IS THERE A LIMIT TO ADOPTION OF DYNAMIC RIDESHARING SYSTEMS?
EVIDENCE FROM ANALYSIS OF UBER DEMAND DATA FROM NEW YORK CITY

Raymond Gerte (Corresponding Author)
Department of Civil and Environmental Engineering
University of Connecticut
261 Glenbrook Road Unit 3037, Storrs, CT 06269
Tel: 203-258-3545 Email: raymond.gerte@uconn.edu

Karthik C. Konduri
Department of Civil and Environmental Engineering
University of Connecticut
261 Glenbrook Road Unit 3037, Storrs, CT 06269
Tel: 860-486-2733 Email: karthik.konduri@uconn.edu

Naveen Eluru
Department of Civil, Environmental and Construction Engineering
University of Central Florida
12800 Pegasus Drive, Room 301D, Orlando, FL 32816, USA
Tel: 407-823-4815 Email: naveen.eluru@ucf.edu

Presented for Presentation and Publication to
ABJ70: Committee on Artificial Intelligence and Advanced Computing Applications
Submitted on August 1, 2017

Revised word count: 7,157 words + 2 Figures/Tables = 7,657 words
ABSTRACT
Recent technological advances have paved the way for new mobility alternatives within established transportation networks, including on-demand ride hailing/sharing (e.g. Uber, Lyft) and citywide bike sharing systems. One aspect that is common across these innovative modes is a lack of direct ownership by the user; in each of these mobility offerings, a resource not owned by the end users’ is shared for fulfilling travel needs. This concept has flourished and is being hailed as a potential option for autonomous vehicle operation moving forward. However, substantial investigation into how shared mode offerings impact travel behaviors and integrate into existing transportation networks is lacking. This paper explores if the growth in the adoption and usage of these modes is unbounded, or if there is a limit to the adoption of shared mode offerings. Recent trends and shifts in Uber demand usage from New York City are investigated to explore the hypothesis. Using publicly available data about Uber trips, temporal trends in the weekly demand for Uber are explored on the Borough of Manhattan. A panel based random effects model accounting for both heteroscedasticity and autocorrelation effects is estimated wherein weekly demand is expressed as a function of a variety of demographic, land use, and environmental factors. It is observed that demand appeared to increase initially after the introduction of Uber, but it seems to have stagnated and weaned over time in heavily residential portions of the island, contradicting the observed macroscopic unbounded growth. The implications extend beyond already existing fully shared systems, but also affect the planning of future mobility offerings.

Keywords: Dynamic Ridesharing, Uber, Shared Systems, Open-Source Data, Linear Panel Model
1. INTRODUCTION

Technology has long been the driver for advances in our everyday means of transport. The development of data-enabled and location aware mobile phones, combined with adoption of sharing economies (also referred to as collaborative economies or peer-to-peer networks), has opened the door for novel modal offerings like Uber, Lyft, and citywide bike sharing systems. These specific modal offerings are referred to generally as dynamic ridesharing or ridesourcing systems. They capitalize on the growth of mobile phone adoption and integrated global positioning (GPS) capabilities to actively link riders to drivers in an on-demand setting. Historically, ridesharing has filled unique niches in the transportation landscape (carpooling, preplanned trips, etc.) and has consistently been encouraged for its ability to minimize user costs and reduce congestion. However, the widespread adoption of conventional ridesharing has been linked with an inability to capture impromptu trips (1). Due to the fact that almost all trips need to be prearranged, conventional ridesharing has substantial limitations and users that recognize its benefits are unwilling to make the switch because of inconvenience. With the growth of dynamic ridesharing, the ability to travel on demand in a manner similar to taxi trips opens a wealth of mobility options. In this paper, the temporal trends in one ridesharing offering, namely Uber, are explored.

As the current industry leader, Uber’s growth is readily apparent. At the close of 2014, Uber operated in 230 cities worldwide, and by June 2015, Uber grew to operate in 300 cities (2, 3). In only six years of existence (2009 to 2015), Uber reported facilitating over one billion trips. As of June 2016, Uber reported surpassing two billion total trips in 450 cities (4). The popularity of Uber can be associated with its inherent flexibility and user-perceived efficiency of use. When using the Uber app, users are able to hail a ride remotely from anywhere within the service city, given that a driver is available. This is an improvement over the call ahead or on street hailing methods of its most similar competitor, the taxi, while also offering service in locations that may not have many viable alternatives. In a major metropolitan hub like New York City, the growth of Uber is substantial. From June 2014 to June 2015, Uber’s Manhattan monthly June pickups increased by over 250% with over one million more trips made using the dynamic ridesharing app (5).

