Center for Advanced Multimodal Mobility Solutions and Education

Project ID: 2021 Project 05

Characteristics of Pooled Trips Offered by Ride-sourcing Services in Chicago

Final Report

by

Carol Atkinson-Palombo, PhD (ORCID: https://orcid.org/0000-0003-2870-4213)
Professor, Department of Geography
University of Connecticut
215 Glenbrook Road, Storrs, CT, 06269

Norman W. Garrick, PhD (ORCID: https://orcid.org/0000-0003-3295-7520)
Professor Emeritus, Department of Civil & Environmental Engineering
University of Connecticut
261 Glenbrook Road Unit, 3037. Storrs, CT 06269-3037.

for

Center for Advanced Multimodal Mobility Solutions and Education
(CAMMSE @ UNC Charlotte)
The University of North Carolina at Charlotte
9201 University City Blvd
Charlotte, NC 28223

September 2022

ACKNOWLEDGEMENTS

This project was funded by the Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE @ UNC Charlotte), one of the Tier I University Transportation Centers that were selected in this nationwide competition, by the Office of the Assistant Secretary for Research and Technology (OST-R), U.S. Department of Transportation (US DOT), under the FAST Act. The authors are also very grateful for all of the time and effort spent by DOT and industry professionals to provide project information that was critical for the successful completion of this study.

DISCLAIMER

Table of Contents

List of Figures

List of Tables

EXCUTIVE SUMMARY

Ride-sourcing companies or transportation network companies (TNCs) are still in the process of establishing a role in the transportation system in urban areas across the world. The costs and benefits—both intended and unintended—are still being identified not only in academic literature but also by governments trying to regulate them. Two issues are of particular concern: affordability in low-income communities with limited car ownership and transit options and the increase in vehicles miles traveled (VMT) that 'deadheading' from one trip to another generates. TNCs have responded to these concerns by offering *pooled rides* where customers can choose to share a ride with stranger(s). However, the extent to which these services are used across different communities, especially low-income areas, is not well understood.

This project explored the spatial and temporal patterns of *pooled rides* in the City of Chicago and discusses the resultant equity implications for low-income communities in the city. The analysis uses cluster analysis with demographic, transportation-related, and built environment characteristics at the census tract level and derives key trip statistics using data mining. The results show that the reliance on pooling is greater in low-income neighborhoods and that income and proximity to the downtown negatively correlate with the willingness of the rider to share a ride. Additionally, it was found that the willingness to pool is very sensitive to fare changes in pooled rides. These findings lend themselves to a better understanding of the potential equity impacts of relying on pooled ride-sourcing services in low-income communities.

Chapter 1. Introduction

1.1 Problem Statement

Ride-sourcing companies or transportation network companies (TNCs) are still in the process of establishing a role in the transportation system in urban areas across the world. The costs and benefits—both intended and unintended—are still being identified not only in academic literature but also by governments trying to regulate them. Two issues are of particular concern: affordability in low-income communities with limited car ownership and transit options and the increase in vehicles miles traveled (VMT) that 'deadheading' from one trip to another generates. TNCs have responded to these concerns by offering *pooled rides* where customers can choose to share a ride with stranger(s). However, the extent to which these services are used across different communities, especially low-income areas, is not well understood.

1.2 Objectives

The objective of this project is to explore the spatial and temporal patterns of *pooled rides* in the City of Chicago and to discuss the resultant equity implications for low-income communities in the city.

1.3 Expected Contributions

Ride-sourcing companies, also referred to as Transportation Network Companies or TNCs (CPUC, 2021), have become commonplace in some urban localities over the past decade, providing users with an alternative mode of transportation that offers point-to-point coverage . These services have transformed the transportation landscape in many cities, allowing a high level of control and flexibility over a journey, including the type of vehicle provided, pickup time, and location. Ride-sourcing companies are distinct from transit in this regard, yet also differ from personal vehicles in that there is less personal responsibility and maintenance/operating costs faced by the user. This niche, at the intersection of transit, taxis, and private vehicles, has varied applications that could complement other forms of transportation or replace them entirely. Given the potential for TNCs to significantly impact the mobility and accessibility of the general public, and essentially privatize transportation, the equity implications of these services warrant careful consideration.

