connectivity solution but inter-competing prices among ride-hail services has resulted in competition with transit services, leading to low transit ridership. Models trying to determine the relationship between ride-hailing and public transport cannot capture ‘the confidence of having a ride back’. This behavior complements public transport as users are confident that they have a mode of transport to reach back home even if they do not use their private vehicles. Most ridesharing trips replace trips which otherwise would not have been made, walking, biking and public transport trips. This is reflected in findings that ridesharing facilities complement heavy rail systems but compete with bus services and light rail services.

Ride-hailing users who disposed of a vehicle use ride-hailing more frequently. Ride-hailing has slightly reduced car-ownership, but it has increased total overall VMTs and subsequently greenhouse emissions.

More diverse land-use mix and greater regional accessibility by car have a higher probability of using ride-hailing services. TNCs are less likely to operate in cities where background checks are mandatory for drivers.

One of the critical limitations of the literature studied was the absence of neighborhood context or urban form as potential factor which could impact the relationship of ride-hailing and car ownership. This is addressed in our research.

CHAPTER 3: RESEARCH DESIGN

The biggest shortcomings seen through the literature analysis was the lack of data provided by the TNCs. In an effort to capture nation-wide trends with relative accuracy, the NHTS 2017 data set was selected as the primary data source. Data compilation was also undertaken to create a data set of all cities in the US which had operational TNCs. The data set was compiled up to May 2017 as the NHTS 2017 data collection was done between April 2016 to May 2017.

3.1. National Household Travel Survey (NHTS), 2017

The NHTS is a national level sample survey conducted by the Federal Highway Administration (FHWA). It was first conducted in 1969 and was originally known as the Nationwide Personal Transportation Survey (NPTS). It has since been conducted 7 more times, the latest being 2017. It is conducted every 5-9 years and is the primary source of travel behavior for the American citizens[7]. As per its website, it is the only source of national data that allows one to analyze trends in personal and household travel. The 2017 survey was conducted from March 2016- May 2017 and was the first survey to include questions about ride-hailing services. Users were asked how many ride-hailing trips they made in the last 30 days through the survey question "how many times have you purchased a ride-hailing service with a smartphone ride-hailing application (e.g., Uber, Lyft, or Sidecar) in the past 30 days?". Ridehailing was also included in the taxi trip mode[8].

---

[7] Federal Highway Administration. (2017). 2017 National Household Travel Survey, U.S. Department of Transportation, Washington, DC. Available online: https://nhts.ornl.gov.

[8] Trends in Taxi Usage and the Advent of Ridehailing, 1995–2017; M. W. Conway, D. Salon, D. King, Arizona State University, 2018; NHTS Data for Transportation Applications Conference Washington, DC

The NHTS sample is selected through a random sampling based on addresses. It is entirely voluntarily and in 2017 the survey had a response rate of 15%. A total of 129,696 households were surveyed which included 264,234 individuals.

3.1.1. Data Cleaning and Management

The sample for this study contains 85920 households, about 66% of the total sample set. The original data set were present in 4 different subsets: At a household level, at a person level, at a trip level and at a vehicle level.

First these files were cleaned so they included only the variables selected. Then, the files at the person, trip and vehicle level were aggregated to the household level and matched to the corresponding entries. This resulted in a single file which contained all the selected variables present in the NHTS 2017 database. The public database did not contain any information about the 'location' of the individual entries.

NHTS was then approached and we were provided a database that contained the census tract for each individual entry. This information was joined to the information we had isolated from the public database. The details of the variables used in the data set are given in Appendix A.

Based on the census tracts, each entry was then assigned a Composite Index Score i.e. Ewing's Sprawl Index calculated at a census tract level[9]. As only Ewing's sprawl Index only exists for 85% of census tracts. Some entries which did not receive a score had to be removed.

---

[9] Geographic Information Systems and Science for Cancer Control. (n.d.). Retrieved from https://gis.cancer.gov/tools/urban-sprawl/

The data was then further cleaned and all NHTS entries which contained information such as 'Not ascertained, don't know, I prefer not to answer' etc. for any of the selected variables were removed. The final data set for this study contained 66% of the original entries.

The next step was to fix the categorical variables. The reference categories were recoded to suit the study needs. Certain variables had to be disaggregated and coded in binary in order to be used.

Finally, each entry was assigned a variable based on the number of years since a TNC was introduced to the city. This data had been compiled manually and was matched to the dataset through the census tracts level.

Finally, mean centering was done on 'Urban Form' (Ewing's Sprawl index) and 'Number of years since TNC was introduced' variables as their moderated effect was to be calculated in the final model.

It should be noted, that the variable 'Number of years since TNC was introduced' contains information regarding the duration of only UberX and Lyft in an area. While uber (formerly called UberCab) started in 2009, it was initially a luxury car service which charged 1.5 times the price of a regular taxi. Uber X, which is the affordable, most widely used service, started in 2010 and was more accurate as a starting point than Uber luxury cars.

3.1.2. Variable Selection

Several variables were identified through which the impact of ridesharing services on travel behavior can be evaluated. It is evident that socio-economic factors play an important role in the usage of ride-hailing services, hence the socio-economic variables selected for this study are age, race, household size, income level, educational attainment and house ownership. Some other variables which were identified to be important from the literature study are ‘Online Services’, defined by variables like web use, use of smartphones, frequency of online delivery etc.

Urban form, urban-rural locations and density were other recurring factors across the empirical studies. There was strong evidence that the relationship between public transit and ride hailing was influenced by urban sprawl. For this research Ewing’s sprawl Index was selected to characterize Urban form.

There was also evidence uncovered from the literature analysis that introduction of TNCs, and number of TNCs played an important role in how ride hailing impacted car ownership. As the studies looked at, gave inconclusive results, it was decided to incorporate ‘time since introduction of TNC’ as an important variable

The multiple independent variables that were selected underwent basic descriptive analysis (Expanded on in Chapter 4) to determine relationships. Certain variables were then dropped from the model as it was found that they did not have a significant association or were causing multicollinearity errors.

3.2. Methodology

As stated in Chapter 1, two types of analysis were conducted. The preliminary analysis includes basic descriptive data to understand the user base. An advanced statistical analysis in the form of Path analysis through Structural Equation Modelling was then conducted to understand the precise impacts of ridesharing on public transit and car ownership.

Although various statistical methods were considered such as Machine Learning, Propensity score matching and Multiple-multivariate regression, Path Analysis through Structural Equation Modelling was chosen over other methods. It was chosen as Path Analysis helps build causal relationships between variables and these relationships can be bi-directional. SEM affords multivariate analysis while incorporating other analysis methods which gives it a flexibility to create detailed, complex models. It also allows for mediation and moderation effects of certain variables to be calculated as well. The Path Analysis was conducted through the LAVAAN package in R, which gave it the ability to standardize the parameters and calculate fit measures for the model.

3.2.1. Path Model

The path model constructed, composes of one latent variable (Online Activities), three regressions and one residual covariance. 'Frequency of Public Transit use in a month per household', 'Frequency of ride-hailing use in a month per household' and 'car ownership per driver per household' are the three dependent variables. Car-ownership acts as an independent variable for both the public transit and ride-hailing variables. 'Number of years since ridesharing was introduced' is the independent rideshare variable which is used for calculating the moderating effect of neighborhood context on car ownership.