While Uber’s growth is noteworthy, there has also been substantial controversy in regards to the ridesharing giant. Established taxi groups and regulators have filed numerous lawsuits against Uber and other dynamic ridesharing corporations for infringing on market share without the burden of for hire vehicle (FHV) licensure and regulation (6). There has also been considerable debate on the ethics of the dynamic pricing incorporated into Uber’s app (7). Furthermore, there is ongoing discussion regarding Uber’s safety, its policies concerning drivers, and how these policies compare to similar regulations governing other FHV companies (8). The growing impact of dynamic ridesharing is becoming an important piece in the planning of cities and regions around the U.S. and elsewhere. These systems are not only expanding and growing in the short term, but also have substantial implications for the future, namely in regards to the deployment of autonomous vehicles (AVs). AV offerings are projected to perform in a shared, on-demand manner, similar to ridesharing offerings. Also, shared AVs have been identified as one of the key pieces in realizing the complete suite of AV benefits (9).

As such, understanding ridesharing can provide insight into how these dynamic systems are currently performing and being received by the public, while also offering a possible window into the impact of future modal offerings employing the same concept. Substantial investigation into how dynamic ridesharing impacts travel behaviors and integrates into existing transportation
networks is lacking. To address these limitations, this study focuses on quantifying and 
understanding temporal trends in Uber demand within a major metropolitan center. In particular, 
the study explores if growth in the adoption and usage of these modes is unbounded or if there is 
4 a limit to the adoption of dynamic ridesharing. In order to pursue this research, trip pickup data 
made available by Uber through a Freedom of Information Law (FOIL) request by the analytics 
group FiveThirtyEight, is utilized. Weekly demand, aggregated spatially to a taxi zone level, is 
analyzed using a panel based random effects model while also accounting for potential 
heteroscedasticity and autocorrelation issues. The demand is expressed as a function of various 
demographic, land use, and environmental factors to further explore the trends.

The remainder of this paper is organized as follows: Section 2 summarizes the existing 
research and positions the current study, Section 3 describes the data and presents the approach 
to data preparation, Section 4 includes a descriptive analysis of the data, Section 5 outlines the 
model used, Section 6 discusses the results, and finally, Section 7 provides concluding remarks 
and offers some directions for future research.

2. EXISTING RESEARCH

The existing research on dynamic ridesharing demand, and analysis of demand trends in 
particular, is quite limited. These limitations can be attributed to the lack of readily available data 
from the providers of these modal offerings, and limitations in data that is publicly available.

The first vein of research helps to gather a more complete understanding of dynamic 
ridesharing system performance in relation to available supply and demand. Chen et al. and Guo 
et al. (7, 10, 11) explore the characteristics of Uber’s surge pricing and assess its impact on 
supply and demand. The authors collected four weeks of comprehensive data on all Uber vehicle 
types using the Uber App API and Client App API. The data gathered included information on 
the supply of vehicles, the demand for vehicles, and the associated surge price multiplier in both 
downtown San Francisco and Midtown Manhattan. They found that the majority of vehicles in 
their study areas were concentrated in commercial and tourist areas within each study zone. 
Further, they observed that surge pricing did in fact negatively impact demand while 
encouraging supply at the study scale. Correa et al. (12) used the same data to quantify changes 
in Uber and taxi demand over space and time and found substantial significant growth in demand 
over time with high correlation between taxi and Uber modes in the core of the city. Increases in 
roadway length, income, and job opportunities, as well as decreases in transit access time and 
vehicle ownership, were observed with increased demand for Uber.

Another major line of research is the comparison between dynamic ridesharing and 
traditional mode offerings such as taxi and transit. Rayle et al. (13) found that dynamic 
ridesharing users are typically younger and more educated than those that use taxis. Also, while 
ridesharing does remove potential trips from taxis, there is substantial evidence that users opted 
to switch from available transit modes citing travel time as the reason. Survey respondents using 
dynamic ridesharing were skewed towards no vehicle ownership with multiple users traveling 
together. Researchers Cramer and Krueger (14) compiled a direct comparison of Uber’s 
utilization rate (percentage of time drivers are serving passengers) against that of taxis in five 
major U.S. cities. Their results showed that as a whole, Uber is the more efficient mode of 
transportation compared to taxis. Poulsen et al. (15) explored the relationship between Uber and 
green cabs in New York City and found that while demand for both modes is growing, Uber’s is 
growing at a much faster rate and green cabs perform better in poor neighborhoods. These 
studies highlight the fact that Uber is in fact similar to existing taxi infrastructure, however, they
highlight key differences that warrant further investigation into how people are using dynamic ridesharing on a citywide scale.