Evidence has emerged about some of the unintended negative consequences of ride-sourcing services. For instance, according to the Union of Concerned Scientists (Anair et al., 2020) ride-hailing trips increase pollution by about 69% more than the trips they replace, and non-pooled rides produce 47% more emissions than if the same trips were taken in private vehicles. Furthermore, Henao and Marshall (2019) found that ride-hailing adds about 83.5% more vehicle miles traveled (VMT) than if the services did not exist and that deadheading (driving around without passengers) makes up at least 40.8% of ride-hailing trips. Pooled rides, ride-hailing trips where users share part of the ride, are one way to reduce the magnitude of emissions and congestion, particularly if ride-hailing is used as a link to public transportation. Yet, for urban users, ride-sourcing services offer an opportunity to switch from transit and walking/biking

(Circella et al., 2018), and additional VMT has the potential to bring more congestion and emissions. Note that though companies are beginning to move toward electric vehicles, those fleets represent a very small share of those used for ride-sourcing. For example, less than 1% of ride-hailing trips in California in 2018 took place in an EV (Anair et. al., 2020).

Similarly, ride-sourcing companies filling in gaps in public transportation could have inequitable impacts on communities. Shwieterman and Livingston (2018) showed that TNC private rides were $15-16 more expensive than the Chicago Transit Authority fares, and pooled rides were $6-11 more expensive. This raises significant concerns over mode replacement in low-income areas underserved by transit. For instance, Ermagun and Tilahun (2020) found that transit accessibility to jobs is low in Chicago's south-side neighborhoods, which are predominantly Black. People in these neighborhoods may turn to ride-sourcing companies as alternatives to poor transit options but would have to spend more money, potentially generating equity issues. Therefore, understanding how people are using ride-sourcing services is a critical question when considering the equity and sustainability implications of this mobility option.

To better assess the impact of ride-sourcing companies on low-income communities, this project conducts a comprehensive overview of where and how these services, specifically *pooled services*, are being used in the City of Chicago. *Pooling* is a multi-step process that involves both consumer behavior and the ability and willingness of a company to meet the demand. This anlysis assesses trips by ride-sourcing companies that occurred in Chicago in 2019, in relation to spatial and temporal factors, as well as demographic characteristics, through a cluster analysis of similar census tracts. The analysis first addresses the question of what factors may influence people's decision to choose the pooled ride option offered by ride-sourcing companies. Then the analysis examines the effect of a fare increase that began in April 2019 and how this affected the motivation to pool rides.

1.4 Report Overview

The remainder of this report is organized as follows: Chapter 2 presents a comprehensive review of literature on ridesourcing. Chapter 3 contains a detailed description of the data used in the study, followed by Chapter 4 which details the methodology. The Results, and Discussion & Conclusion are then set out in Chapters 5 and 6 respectively.

Chapter 2. Literature Review

This chapter provides a review and synthesis of the state-of-the-art and state-of-the-practice literature on ride-sourcing, focusing specifically on those studies that address equity implications.

2.1 Overview

Most of the reviewed literature on ride-sourcing is very recent, reflecting the newness of this type of service. Yet, some of the attributes of why these services are chosen are associated with traditional ridesharing services that have been in use since the seventies (Ferguson, 1997). The availability of comprehensive data detailing trips by ride-sourcing companies remains limited, and current datasets that are publicly available exclude information such as occupancy and information about the riders. This lack of information poses a limitation to thorough assessment regarding how and why trips are pooled. The existing literature attempts to fill in some of these gaps by assessing the factors that influence the adoption of ridesharing and ride-hailing services (Lavieri & Bhat, 2019); characterizing the current travel patterns displayed by ride-hailing services (Atkinson-Palombo et al., 2019); and developing methodological approaches to understand ride splitting behavior (Chen et al., 2017; Saadi et al., 2017). The following section describes the most important trends and findings in this literature.

TNC supply and demand rest on the attributes and behavior of both passengers and drivers and temporal and spatial characteristics of the trips, which results in different levels of trip quality, drivers' and providers' utility, and ridership. On the passenger side, demand is also influenced by the presence of alternative travel modes and their relative quality and fare. Driver supply is affected by preferences such as when, where, and how long drivers wish to work, and their decision to drive for ride-sourcing companies is weighed against other job options. The benefits increase with a rise in demand level, sharing discount, and scheduling horizon (time before departure that the trip is requested), as shown by Kucharski and Cats (2020), who developed an algorithm to match shared rides in Amsterdam.