The variables are explained in Appendix A

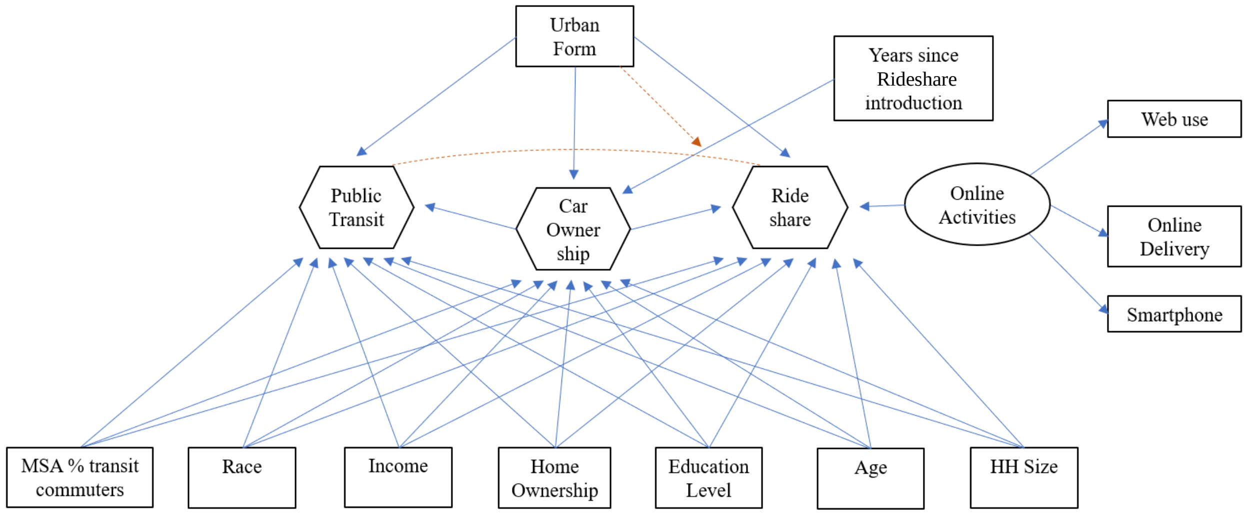

Figure 1: Path Model

The final model is described as (abbreviated):

Model <- ʻ Online Activities = ~ Web-Use + Smartphones + Delivery

Car Ownership ~ PT% $_{MSA}$ + HH_{size} + HH_{Income} + HO + EAttain + Age + Race + CI + $Years_{TNC}$ + CI: $Years_{TNC}$

Public Transit ~ PT% $_{MSA}$ + HH_{size} + HH_{Income} + HO + EAttain + Age + Race + CI + Car Ownership

Rideshare ~ PT% $_{MSA}$ + HH_{size} + HH_{Income} + HO + EAttain + Age + Race + CI + Car Ownership + Online Activities ʼ

Where Online Activities in a latent variable created;

Web-Use = frequency of internet use

Smartphones = frequency of smartphone use

Delivery = frequency of online shopping in a month;

Car Ownership = No. of Cars per driver in a household

$PT\%_{MSA}$ = Percent of Transit Commuters in an MSA

HH_{size} = household size

HH_{Income} = household income

HO = home ownership status

EAttain = educational attainment of head of household

Age = age of head of household

Race = race of head of household

CI = Ewing's sprawl index for the appropriate census tract

YearsTNC = No. of years since introduction of either UberX or Lyft

Public transit = No. of times public transit is used by a household in a month

Rideshare = No. of times rideshare is used by a household in a month

For the entire formula used please refer to Appendix B

CHAPTER 4: RESULTS

This chapter presents the findings of the descriptive and statistical analysis conducted and interprets them.

4.1. Descriptive Analysis: Relationship between Ride-hailing Usage and Selected Variables

First a preliminary analysis was conducted to determine the relationships between ride-hailing and some selected variables to understand their impact on frequency of ride-hailing usage. NHTS 2017 was used as the base dataset. The data was categorized according to city types. Out of the 5 city types present, only 4 were taken into consideration. They were small-town, second city, suburban and urban. The rural city type was excluded. Their distribution is shown in the following graph:

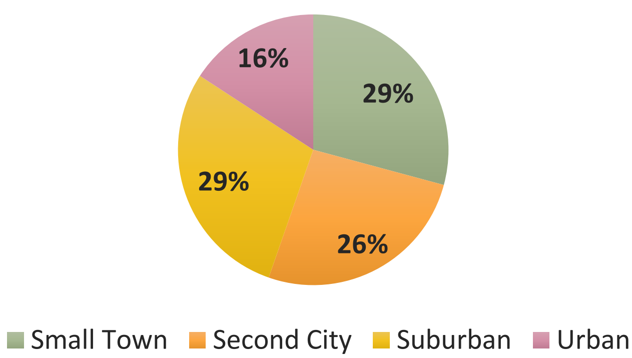

Figure 2: Distribution of City Types in NHTS 2017 data set

4.1.1. Transit Mode to Work at different city types

The first variable to be considered was the mode to work, specifically the frequency of people using transit services to go to work. As the original dataset included various modes, it was recoded into 2 classes: 1 = people who take transit to work and 0 = people who do not take transit to work. This variable was then compared to the number of times these users used ridesharing facilities in a month. The results are displayed in Figures 3 and 4.

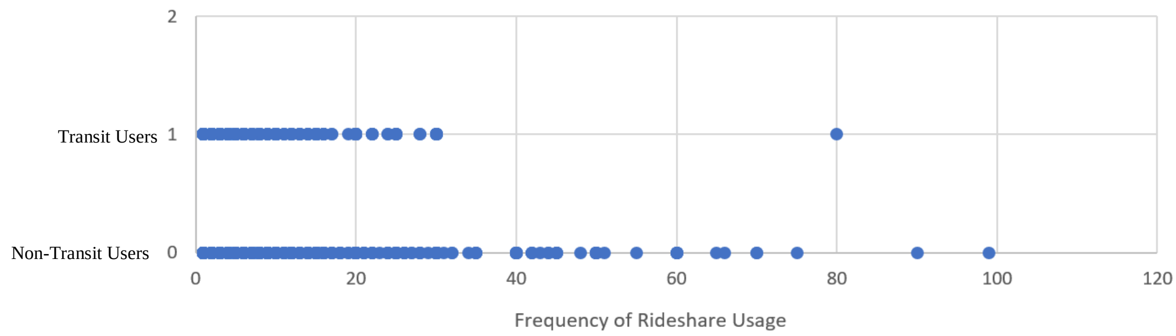

Figure 3: Frequency of Ride-hailing Usage for individuals who use transit and those who do not

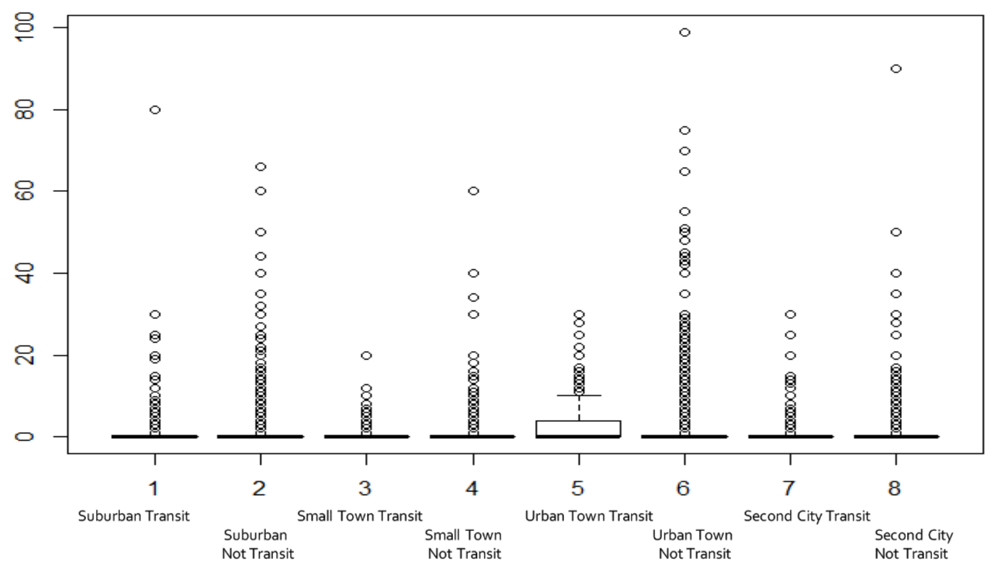

Figure 4: Mode to work by different city types and frequency of ride-hailing facility usage