The final set of literature pertaining to dynamic ridesharing and demand is informal in the form of blog posts and news articles. These are mostly based on Uber demand information for all of New York City made openly available by researchers at FiveThirtyEight; the same data also used in this paper. FiveThirtyEight researchers have published numerous articles on the topic by analyzing the data. Fischer-Baum and Bialik (16) address how Uber’s growth should be classified. They found that in the urban core of Manhattan, Uber trips seem to come almost entirely from taxis, with the sum of total trips across both modes remaining the same across both study years. Schneider (5) explored the trends in Uber demand and found that Uber pickups in the urban core are growing, and notes that these pickups increase more substantially in times of inclement weather compared to increases in taxi pickups.

The research highlighted above provides interesting insights into the demand for dynamic ridesharing, but there are substantial gaps. One gap this work aims to fill is if growth in demand for these systems is subject to saturation, or if it is unbounded. To this end, the study explores demand over time, while also capturing the impact of demographic, built environment, and environmental variables. In order to explore this idea quantitatively, this study develops a panel-based random effects model that is similar in nature to the works by Faghih-Imani et al. (17). The authors study the relationship between NYC Citi Bike ridership, land use, and urban form amongst other studies focused on Citi Bike and Taxi travel time. Their research focuses on three main types of variables: those related to the environment, temporal variables, and land use variables. A similar structure is also observed in an analysis of bus ridership with the incorporation of real-time location information by Brakewood et al. (18), and work by Campbell and Brakewood (19). These studies, along with some of the variables described in the work above, served to motivate the modeling aspect of this research and helped guide selection of potential demographic, land use, and environmental factors that may influence Uber demand.

3. DATA
As noted in the limitations regarding dynamic ridesharing research, restrictions in data availability and quality have had a substantial impact on the volume of research. Recent adoption of open-data initiatives (wherein data is made publicly available) and the practice of making data available in response to public requests (through FOIL and other similar mechanisms) have made much of this work possible. The Uber trip information was made available by Uber after a FOIL request and was gathered from FiveThirtyEight’s GitHub repository (20). While these data are the only source of information on Uber demand, they also have several limitations that dictated the data preparation and approach to analysis. First, Uber data is only available for six months in 2014 (April through September) and six months in 2015 (January through June). Within this, trip pickup location is recorded at a latitude/longitude level in 2014 and aggregated up to predefined taxi zone level in 2015, including pickup date and time information. Due to the spatially aggregated nature of 2015 data, taxi zone was chosen as the spatial unit of analysis. Subsequently, all remaining data (demographic, land use, and environmental factors) were spatially aggregated to the taxi zone level. Taxi zone delineations are defined by the NYC Taxi and Limousine Commission and line up with neighborhood boundaries in New York City. This analysis focused only on the taxi zones in the Borough of Manhattan. There are 69 unique taxi zones in Manhattan which includes three zones for ferries to Ellis Island, Liberty Island, and Governors Island, and four zones for major parks, namely, Central Park, Battery Park, Inwood
Hill Park, and Highbridge Park. To add context, there are 288 census tracts located in the same
study area. The boundaries of census tracts and taxi zones are coincident. However, taxi zones
are relatively larger in spatial extent compared to tracts and on average each taxi zone is
comprised of about four census tracts. Built environment variables were collected at the tax lot
level and gathered from the NYC Planning Department’s website (27). Information on transit
stations/stops and bike share infrastructure was accessed through NYC open data hub (22).
Environmental variables like temperature and precipitation were gathered at the day level from
2014 through 2015 in Central Park from the National Climate Data Center hosted by the
National Oceanic and Atmospheric Administration (23). All demographic data was collected at
the census tract level from the American Community Survey’s five year estimates for 2014 and
2015 and aggregated to the taxi zone level.
Substantial data processing and data preparation was performed prior to the analysis.
First, the taxi zone aggregated Uber pickups were summed to the week level because of the high
level of variability at the day level. From the 12 months of initial data, there were 54 total weeks,
however, weeks that fell at the beginning and end of the dataset were incomplete. As a result,
these incomplete weeks were dropped leaving 49 total complete weeks of demand data. For the
remainder of the paper, the weekly pickup counts will be used to constitute the demand and will
be referred to as such. Daily weather information was treated similarly. Built environment
variables were assumed to not vary across the investigation period and the demographics of each
taxi zone were assumed to be represented by the 2014 and 2015 ACS estimates, respectively.