In terms of the factors influencing the use of shared ride-hailing services, Kucharski & Cats (2020) found a positive correlation between longer trips and sharing. This aligns with findings from Dong et al. (2018) (who focused on ridesharing service eHitch from Didi in China) and Hou et al. (2020) (focused on ride-hailing pooled services in Chicago), who found that longer, more expensive trips displayed higher rates of willingness to pool, suggesting pooling is used to reduce travel costs. Kucharski & Cats (2020) explained that the correlation between longer trips and sharing occurs because i) a discount applied to shorter trips cannot compensate for the detour that results, but longer trips can have more significant fare reductions, and ii) rides are easier to match with others when they are longer. Alonso-González et al. (2020) investigated the impact of fare discounts, additional travel time, and willingness to share a ride with other passengers on the decision to share a ride. Those researchers found that sharing disutility is less of a deterrent than the time-cost tradeoff (essentially, the cost savings do not outweigh the additional time gained due to pooling) but offering rides that assure only one or two additional passengers will be present can increase pooling requests.

A similar study by Abouelela (2020) studied a commercial pooling service in Mexico called Jetty, in which passengers are matched to nearby routes traveled by vehicles with 3-45 seats. They found that 95% of riders use it to get to work and noted that the likelihood of using this service increases in areas where public transportation is less accessible. The authors also described an array of demographic factors that contribute to the decision to use shared mobility services, with many users being young, full-time employees with high education, income, and car ownership. A quarter of women reported that they use it for security against harassment experienced on other modes, such as public transportation. Those authors also noted that because Jetty is a more direct travel option, it replaced trips that otherwise would have been made using multiple modes of transportation. Similarly, Lavieri & Bhat (2019), who specifically focused on ride-hailing pooled services in the Dallas-Fort Worth Metropolitan Area, found that higher use of ride-hailing services is associated with productive use of time while traveling, higher household income, individuals living alone, urban residents, and commuters using non-car modes.

Various studies have pointed out that the willingness to share varies significantly in space and time. Hou et al. (2020), examining data from Chicago, found that willingness to pool was highest in the midweek and during morning peak hours and lowest on Saturdays and nights, consistent with Abouelela (2020). Additionally, Hou et al. found that trips to and from the airport and zones with high population/job density had a low willingness to pool. Meanwhile, economically disadvantaged areas have higher rates of people willing to pool their rides, with income at the pickup and drop off location being some of the most important explanatory variables. Similarly, Chen et al. (2017), who studied ride-splitting behavior using decision trees in Hangzhou, China, found that travel time along with surge pricing ratio, trip fee, trip distance, pick up time or passenger waiting time are the top 5 attributes that affect the decision to split rides with other passengers. Other studies (Jacob and Roet-Green (2017), Yan et al. (2018)) also found that waiting time (i.e., the time the passenger has to wait after he or she requested the service) can significantly influence pooling outcomes.

Regarding the interaction of ride-hailing services with other transportation modes, Babar and Burtch (2017) examined the effect of ride-hailing on transit demand and noted that ride-hailing is associated with declines in the utilization of city bus services but increases in the use of commuter rail. They identified that population density, unemployment, gas prices, crime, and weather are factors in the operating environment that affect this interaction between modes. Zhang and Zhang (2018) also observed that there is a significant positive relationship between public transit use and ridesharing use, but note that in areas without rail, the ridesharing frequency is lower. The authors also observe that usage patterns differ based on a variety of socio-demographic characteristics.

Finally, in terms of the methods used to investigate resourcing services, in addition to econometric models and linear programming approaches, various studies have used machine learning-based methods to explore the factors influencing ride-hailing demand. For example, Chen et al. (2017) applied an Ensemble Learning approach to understanding ride-splitting willingness using data from Didi in China and found that travel time was the most critical variable (greater weight) in predicting ride-splitting behavior. Saadi et al. (2017) found that demand depends significantly on ride fares and the time of day (demand can exceed supply during peak hours and evenings, as discussed by Xu et al.). Similarly, weather variables had lower significance (but were not negligible), and traffic conditions showed that the least congested local areas have a substantial weight in explaining the demand for ride-hailing services.

2.2 Summary and Gap Analysis

The existing literature provides a base upon which this study aims to expand an understanding of generalized consumer preferences and the factors that contribute to a ride-sourcing company's ability to meet the demand for their pooled services. Studies such as Lavieri & Bhat (2019) explored socio-economic aspects of ride-pooling. Still, their sample is limited to commuters with a significant share of drive-alone users (~87%), and a few share of ride-hailing pooled users (~10%). Other studies that used big data (Chen et al., 2017) did not explore socio-economic aspects in their analysis. Therefore, there is a need to add new insights into the demographic relationships to the willingness to pool through the characterization of neighborhood type using cluster analysis, the impact on demand by fare increases, and assessment of spatial and temporal patterns in pooling to determine where and when the pooled rides most often occur.