	Suburban Transit	Suburban Not Transit	Small Town Transit	Small Town Not Transit	Urban Town Transit	Urban Town Not Transit	Second City Transit	Second City Not Transit
Min	0	0	0	0	0	0	0	0
1st	0	0	0	0	0	0	0	0
median	0	0	0	0	0	0	0	0
Mean	1.21	0.32	0.52	0.11	2.68	0.99	0.89	0.27
third	0	0	0	0	4	0	0	0
Max	80	66	20	60	30	99	30	90

Figure 5: Quartile values of mode to work for different city types vs frequency of ride hail services per month

As the box chart cannot be seen clearly due to large number of 0 values (for frequency of ride-hailing services usage), Figure 5 is given to depict the value of the box chart. As can be seen, the average ride-hailing frequency for Transit is much higher in most city types by approximately 4 times. The difference is greatest in small towns and least in Urban Towns indicating that rider sharing services complement public transit more in small towns and less in urban areas. It should also be noted that the max is higher in non-transit categories indicating that some people who do not take transit to work travel almost 99 times per month using ride-hailing services.

4.1.2. VMT and Ride-hailing by City Type

The next variable to be considered is VMT. It is compared at a household level to the frequency of ride hail services usage of the household. Figure 6 to 10 shows the different relations obtained by a simple scatterplot for the 4 city types and as a whole. From the scatterplot, the R^2 value is determined to see the degree of co-relation between the 2 variables. It should be noted that all values have been aggregated to a household level.