4. DESCRIPTIVE ANALYSIS
The growth in Uber overall has been well documented, in both the literature presented above and
in its emergent prevalence in the societal lexicon. When analyzing the data at the borough level,
substantial growth in Uber demand over time is easily observed. The total number of trips made
in Manhattan from April 2014 through September 2014 was just under 140 thousand per week.
In 2015, that number rose to just over 430 thousand trips per week from January through June.
However, it is important to note that this growth is not evenly distributed. At the taxi zone level,
the top five taxi zones in both 2015 and 2016 are as follows: Midtown Center, Union Square,
Tribeca/Civic Center, and both East and West Village. All of these zones are located either in
midtown or lower Manhattan and together amount for about 20 percent of the total trips in 2014
and 2015, respectively. In reviewing the shifts in demand by each taxi zone individually,
differences were noted in zones that were more residential. These differences were typically
manifested as lower overall demand each week, but the growth of that demand appeared lower
than that in commercial zones. This observation is reasonable because residential areas typically
have less tourism or visitor traffic.
To better understand overall demand and the temporal trends, the taxi zones were
classified into five groups based on a residential built environment measure derived from the
land use dataset. The first group, residential taxi zones, are those zones that have a higher ratio of
residential land use than commercial land use. The second group, commercial zones, are those
that have a higher ratio of commercial land use than residential. These two groups were then
further divided into those found in Northern Manhattan (north of 86th street on the east side of
the island and north of Central Park) and those not. Taxi zones that were exclusively parks were
treated as the fifth group. Figure 1 highlights these trends with respect to four of the land use
areas classified above, excluding parks. From the map inset, it is shown that residential zones are
surrounding Central Park and the northern portion of the island, as well as in the south east. The
line shading for the demand plots range from dark (less residential/commercial) to light (heavily residential/commercial). Additionally, residential and commercial dominant zones are shaded in blue and orange hues respectively. Therefore, more mixed-use areas i.e. places with an even split of residential and commercial land, are represented by the darker hues. Weekly trends show that there is general increase in demand for all zones from 2014 to 2015. Looking only at the commercial dominant taxi zones, a few notable trends are apparent. First, a banding effect can be observed wherein adjacent zones experience similar levels of demand. The upper bands with the most demand correspond to Midtown and surrounding taxi zones, while the remaining zones are distributed further downtown. Almost all commercial zones located in central and southern Manhattan have an upward trend even after noting seasonal decreases. Only one commercial dominant taxi zone is found in the northern part of the island and that is Morningside Heights. Home to Colombia University, this zone exhibits atypical behavior when compared to the other commercial and residential zones. Interestingly, the substantial increases and decreases observed in 2015 line up with typical semester beginning and ending points.

On the other hand, focusing in on the residential areas in central and lower Manhattan, darker lines indicate the lesser residential of the residential areas. These zones have the largest increase in demand of the residential zones and have higher volatility and clear seasonal shifts. The lighter color lines indicate heavily residential zones and have fewer weekly trips and less overall variability at the week level. Perhaps the most interesting observed trend within these zones is that demand appears to stagnate over time in the heavily residential areas. The results are similar in the residential zones in northern Manhattan. The more mixed zones have the highest demand and growth while the most residential of the group fall within the middle of the graph. Here too, the lines below the 1000 trip mark appear to decrease. The observations of flattening demand lend credence to the hypothesis that growth in Uber demand may not be unbounded and that it may be subject to a limit. We further explore this hypothesis using a random effects panel model below.
The demand lines are shaded by the ratio of residential floor area to total floor area for the residential dominant zones, and by ratio of commercial floor area to total floor area for the commercial dominant zones. Darker lines indicate a less intense relationship, while lighter lines indicate a zone is almost entirely residential or commercial, respectively.
5. PANEL MODEL
The nature of the weekly Uber demand data (i.e. panel) lends itself to a linear panel based analysis (24). The demand data are available temporally by week for 25 complete weeks in 2014, and 24 complete weeks in 2015. However, not all elements of the demographic, land use, and environmental data used to explain the demand are time variant. As a result, a random effects structure was employed that allows consistent estimation of coefficients associated with time invariant variables.