Chapter 3. Data Description

3.1 Trip Data

Trip information was sourced from the "Chicago Transportation Network Providers" trips dataset, available on the Chicago Data Portal (2020). It includes trip information that has been provided by TNCs as part of an ordinance required by the city. Ride-hailing services can be requested as "private" or "pooled" (or *shared trips*), where the rider is willing to share the vehicle with another customer during part of the trip. The database uses the designation of *"Shared Trip Authorized"* for trips requested as pooled and an additional field (number of trips pooled) to show whether the trip was successfully pooled (values greater than 1) or if it remained to be individual (value equal to 1). The latter value does not reflect occupancy and only indicates the number of trips combined together. The dataset also gives information about the starting and ending Census Tract, time, mileage, tips, and additional charges made to the customer. However, it does not include the exact latitude/longitude of the pickup/drop off location and rounds fares to the nearest $2.50.

The data retrieved contains 111,853,251 trips taken between 1/1/2019 and 1/1/2020. Trips with missing information (e.g., pick up or drop off census tract) or those marked as pooled but requested as private were removed from the dataset. The resulting data includes 74,850,391 trips (77% of the original dataset), 15.94% of which were requested as pooled. From those trips requested as pooled, 8 million (73.6%) were pooled successfully (i.e., two or more customers shared the same vehicle for some segment of their trip). Note that the requests for pooled rides and pooling success rate declined significantly starting in April 2019. Because trips were identified only by their census tract, the centroid of the tract was used as a rough estimate of the location. In addition, the distances between that centroid and the center of Chicago ("Chicago Loop.") were computed using the Haversine formula, or "as-the-crow-flies" distance and assigned to each trip for both origin and destination.

3.2 Socioeconomic Data

Socio-demographic and community variables were included in the analysis to provide a richer picture of the population characteristics at the origin and destination neighborhoods. Eighteen demographic variables were collected from the 2017 American Community Survey (ACS) 5-year estimates and included information about race, income, age, employment, educational status, household characteristics, car ownership, and population. Information about jobs from 2017 was obtained through the Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) dataset provided by the US Census (2017). Transit (bus and rail), population, and jobs density were calculated using data from the Chicago Data Portal (2020). Transit density reflects the number of stops per area of tract and population and job densities are calculated using an area that excluded spaces not permanently occupied by people, like highways or lakes. A comprehensive list of the variables and some descriptive statistics can be found in Table 1.

Table 1. Descriptive Statistics of Variables Aggregated to the Census Tract

Variables	Minimum	Maximum	Mean	Standard Deviation
Transportation-related variables				
Rail Stops/square mile	0	37.06	0.87	3.1
Bus Stops/square mile	0	312.1	57.06	31.55
% Car-Free Households	0.76	76.37	26.84	14.91
Population/square mile	422	306,484	18,195	15,779
Land-use mix and intensity-related variables				
Jobs/square mile	0	723,432	7,012	32,771
Activity density: (pop+job)/square mile	1,121	738,193	25,207	39,028
Social, economic, and demographic variables				
Median Income	10,471	160,833	54,121	29,752
Average HH Size	1.34	33.6	2.79	1.34
% HH with people < 18 years old	0.7	63.6	29.64	12.87
% HH with people living alone	7.4	77.6	34.34	13.56
% people > 25 with bachelor's degree	0.5	63.5	20.7	14.31
% Unemployed	0	48.5	10.97	8.63
% White	0	97.34	45.71	33.03
% Black	0	100	36.28	40.04
% Latino	0	99.6	25.99	28.99
% Asian	0	89.84	5.58	9.34
% Older adults	0.5	51	11.88	6.27

Note: HH = households. Statistics exclude airports (Cluster H) and tracts not used in the analysis.

Chapter 4. Methodology

4.1 Overview

The methodology for this study can be seen in Figure 1. The first step in the analysis was to categorize the trips in the database into one of three types. The process was conducted using data mining algorithms with *Python* version 3*.0*. The resulting trip types are as follows:

- *Private or individual* (P) (field *Shared Trip Authorized* = *False*) - this is a trip where the user did not choose the option to share a trip with another party. Thus, the trip was occupied by only one customer (note that the customer could be more than one rider).
- *Requested as pooled and successfully pooled (P-P)* (field *Shared Trip Authorized* = *True* and field *Trips Pooled* > 1) – here, the user chose the option to share a trip with another customer and was successfully matched with another customer.
- *Requested as pooled, but not successfully pooled (P-NP)* (field *Shared Trip Authorized* = *True* and field *Trips Pooled* = 1) – here, the user chose the option to share a trip with another customer but was not successfully matched with another party. Thus, the ride only carried one customer for the entire journey.