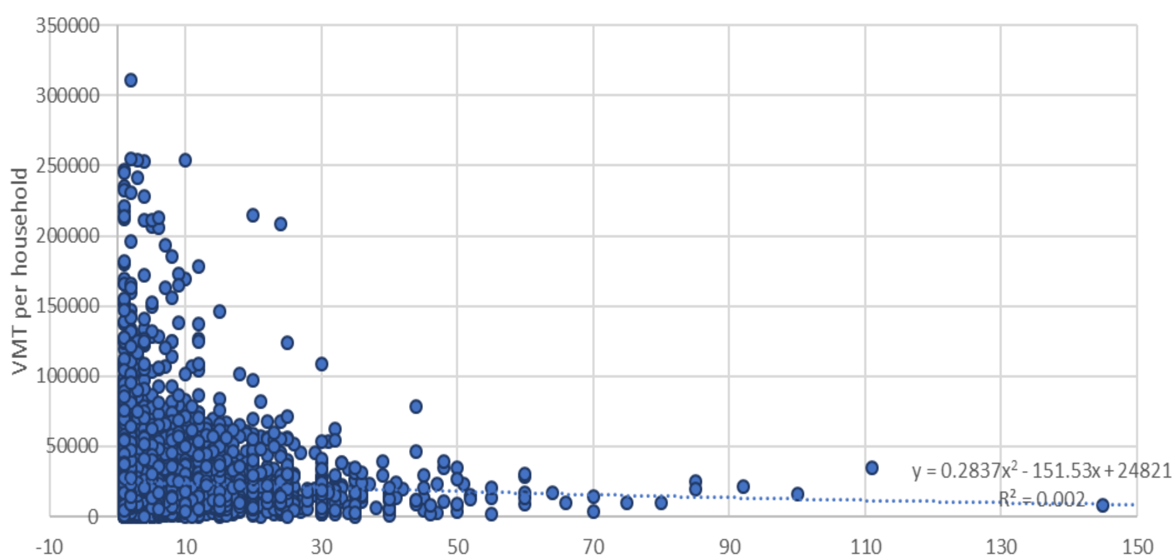

Figure 6: VMT by Total Households

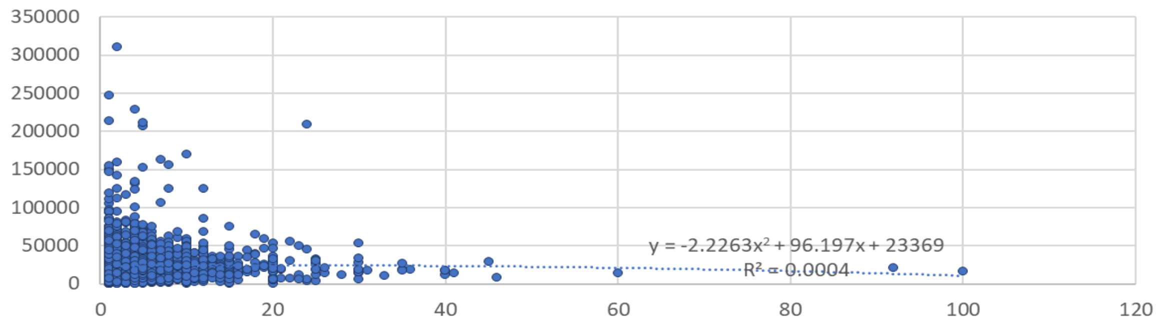

Figure 7: VMT by Second City Households

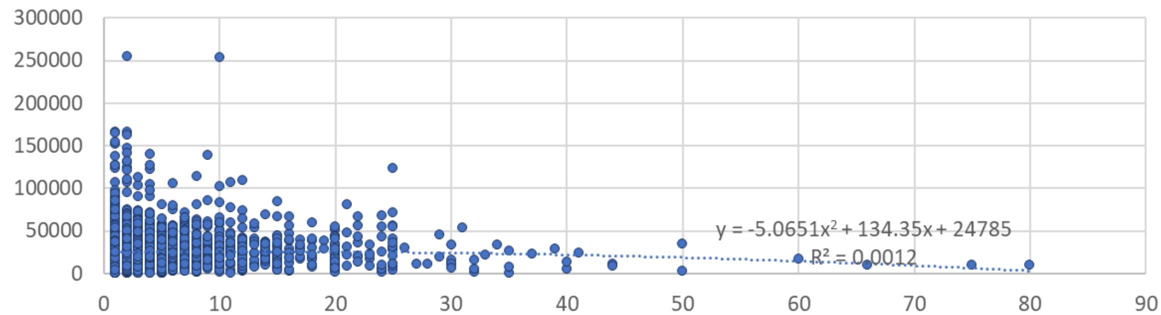

Figure 8: VMT by Suburban Households

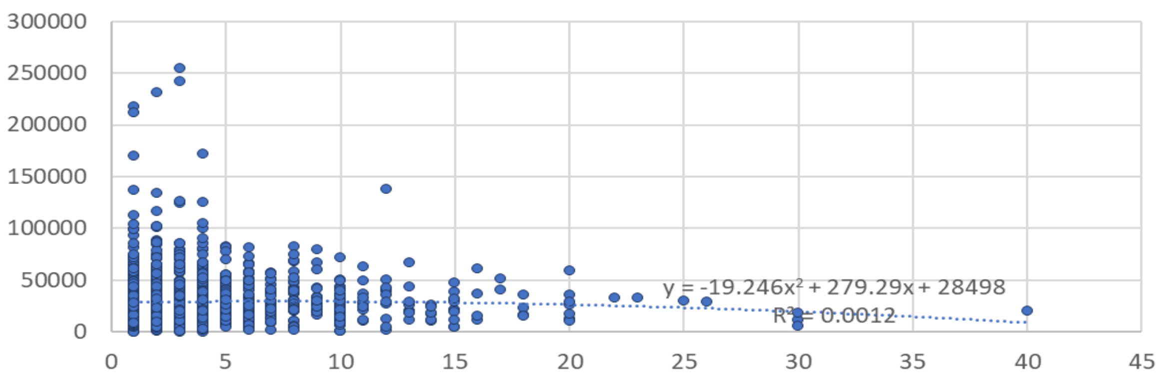

Figure 9: VMT by Small Town Households

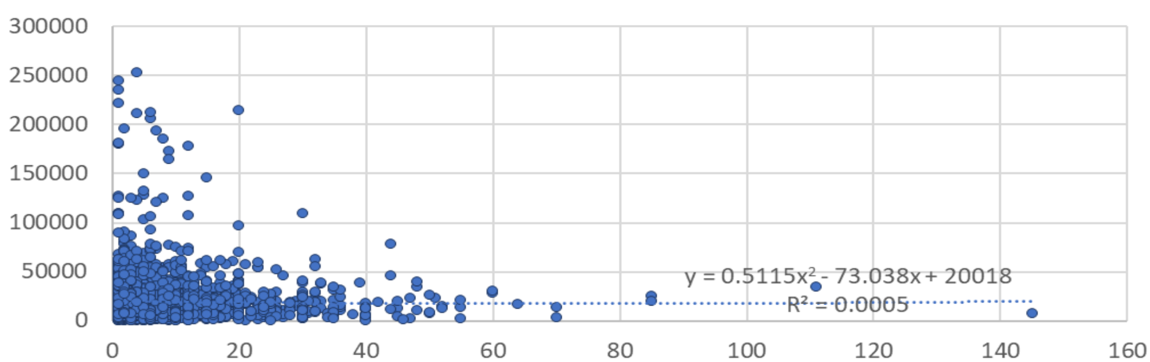

Figure 10: VMT by Urban Households

As is observed from the graphs, the R^2 values range from 0.0005 to 0.002 indicating a very small relationship between the two variables.

Urban Areas had the only positive co-relation between frequency of ride-hailing usage and VMT in all four city types. This indicates that the more miles travelled in a household, the more likely they are of using ride-hailing services. Whereas, in other city types, the more miles travelled in a household, the less likely they are of using ride hail services. Among the insignificant R2 values, small towns have a slightly higher co-relation than urban areas indicating that VMT plays a larger role in ride hail frequency usage in small towns and very little in urban towns.

As the R^2 values are so small, VMT is not considered a good fit and is dropped from the final model.

4.1.3. No. of Cars per Household and Ride-hailing Frequency by City Type

The last variable to be considered is the number of cars per household. This is again compared to the frequency of ride-hailing services usage aggregated to a household level. In Figures 11 to15, this relationship is depicted through simple scatterplots categorized by different city types. The R^2 value of each relationship is determined to calculate the extent of co-relation between the variables.

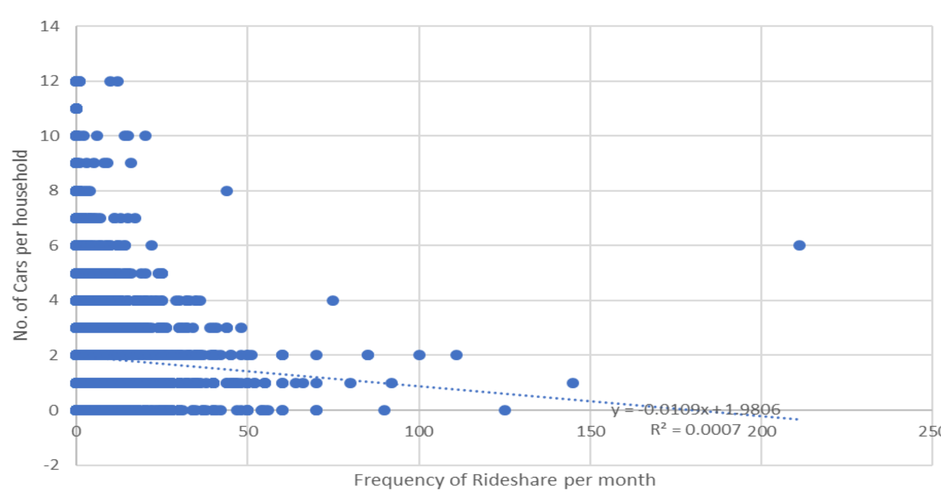

Figure 11: No of Cars per total household vs frequency of ride-hailing service usage per month