Following the exposition of linear panel models in Cameron and Trivedi (24), an outline of the model structure and assumptions are presented below. Let \( z = 1, 2, ..., Z \) represent an index for the unique taxi zones and \( t = 1, 2, ..., T \) represent an index for the week. The dependent variable, weekly pickup demand, is then modeled using a linear panel structure similar to the one below:

\[
y_{zt} = \alpha_z + \beta'X_{zt} + \varepsilon_{zt}
\]

Where \( y_{zt} \) represents the weekly pickup demand in taxi zone \( z \) at week \( t \), \( \alpha_z \) captures the zone specific unobserved heterogeneity, \( X_{zt} \) represents the vector of independent variables and time variables, and \( \beta \) represents the regression coefficients associated with the independent variables, \( \varepsilon_{zt} \) represents the errors within each taxi zone \( z \) for time period \( t \). \( \varepsilon_{zt} \) is assumed to be strictly exogenous and normally distributed as \( N(0, \sigma_{\varepsilon}^2) \). Further, specific assumptions on \( \alpha_z \) result in the fixed effect and random effect variants. In a random effects model, used in the analysis, \( \alpha_z \) is assumed to be random and distributed normally as \( N(0, \sigma_{\alpha}^2) \). Also, \( \varepsilon_{zt} \) and \( \alpha_z \) are assumed to be independent. Further, heteroscedasticity and autocorrelation are allowed in specification of \( \varepsilon_{zt} \).

The model was estimated using Stata software using the feasible generalized least squared routine (i.e. xtgls). In order to test the autocorrelation and heteroscedasticity specification of the error term, separate hypothesis tests were run. First, to test heteroscedasticity, a likelihood ratio test proposed by Poi and Wiggins (25) was used that compares heteroskedastic treated panel model against the same model assumed to be homoscedastic. Second, to test autocorrelation, the procedure outlined in Wooldridge (26) was adopted.

6. DISCUSSION OF RESULTS
As noted earlier, a linear panel model was estimated using the generalized least squares approach in Stata software. The model allowed heteroscedasticity and autocorrelation in the specification of the disturbance term. Due to the presence of autocorrelation and gaps in the data (i.e. missing weekly data between 2014 and 2015), standard procedure of maximum likelihood estimation was not applicable. The final panel contained 61 taxi zones over 49 weeks for a total of 2,989 observations. Parks and zones with missing independent variables were dropped from the model, including Ellis Island, Liberty Island, Governors Island, Randall’s Island, Central Park, Battery Park, Inwood Hill Park, and Highbridge Park. Before running the panel model, the tests for heteroscedasticity and autocorrelation identified above were performed. The results indicated that both autocorrelation (with an F-test statistic of 331 with degrees of freedom of 1 and 60 and p-value of 0.00), and heteroscedasticity (with a chi squared statistic of 2242.47 with degrees of freedom of 60 a p-value of 0.00) were present and as such the error term specification needed to accommodate these issues. In the interest of space, the complete listing of autocorrelation and heteroscedasticity parameters is not presented here.

A variety of explanatory factors, including demographic, land use, and environmental, were explored. Additionally, temporal variables were explored to capture the impact of
unobserved time varying factors. The final model is presented in Table 1. All the variables included are significant at least at the 10 percent level of significance. Below, the influence of the various explanatory variables on the weekly Uber pickup demand are discussed.

**Table 1. GLS Random Effects Panel Model Results**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Random Effects Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
</tr>
<tr>
<td>Constant</td>
<td>-19056.180</td>
</tr>
<tr>
<td><strong>Time Variables</strong></td>
<td></td>
</tr>
<tr>
<td>Summer Dummy Variable</td>
<td>-278.133</td>
</tr>
<tr>
<td>Winter Dummy Variable</td>
<td>89.647</td>
</tr>
<tr>
<td>Difference in weeks from initial week</td>
<td>93.323</td>
</tr>
<tr>
<td><strong>Environmental Variables</strong></td>
<td></td>
</tr>
<tr>
<td>Weekly Precipitation (in.)</td>
<td>25.892</td>
</tr>
<tr>
<td>Weekly Average Max Temperature (°F)</td>
<td>-10.143</td>
</tr>
<tr>
<td><strong>Built Environment Variables (Taxi Zone Level)</strong></td>
<td></td>
</tr>
<tr>
<td>Count of Citi Bike Stations</td>
<td>242.850</td>
</tr>
<tr>
<td>Total Built Area (10,000 sq. ft.)</td>
<td>0.936</td>
</tr>
<tr>
<td>Percent Residential by Floor Area</td>
<td>64.247</td>
</tr>
<tr>
<td>Percent Retail by Floor Area</td>
<td>151.426</td>
</tr>
<tr>
<td><strong>Demographic Variables (Taxi Zone Level)</strong></td>
<td></td>
</tr>
<tr>
<td>Percent African American</td>
<td>-58.925</td>
</tr>
<tr>
<td>Percent Male</td>
<td>60.847</td>
</tr>
<tr>
<td>Percent Age Under 19</td>
<td>220.657</td>
</tr>
<tr>
<td>Percent Age Over 65</td>
<td>-0.213</td>
</tr>
<tr>
<td>Percent of Households with No Vehicle</td>
<td>126.400</td>
</tr>
<tr>
<td>Percent of Households with One Vehicle</td>
<td>131.103</td>
</tr>
<tr>
<td>Average Household Size</td>
<td>-2991.058</td>
</tr>
<tr>
<td>Percent of Adults with at least Some College</td>
<td>17.922</td>
</tr>
<tr>
<td><strong>Interaction Variables</strong></td>
<td></td>
</tr>
<tr>
<td>(Percent Residential by Floor Area) x (Difference in weeks)</td>
<td>-0.909</td>
</tr>
<tr>
<td>(Percent Retail by Floor Area) x (Difference in weeks)</td>
<td>9.581</td>
</tr>
</tbody>
</table>

(* p < 0.10, ** p < 0.05, *** p < 0.01)

6.1 Influence of Temporal Variables

A variety of temporal variables were explored to capture the influence of unobserved time varying and seasonal factors. As the scope of the data was limited to only two six month periods, the extent of seasonal coverage was also limited to only include winter, spring, and summer. In the model, seasonal effects are reported with reference to spring. Spring was chosen because it was the only season present in both years of the data used in the analysis. It appears that Uber demand at the taxi zone level decreases during the summer as noted by the negative coefficient.
This observation is also confirmed by the descriptive analysis wherein a clear drop was observed during the summer months of 2014. This may be attributed to individuals’ preference to use non-motorized modes, such as walking or biking, when the weather is warmer compared to other seasons. While reviewing research on seasonal influences on motorized and non-motorized travel, Guo et al. (27) found that during the summer, people make more discretionary trips via non-motorized modes. On the other hand, an opposite reasoning can be offered to explain the positive effect of the winter dummy variable on Uber demand: given the extreme weather that New York City experiences during the winter months, people may seek the comfort of a climate controlled vehicle in order to escape cold conditions.

Additionally, the difference in week from initial week variable is defined as the number of weeks past the first complete week (i.e. week ending on April 6th, 2014) of demand data. Here, a positive influence is apparent and can be interpreted to mean that as time progresses, the pickup demand within taxi zones is increasing. This finding is in line with both the observed trends at the descriptive level, as well as research and reports presented earlier on the growth of Uber overall.

One of the primary hypotheses was to explore potential tapering off of Uber demand with time. To this end, the impact of higher order versions of the difference in week variable were explored after controlling for the other explanatory variables. For example, a negative coefficient associated with a second order influence will lend evidence in support of the hypothesis. It was found that higher order versions of the difference in week variable were not significant. However, this contradicts the observations from the trend analysis by taxi zones that showed that there is evidence of a tapering effect for residential zones. To further explore the tapering off effect, interaction variables between difference in week and other explanatory factors, namely demographic and land use variables were explored. The results for the interaction variables are discussed at the end of this section.

6.2 Influence of Environmental Variables
The environmental variables are included to extend interpretation beyond just seasonality and quantify how weekly weather trends influence zonal level Uber demand. The positive influence related to total weekly precipitation was expected and is intuitive. As the overall total of weekly precipitation increases, demand for Uber also increases. This is plausible because rain/sleet/snow move people away from other non-automobile modes and towards trip making in the comfort of an on-demand vehicle. A substantial amount of users in Manhattan use non-automobile modes like Citi Bike, walk, and transit among others (17, 19). These riders may potentially opt for an Uber trip if it means avoiding exposure to the elements. As the average high temperature by week increases the demand for Uber also increases. Increased temperatures, especially those outside of the normal weather condition, may be prompting individuals to trend towards non-motorized travel in a similar manner to the summer influence.