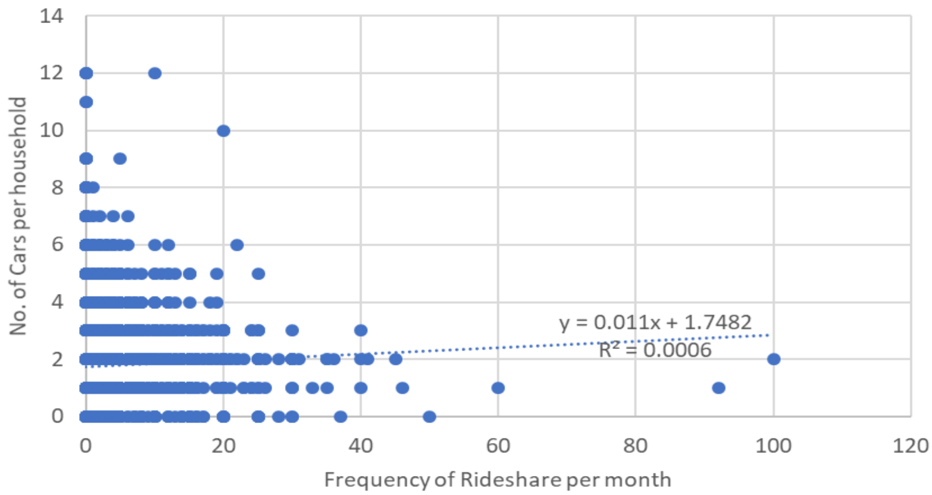

Figure 12 : No of Cars per Second City households vs ride-hailing

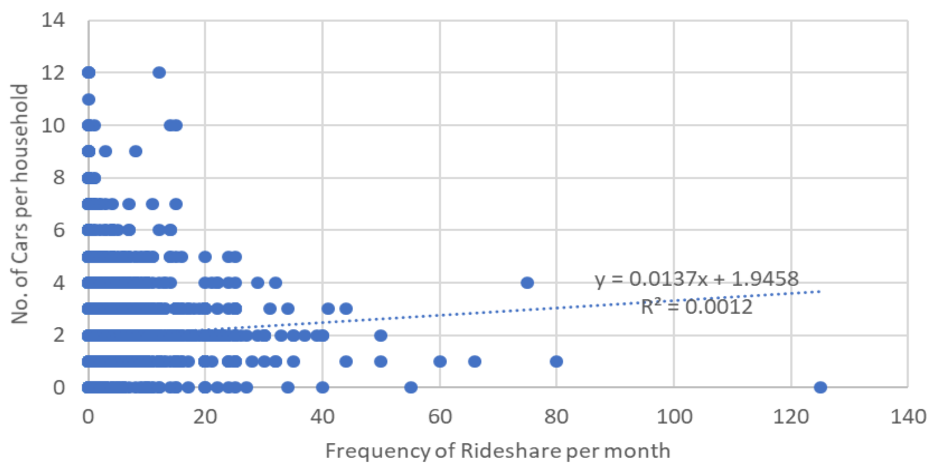

Figure 13: No of Cars per Suburban households vs ride-hailing

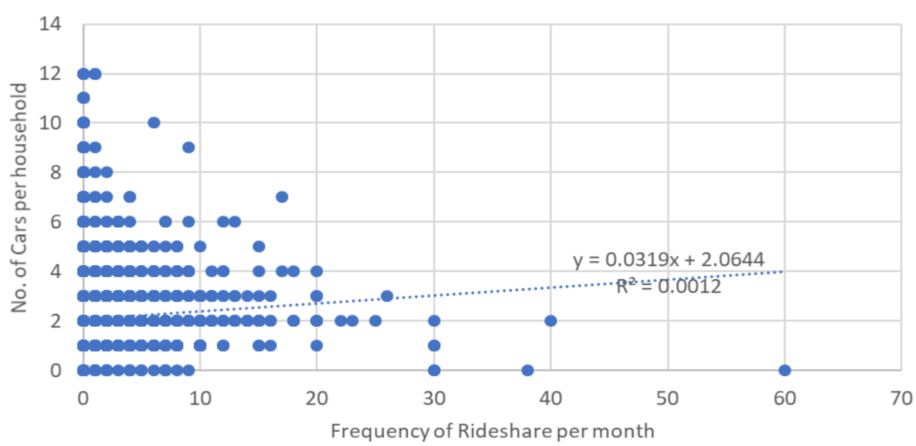

Figure 14: No of Cars per Small Town households vs ride-hailing

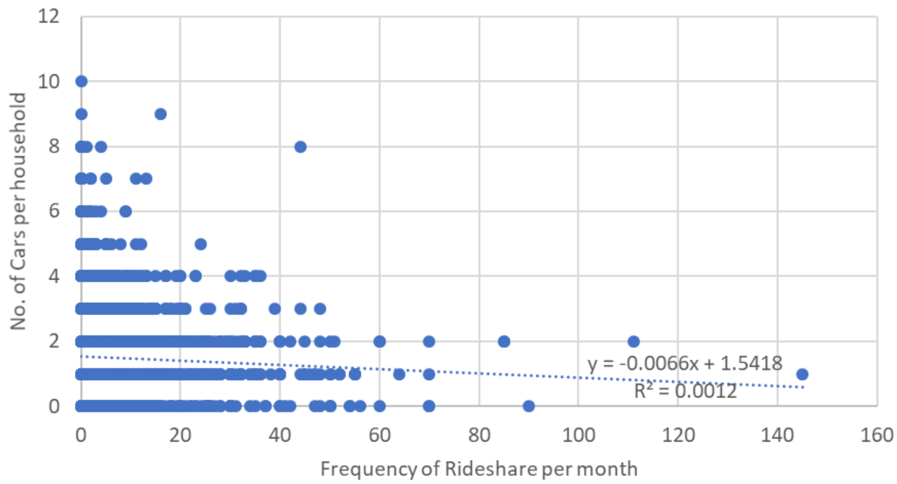

Figure 15: No of Cars per Urban households vs ride-hailing

As can be seen from the figures above, the two variables had a positive co-relation for all city types except urban yet the whole has a negative co-relation. This is interesting, because the urban city is the smallest category for this case. The R2 values are insignificant indicating little co-relation. They range between 0.0006 to 0.0012.

Since the descriptive analysis showed little co-relation between car ownership and use of ridesharing services, there arose a need to consider other factors. The city types also offered interesting findings and it was evident that they were important, but it was decided that the NHTS classification was not expansive enough to use, urban form was therefore the important factor and a new variable for it was needed. Ewing's Sprawl Index was therefore introduced to the model, replacing the city type classifications of NHTS.

4.2. Path Analysis

The model for this research as described in Chapter 3 was run using the LAVAAN package in R. The results of the SEM are given in detail in Annexure 3. The results of the model are as follows :

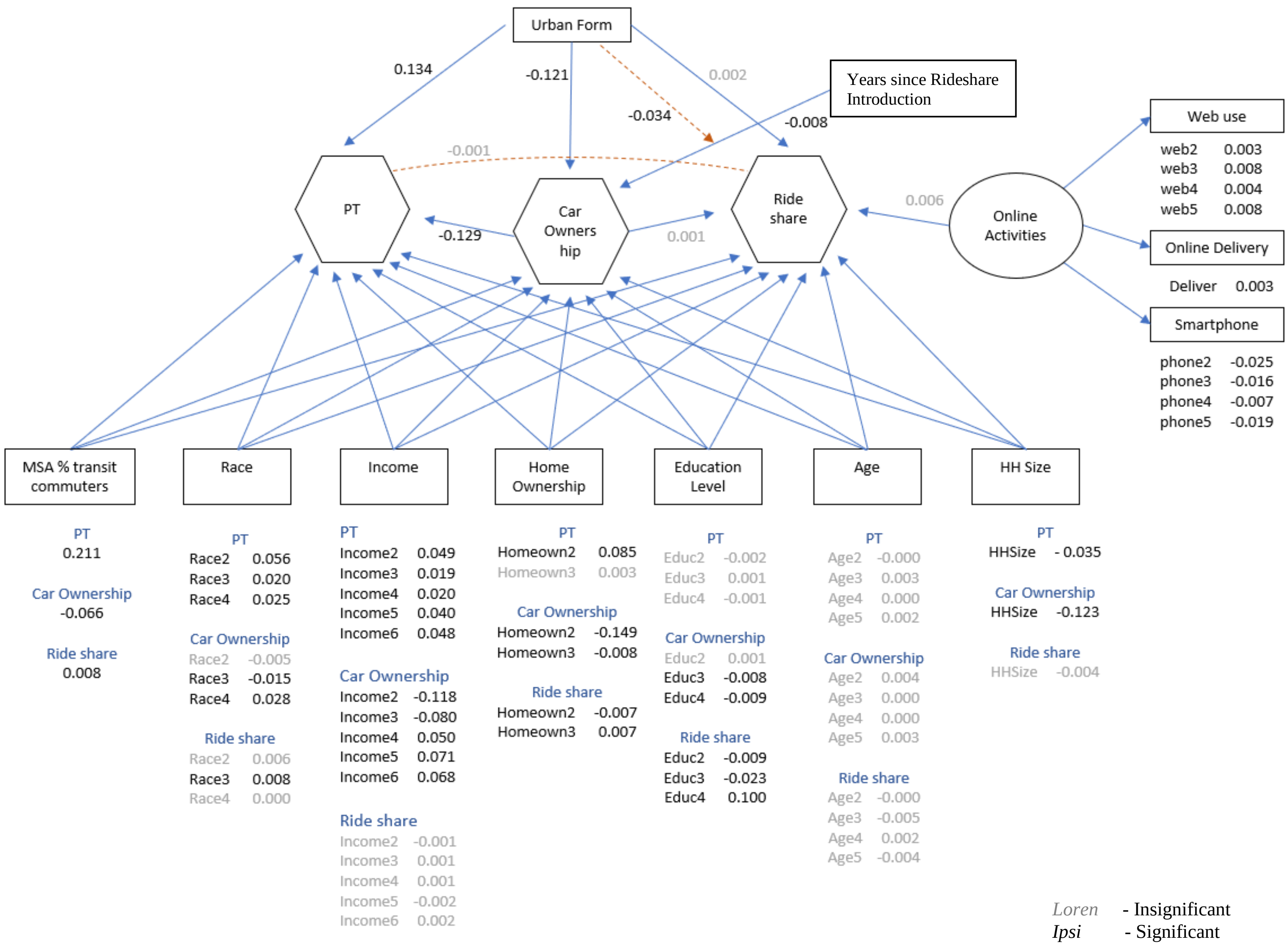

Figure 16: Path Analysis of Model with Outputs

It can be seen from Figure 16 that the model yielded interesting results. In terms of the socio-demographic variables, age of the household head plays no significant role on any of the dependent variables. Race is mostly only significant towards frequency of public transport use. Educational attainment is significant only towards ride-hailing frequency. Having the highest educational attainment till high school or less makes an individual have a negative impact on frequency of ride-hailing services, whereas if a person has some college attainment, they have a positive impact on ride-hailing frequency. The relationship of income with ride-hailing use is insignificant and is different categories of income have different impacts on ride-hailing use with no discernable pattern.

Household size has a negative relationship with both public transit use and car ownership. It has an insignificant impact on ride-hailing use. Renters have a negative relationship with car ownership and ride-hailing usage but a positive relationship with public transit use.

The percent of transit commuters to work have a very slight but positive relation to ride-hailing usage. They have a negative impact on car ownership.

When looking at the latent variable: Online Activities, an interesting trend is seen. Although web use and online delivery have positive relations with the latent variable, use of smart phone has negative impacts. The overall impact of online activities is insignificant on ride-hailing usage.

4.2.1. Impact of Urban Form and Duration of TNC presence on Dependent Variables

One of the key findings of this research is the link between car-ownership and ride-hailing and how urban form affects it.

As shown in Figure 17, both Sprawl Index and Year since start of a Rideshare[11] have negative link with Car ownership. When the moderation effect of time on the sprawl index is considered, this negative relationship becomes larger.

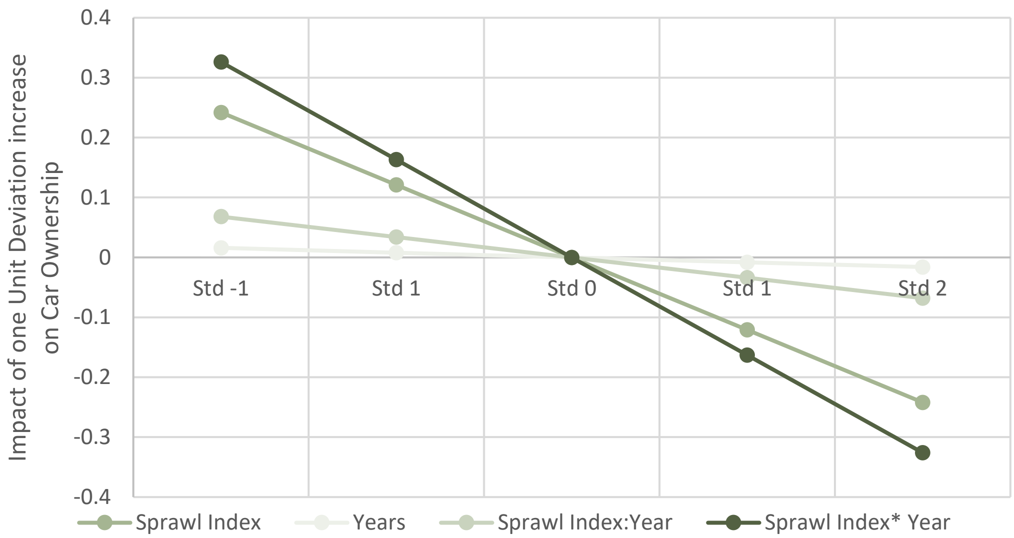

Figure 17: Change in Car Ownership through 1-unit deviation in Sprawl Index and Years since TNC

This can be read as ΔSprawl Index + ΔYears +ΔSprawl Index: Years = Δ Sprawl Index* Years

Therefore, ΔCar Ownership = Δ Sprawl Index* Year

ΔCar Ownership = - 0.121 -0.034-0.008 = -0.163

So a one unit change in urban form and duration of TNCs will result in a decrease by 0.163 in car ownership.

This suggests that neighborhood context does impact the link between car ownership and ride hailing services. This suggests that ride hailing in denser neighborhoods reduces car ownership more strongly than in sprawled neighborhoods

---

[11] Either UberX or Lyft

Next let us consider the dependent variable: ride-hailing use. Both Urban Form and Car Ownership have insignificant effects on ride-hailing. This shows that although the impact of car ownership on ride-hailing is not explained through this model, the impact of ride-hailing on car ownership is explained.

Finally, let us consider Public Transit Use. Urban form has a positive, significant and comparative large impact on public transit use. This means that denser areas induce public transit use. Car ownership plays a negative impact on public transit and public transit and ride-hailing have an insignificant covariance.

CHAPTER 5: CONCLUSION

Ride-hailing services have transformed today's transportation systems. In order to make informed decisions to manage and cultivate the change bought by ride-hailing, the transforming travel patterns need to be understood. Unfortunately, the rapid growth of such services has left us with a severe data shortage. The purpose of this research is to try and fill this data gap by defining the links between ride-hailing, public transit and car ownership.

Using Path Analysis implemented through Structural Equation Modeling to formulate a model describing the relationships between sociodemographic variable, technology dependent variables, use of public transit, car ownership and ride-hailing services, this research was able to defining these links and study their causal relationships.

The first of the three key questions analyzed in this study was the relationship between public transit and ride-sharing. Although the statistical model used in the research estimated that rideshare and public transit had an insignificant covariance, initial findings suggested that ride sharing services complement public transit more in small towns and less in urban areas. This gives scope for further research about the topic.

The second topic analyzed through this model was the link between ridesharing and car ownership. The model estimated that ride-hailing had a significant and comparatively large impact on car ownership. Due to the bidirectional nature of the model, we were able to study the reverse relationship as well. The model did not show car-ownership having a significant impact on frequency of rideshare use.

The final relationship to be studied was the impact of neighborhood context on the links between car ownership and ride-hailing. It was suggested through a moderation estimation that urban form does play a significant role in impacting the role of rideshare on car ownership. The length of duration since the introduction of TNCs in a city plays an important role on car ownership. The longer TNCs have been around, the smaller the value of car ownership is. Denser Urban forms deepen this relationship while sprawled neighborhoods weaken the correlation.

5.1. Research Limitations

The study is limited to the data provided by the National Household Travel Survey, 2017. The sample size consists of 129,696 households (0.15% of all households in the US) with a 15% response rate. There is a massive amount of mobility data owned by private mobility agencies which is not shared with the public. This data can give an in-depth overview of the impact of ride-hailing services. Apart from the variables considered in the study, there are several more contextual variables which undoubtedly have not been considered. Only data from 2017 has been analyzed and therefore the study does not give any temporal trends. Data regarding establishment of ride hailing services in individual cities has been compiled using various sources such as news articles and blogs as no such public dataset exists.

5.2. Future Research and Policy Implications.

Although this research is unable to provide answers for all the questions, it gives enough evidence to warrant further studies in certain areas such as exploring the relationship between city types, rideshare usage and public transit. This research can serve as a base point for further research to completely understand the links between the three.