6.3 Influence of Built Environment Variables
Unfortunately, in the final model formulation, the influence of number of bus stops and subway stops at the taxi zone level was found to be statistically insignificant. However, the large positive influence of Citi Bike stations at the taxi zone level is an interesting result. At a high level, the positive influence makes sense as Citi Bike stations are primarily located in lower Manhattan and these taxi zones correspond to the areas with the highest trip attraction. However, upon closer inspection, it can be seen that even with the inclusion of other built factors like total floor area...
which are more direct determinants of trip attraction, Citi Bike stations have a significant and
sizeable impact. One plausible reason for this is that the users of the on-demand shared Citi Bike
system may also be inclined to use shared on-demand services provided by Uber. It is also
plausible that Uber is used by such individuals to complete trips that Citi Bike could not. This is
consistent with research by Faghih-Imani et al. (17) – in their analysis of Citi Bike and Taxi
modes, the authors found that for trips with substantial distance taxi modes proved to be more
suitable option – a similar scenario could be at work here.

The total built area is assumed to capture the overall intensity of each taxi zone, and the
positive relation to Uber demand follows from our expected results. With more built area, there
is more to see, do, and places to live and this falls within other findings that Uber demand is
highest in the core of Manhattan and in tourism driven taxi zones. The percent of residential
floor area has a positive influence on Uber pickup demand. Residents, just like visitors and
tourist, make trips and there are substantial amounts of demand in these residential zones –
referring to Figure 1, it can be seen that residential zones like East and West village have some
of the highest demand of all observed taxi zones.

Variables representing other land use, namely, percentages of overall commercial, retail,
garage (parking), and office areas were explored. These remaining land use percentages were
also included in the model but the only other type that proved significant was the percentage of
retail by floor area. This strong positive interaction shines light on previously noted trends
regarding Uber’s propensity to serve tourism and central business districts. Within Manhattan,
these locations coincide with higher proportions of retail area and as such higher Uber demand.

6.4. Influence of Demographic Variables
Understanding how demographic makeup of taxi zones influence Uber demand can provide
valuable insights into how different groups are adopting dynamic ridesharing services. When
exploring the link between Uber demand and percentage of the population that is African
American in a taxi zone, a negative relationship was observed. This observation raises questions
on equity. A better understanding of the factors influencing (or lack thereof) African American
users to Uber is needed.

The second demographic factor explored is that of gender on Uber demand. Here the
percentage of males within the population was found to have a positive relationship with Uber
pickup demand. A possible explanation for this can be safety. As mentioned previously (8), Uber
has been dealing with issues regarding safety. As a result, male riders may feel more comfortable
using the dynamic ridesharing system than their female counterparts.

Age has been an important topic in regard to Uber adoption, and previous research
indicates that Uber users tend to skew younger (13). When holding the typical adult age range as
the constant categorical factor and exploring the effects on individuals under 19 and over 65, our
findings fall in line with existing research. The percentage of the population under 19 years of
age is positively associated with demand for Uber than their older counter parts. On the other
hand, senior individuals exhibit a lower demand for Uber. Propensity for technology and existing
mode bias are potential explanations for this relationship.

Vehicle ownership was investigated over conventional income variables due to findings
in the Uber literature. In San Francisco, survey respondents that utilized Uber identified as
owning few vehicles (13). As such, we investigated the influence of household vehicle
ownership on Uber demand. It was observed that both no vehicle and one vehicle households
exhibit heightened Uber demand compared to households with two or more vehicles. Overall
vehicle deficiency can explain this, as households still need to make trips to fulfill their activity needs.

The modeled effects of household size in regard to demand indicate that larger households have a substantially lower demand for Uber than smaller ones. As household size increases, traveling in general typically becomes more difficult and potentially also expensive especially for a mode like Uber that has limited capacity per vehicle.

The final significant demographic variable explored was aimed at quantifying the relationship of education on Uber demand. While only marginally significant, the findings indicate that individuals with higher education exhibit higher demand for Uber. Once again in line with preliminary survey findings, this can be can be attributed to a higher propensity for adopting and using technology by those associated with higher levels of education.

6.5. Influence of Interaction Variables
As noted earlier, the main hypothesis was to explore if there is a potential tailing off effect in the growth of Uber demand with time. As noted before, higher order versions of the weekly difference variable were found to be insignificant despite the evidence from the descriptive analysis. In an effort to further explore the tailing off effect, interaction variables between the weekly difference variable and different demographic, and land use variables were investigated to see if there are any potential negative shifts in the demand for specific tax zones and/or population groups over time. This exploration provided substantial evidence in support of our proposed hypothesis. We found that the interaction between percent retail and weekly difference had a substantial positive impact on Uber demand, suggesting that Uber demand in high retail areas is growing with time. As Uber becomes more readily adopted, a greater proportion of users are making trips from and to areas with higher retail using the service. However, when looking at the interaction of the percentage of residential area by weekly difference there is a clear negative impact. While the magnitude of the relationship is small, the negative coefficient indicates that residential areas over time have a waning or decline in overall Uber demand. There are a couple of possible reasons for this observation. First, some residents may have preference towards existing modes (and bias against Uber) and as such are not going to adopt new systems. Second, if Uber only replace SOV trips, and/or borrows from other non-automobile modes, it will not improve congestion on roadways. As a result, those who are conscious to travel time performance and are price conscious may prefer the cheaper alternative with similar travel time performance (e.g. subway, bus).