The evolving transport systems of today's world are leaving planners with little to no data to make informed, viable decisions. This research presents opportunities for planners in cities where TNCs have recently been introduced to evaluate their policies and regulations to better respond to the undeniable rise of shared economy. Ridesharing and other Shared economy models like Airbnb, crowdfunding and couch surfing can have significant externalities on urban development aspects such as neighborhood context, parking requirements, transportation networks, retail sector etc. and proper research needs to be done to effectively manage them.

REFERENCES

1. Alemi, F., Circella, G., Handy, S., & Mokhtarian, P. (2018). What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California. Travel Behaviour and Society, 13, 88-104.

2. Beer, R., Brakewood, C., Rahman, S., & Viscardi, J. (2017). Qualitative Analysis of Ride-Hailing Regulations in Major American Cities. Transportation Research Record: Journal of the Transportation Research Board, 2650, 84-91.

3. Boisjoly, G., Grisé, E., Maguire, M., Veillette, M.-P., Deboosere, R., Berrebi, E., & El-Geneidy, A. (2018). Invest in the ride: A 14 year longitudinal analysis of the determinants of public transport ridership in 25 North American cities. Transportation Research Part A: Policy and Practice, 116, 434-445.

4. Clewlow, R. R., & Mishra, G. S. (2017). Disruptive transportation: The adoption, utilization, and impacts of ride-hailing in the United States. University of California, Davis, Institute of Transportation Studies, Davis, CA, Research Report UCD-ITS-RR-17-07.

5. 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. Transport Policy, 45, 168-178.

6. Rodier, C. (2018). The Effects of Ride Hailing Services on Travel and Associated Greenhouse Gas Emissions.

7. Sadowsky, N., & Nelson, E. (2017). The impact of ride-hailing services on public transportation use: A discontinuity regression analysis.

8. Singh, A. C., Astroza, S., Garikapati, V. M., Pendyala, R. M., Bhat, C. R., & Mokhtarian, P. L. (2018). Quantifying the relative contribution of factors to household vehicle miles of travel. Transportation Research Part D: Transport and Environment, 63, 23-36. T

9. Conway M. W., Salon D., King D.(2018).Trends in Taxi Usage and the Advent of Ridehailing, 1995–2017; Arizona State University. NHTS Data for Transportation Applications Conference Washington, DC

10. Muheim, P., & Reinhardt, E. (1999). Carsharing: the key to combined mobility. World Transport Policy & Practice, 5(3).

11. Klintman, M. (1998). Between the Private and the Public: Formal Carsharing as Part of A Sustainable Traffic System. An Exploratory Study (No. KFB-MEDD-1998-2).

12. Cervero, R. (2003). City CarShare: First-year travel demand impacts. Transportation research record, 1839(1), 159-166.

13. 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(6), 1307-1323.

14. Feng, G., Kong, G., & Wang, Z. (2017). We are on the way: Analysis of on-demand ride-hailing systems. Available at SSRN 2960991.

15. Korolko, N., Woodard, D., Yan, C., & Zhu, H. (2018). Dynamic pricing and matching in ride-hailing platforms. Available at SSRN.

16. Li, S., Fei, F., Ruihan, D., Yu, S., & Dou, W. (2016, December). A dynamic pricing method for carpooling service based on coalitional game analysis. In 2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS) (pp. 78-85). IEEE.

APPENDIX A: SUMMARY OF VARIABLES

This section gives a detailed summary of all the variables used in the descriptive analysis and the final model. Some of the Categorical variables are unordered to set the reference term as a median value.

Table A.1. Summary of Variables

Name and Description	Variable Code	Source	Type	Categories
Home Ownership Status	HOMEOWN	NHTS	C	01=Own 02=Rent 03=Some other arrangement
Count of household vehicles	HHVEHCNT	NHTS	N	
Count of household vehicles per driver	CarsPerDriver	-	N	
Household income	HHFAMINC	NHTS	C	2 = Less than $10,000 3 = $10,000 to $24,999 1 = $25,000 to $49,999 4= $50,000 to $99,999 5 = $100,000 to $199,999 6 = $200,000 or more
Frequency of Smartphone Use to Access the Internet	SPHONE	NHTS	C	02 = Daily 03 = A few times a week 01 = A few times a month 04 = A few times a year 05 = Never
Number of drivers in household	DRVRCNT	NHTS	N	
Race of household respondent	HH_RACE	NHTS	C	1 = Non-Hispanic White 2 = Non-Hispanic AA 3 = Hispanic 4 = Others
Age of household respondent	R_AGE	NHTS	C	2 = 24 and under 3 = 25-34 1 = 35 -54 4 = 54-65 5 = 65+
Educational Attainment of head of household	EDUC	NHTS	C	2 = less than high school 3 = high school 1 = some college 4 = bachelors and above
Count of Ride-hailing App Usage in a month aggregated at household level	RIDE-HAILING	NHTS	N	

Name and Description	Variable Code	Source	Type	Categories
Count of household members	HHSIZE	NHTS	N	
Frequency of internet use	WEBUSE17	NHTS	C	02=Daily 03=A few times a week 01=A few times a month 04=A few times a year 05=Never
Count of Times Purchased Online for Delivery in Last 30 Days	DELIVER	NHTS	N	
Count of Public Transit Usage in Last 30 Days	PTUSE	NHTS	N	
Ewing's Sprawl Index at Census Tract level	Composite Index	Ewing	N	
Percentage of work trips taken by public transit on total work trips at an MSA level	msa_pt	American Fact Finder	N	
Years since Introduction of TNCs Uber X or Lyft to city up to June 2017	Years	-	N	
Trip distance in miles for personally driven vehicle trips	VMT	NHTS	N	
Urban / Rural indicator - Block group (City Type)	HBHUR	NHTS	C	C=Second City R=Rural S=Suburban T=Small Town U=Urban

Code: In Column 'Type' N = Numeric Variables; C = Categorical Variables

In column ' categories', classes in red are the base reference class in the multivariate model.

APPENDIX B: CODE USED TO DESCRIBE FINAL MODEL

The model for the research is described by 'final.model'. The dataset is called 'tryfive' and the fitted output is termed 'final.fit'. The model is standardized, and fit measures are calculated.

```
final.model <- '

OnlineActivities =~ WEBUSE17+ DELIVER + SPHONE

CarsPerDriver ~ msa_pt + HHSIZE + income2 + income3 + income4 + income5 + income6 +
homeown2 + homeown3 + educ2 + educ3 + educ4 + Age2 + Age3 + Age4 + Age5 + Race2 + Race3 +
Race4 + CIMean:YearsMean  + YearsMean+ CIMean

PTUSE ~ msa_pt + HHSIZE + income2 + income3 + income4 + income5 + income6 + homeown2 +
homeown3 + educ2 + educ3 + educ4 + Age2 + Age3 + Age4 + Age5 + Race2 + Race3 + Race4 +
CIMean + CarsPerDriver

RIDE-HAILING ~ msa_pt + HHSIZE + income2 + income3 + income4 + income5 + income6 +
homeown2 + homeown3 + educ2 + educ3 + educ4 + Age2 + Age3 + Age4 + Age5 + Race2 + Race3 +
Race4 + CIMean + CarsPerDriver + OnlineActivities

PTUSE ~~ RIDE-HAILING

'

final.fit<- sem(final.model, data= tryfive)

summary (final.fit, standardized=TRUE, fit.measures=TRUE, rsquare=TRUE)
```