7. CONCLUSION AND FUTURE RESEARCH
With the substantial growth in dynamic ride sharing offerings over the last few years, a comprehensive investigation into understanding the underlying factors impacting demand is of interest. Using multiyear data from an existing ride sharing service, namely, Uber from New York City, the temporal trends in demand for this service are explored. Existing research supports the growing adoption of Uber, but here it is hypothesized that while macroscopic trends may suggest unbounded increases in the demand, meso/microscopic levels may offer a different picture. It is posited that the demand for ride sharing services has a limit. To explore these underlying factors and test our hypothesis, a random effects panel model was developed.

The findings of our analysis indicate that there is a significant growth over time, but that seasonality in demand exists with an increase in demand in winter months and decrease in demand in summer months. Weather was also noted to have a statistically significant impact,
with increased precipitation positively influencing demand and warmer temperatures decreasing the weekly demand. At a built environment level, the presence of transit was not determined to have significant impact, but bike share infrastructure had a positive influence. Similarly, the more built area a taxi zone had, as well as the higher percentage of the floor area that is residential and retail based land use, indicated a higher demand for Uber service. The influence of zonal level demographics was also investigated. The percentage of population that is male, under 19 years of age, has some college education, and has one or no household vehicles showed positive impacts on weekly Uber demand. The percentage of the population that is African American, over 65, and larger households all had negative impacts on Uber demand. Within the model, the weekly difference had a positive effect on Uber demand, it was included at higher order to capture the observed tailing over time, however the results were insignificant. Interaction variables were then included to test our hypothesis for observed declines in demand over time with respect to various built environment variables. A positive influence of retail area across time was determined. Perhaps the most substantial finding, a negative relationship was observed when exploring the influence of percent residential area across time. This negative relationship lines up with the decreasing trends observed in the descriptive analysis. Together, these results offer substantial evidence in support of a trailing in Uber demand over time in residential dominant areas.

These last findings have substantial impact beyond the scope of just Uber demand analysis and have potentially important policy implications for the overall adoption of dynamic shared modes. The significant tapering off within residential zones indicates that even as the demand grows at a citywide scale, residents’ inertial modal biases and familiarity with existing modes may limit the dynamic shared ride adoption by all individuals within individual subareas (28). A second policy implication that can be drawn here is that users are typically motivated to change modes if the alternative provides improvement over existing offerings (e.g. lower costs, or shorter travel times). In the current operational framework, Uber and other dynamic shared ride providers most often replace single occupancy (SOV) auto-modes further contributing to the effects of congestion and hindering overall adoption (13, 16). Further investigation into these effects is warranted.

This research presents a substantial first step in understanding dynamic ridesharing in a real world context, but it is important to recognize some of the limitations of this work. First, holding Uber pickups as a stand in for the demand can be problematic as it inherently omits any demand by users who attempted to request rides but were not able to have those requests filled. It also inherently omits any user without a smartphone or knowledge of the Uber interface, which may be disproportionately represented by elderly and low income groups. Second, with data only on the pickup event of each trip, a rich understanding of how these trips are integrating into the existing network is not possible. Third, the spatially aggregated nature of the data limited this analysis to the neighborhood level offered by taxi zones. The data quality and structure issues noted with the Uber data motivates an important discussion. As open data becomes more readily available, it is vital for open data providers to adopt a standard of practice for quality and structure that can in turn promote additional research endeavors.

Additionally, the panel model used in this research assumes strong exogeneity which may not be supported by the data and alternate model forms must be explored. A comprehensive spatial analysis will help to further explore and characterize the spatial relationships to Uber demand. Also, to supplement and confirm the findings presented here, we anticipate further exploration through a primary survey which will collect disaggregate level information on users.
and their preferences. Lastly, a comprehensive study of the temporal trends in all modes including taxis, subways, buses, and other dynamic rideshare providers from the region can help understand the tradeoffs between on-demand modes, like Uber, and existing shared ride services, and will be explored in future work.
REFERENCES