APPENDIX C: R SUMMARY OUTPUT FOR MODEL

Results of SEM for the model:

lavaan 0.6-4 ended normally after 300 iterations

Optimization method NLMINB

Number of free parameters 75

Number of observations 85919

Estimator ML

Model Fit Test Statistic 22508.049

Degrees of freedom 78

P-value (Chi-square) 0.000

Model test baseline model:

Minimum Function Test Statistic 75389.377
Degrees of freedom 147

P-value 0.000

User model versus baseline model:

Comparative Fit Index (CFI) 0.702

Tucker-Lewis Index (TLI) 0.438

Loglikelihood and Information Criteria:

Loglikelihood user model (H0) -1008250.218

Loglikelihood unrestricted model (H1) -996996.193

Number of free parameters 75

Akaike (AIC) 2016650.435

Bayesian (BIC) 2017352.522

Sample-size adjusted Bayesian (BIC) 2017114.170

Root Mean Square Error of Approximation:

RMSEA 0.058

90 Percent Confidence Interval 0.057 0.058

P-value RMSEA <= 0.05 0.000

Standardized Root Mean Square Residual:

SRMR 0.032

Parameter Estimates:

Information Expected

Information saturated (h1) model Structured

Standard Errors Standard

Latent Variables:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
OnlineActivities =~						
WEBUSE17	1.000				0.416	0.620
DELIVER	-2.818	0.048	-58.313	0.000	-1.172	-0.252
SPHONE	2.270	0.045	50.397	0.000	0.944	0.805

Regressions:

	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
CarsPerDriver ~						
msa_pt	-0.583	0.032	-18.273	0.000	-0.583	-0.066
HHSIZE	-0.061	0.002	-36.168	0.000	-0.061	-0.123
income2	-0.335	0.010	-33.274	0.000	-0.335	-0.118
income3	-0.142	0.007	-20.820	0.000	-0.142	-0.080
income4	0.063	0.005	11.760	0.000	0.063	0.050

income5	0.099	0.006	16.670	0.000	0.099	0.071
income6	0.157	0.009	18.314	0.000	0.157	0.068
homeown2	-0.201	0.005	-41.868	0.000	-0.201	-0.149
homeown3	-0.060	0.024	-2.539	0.011	-0.060	-0.008
educ2	0.002	0.013	0.173	0.863	0.002	0.001
educ3	-0.014	0.006	-2.209	0.027	-0.014	-0.008
educ4	-0.011	0.004	-2.567	0.010	-0.011	-0.009
Age2	0.016	0.013	1.196	0.232	0.016	0.004
Age3	0.000	0.006	0.004	0.997	0.000	0.000
Age4	0.000	0.005	0.046	0.963	0.000	0.000
Age5	0.004	0.005	0.866	0.386	0.004	0.003
Race2	-0.011	0.007	-1.596	0.110	-0.011	-0.005
Race3	-0.044	0.009	-4.625	0.000	-0.044	-0.015
Race4	0.053	0.006	8.508	0.000	0.053	0.028
CIMean:YearsMn	-0.000	0.000	-10.364	0.000	-0.000	-0.034
YearsMean	-0.002	0.001	-2.273	0.023	-0.002	-0.008
CIMean	-0.003	0.000	-34.707	0.000	-0.003	-0.121
PTUSE ~						
msa_pt	16.108	0.254	63.353	0.000	16.108	0.211
HHSIZE	-0.149	0.014	-10.290	0.000	-0.149	-0.035
income2	1.213	0.086	14.052	0.000	1.213	0.049
income3	0.292	0.058	5.035	0.000	0.292	0.019
income4	0.216	0.046	4.703	0.000	0.216	0.020
income5	0.484	0.050	9.588	0.000	0.484	0.040
income6	0.956	0.073	13.096	0.000	0.956	0.048

homeown2	0.999	0.041	24.139	0.000	0.999	0.085
homeown3	0.217	0.201	1.081	0.280	0.217	0.003
educ2	-0.063	0.113	-0.554	0.580	-0.063	-0.002
educ3	0.015	0.053	0.291	0.771	0.015	0.001
educ4	-0.007	0.037	-0.184	0.854	-0.007	-0.001
Age2	-0.011	0.112	-0.094	0.925	-0.011	-0.000
Age3	0.052	0.055	0.949	0.343	0.052	0.003
Age4	0.001	0.044	0.024	0.981	0.001	0.000
Age5	0.017	0.041	0.412	0.681	0.017	0.002
Race2	1.034	0.061	17.030	0.000	1.034	0.056
Race3	0.497	0.080	6.186	0.000	0.497	0.020
Race4	0.408	0.053	7.708	0.000	0.408	0.025
CIMean	0.029	0.001	38.944	0.000	0.029	0.134
CarsPerDriver	-1.123	0.029	-38.668	0.000	-1.123	-0.130
RIDE-HAILING ~						
msa_pt	0.362	0.167	2.172	0.030	0.362	0.008
HHSIZE	-0.011	0.009	-1.202	0.230	-0.011	-0.004
income2	-0.012	0.057	-0.207	0.836	-0.012	-0.001
income3	0.013	0.038	0.344	0.731	0.013	0.001
income4	0.004	0.030	0.118	0.906	0.004	0.001
income5	-0.017	0.033	-0.523	0.601	-0.017	-0.002
income6	0.024	0.048	0.493	0.622	0.024	0.002
homeown2	-0.053	0.027	-1.961	0.050	-0.053	-0.007
homeown3	0.262	0.132	1.989	0.047	0.262	0.007
educ2	-0.190	0.074	-2.563	0.010	-0.190	-0.009

```
  educ3          -0.208   0.035   -5.974   0.000   -0.208   -0.023
  educ4           0.626   0.024   25.885   0.000    0.626    0.100
  Age2           -0.004   0.074   -0.053   0.958   -0.004   -0.000
  Age3           -0.052   0.036   -1.449   0.147   -0.052   -0.005
  Age4            0.014   0.029    0.480   0.631    0.014    0.002
  Age5           -0.025   0.027   -0.931   0.352   -0.025   -0.004
  Race2           0.070   0.040    1.760   0.078    0.070    0.006
  Race3           0.114   0.053    2.161   0.031    0.114    0.007
  Race4          -0.001   0.035   -0.023   0.982   -0.001   -0.000
  CIMean          0.000   0.000    0.523   0.601    0.000    0.002
  CarsPerDriver   0.005   0.019    0.273   0.785    0.005    0.001
  OnlineActivits  0.047   0.030    1.555   0.120    0.019    0.006

Covariances:
              Estimate  Std.Err  z-value  P(>|z|)  Std.lv  Std.all
 .PTUSE ~~
  .RIDE-HAILING     -0.021   0.050   -0.419   0.675   -0.021   -0.001

Variances:
              Estimate  Std.Err  z-value  P(>|z|)  Std.lv  Std.all
  WEBUSE17        0.277   0.004   76.534   0.000    0.277   0.616
  DELIVER        20.291   0.101  199.938   0.000   20.291   0.937
  SPHONE          0.485   0.018   27.668   0.000    0.485   0.352
  CarsPerDriver   0.307   0.001  207.267   0.000    0.307   0.889
  PTUSE          22.258   0.107  207.267   0.000   22.258   0.859
```

RIDE-HAILING	9.550	0.046	207.264	0.000	9.550	0.987
OnlineActivits	0.173	0.004	45.797	0.000	1.000	1.000

R-Square:

	Estimate
WEBUSE17	0.384
DELIVER	0.063
SPHONE	0.648
CarsPerDriver	0.111
PTUSE	0.141
RIDE-HAILING	0.013