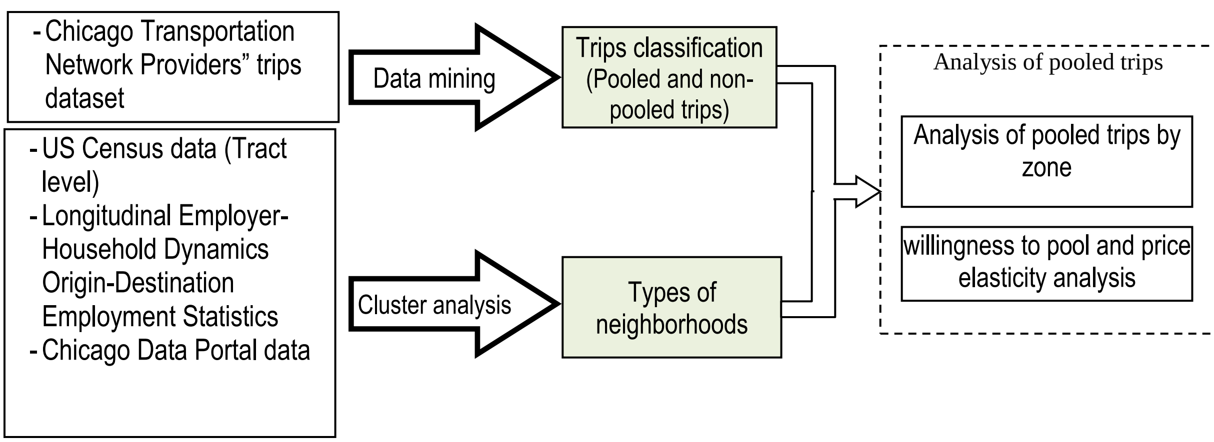

Figure 1. Methodology

The individual trips were grouped by census tract and by the time of the ride ("*Start Hour,*" which was the time rounded down to the closest hour).

The second step was to develop clusters of similar tracts using the demographic variables shown in **Table 1**. The clusters were developed by classifying the normalized raw data into a hierarchical cluster using Ward's Method[1] with an interval of squared Euclidean distance. A

[1] Ward's method is a hierarchical cluster analysis method to identify and classify homogenous groups ("clusters") based on selected variables (SPSS Statistics, 2017). The cluster analysis was used because by consolidating the many census tracts into a smaller number of categories with shared attributes, the data could more easily be assessed in terms of environmental factors. Creating a typology in this manner allowed for a very clear understanding of the spatio-demographic patterns emerging from the ride-sourcing data.

dendrogram was created, starting from one group containing all of the observations and continuing until each individual obervation has its own group.

In the next step, trip variables from the Chicago dataset were summarized and analyzed in relation to their cluster group. These variables included miles traveled, trip duration, fares (per mile, per second, tips, and any additional charges) as well as the percentage of trips requested as pooled (*%* trips r*equested as pooled* per tract) and the success rate of pooled (% trips *requested as pooled and pooled* per tract). These trip variables and statistics were compared across cluster groups using a one-way Analysis of variance (ANOVA) test, which identified statistically significant differences in trip attributes between the cluster groups. Attributes of each of the cluster groups, discussed in the following section, are representative of the average attributes for tracts in each group.

Finally, to determine what factors increase the likelihood of higher willingness to pool, the percent of requested pooled trips (i.e., trips r*equested as pooled)* was analyzed in conjunction with other variables. Also, considering the variability of prices over time (between 1/1/2019 and 1/1/2020), this study assessed the price and cross-price elasticity of demand for each of the cluster groups using Equations 1 and 2, respectively:

$$Price\ elasticity\ of\ demand\ = \frac{\%\ change\ in\ quantity\ demanded\ of\ Authorized\ rides}{\%\ change\ in\ price\ of\ Successfully\ Pooled\ rides} \tag{1}$$

$$Cross-price\ elasticity\ of\ demand\ = \frac{\%\ change\ in\ quantity\ demanded\ of solo\ rides}{\%\ change\ in\ price\ of\ Successfully\ Pooled\ rides} \tag{2}$$

Chapter 5. Results

5.1 Neighborhood Typology

We chose eight groups that differed based on socio-demographic, transportation, and density characteristics as shown in **Table 2**. Three census tracts fell outside the characteristics of any of these groups. These uncategorized tracts were omitted because they had characteristics that were not generalizable to any of the clusters; for example, they had missing information about Median Income. In table 2, Clusters A and B both have very low activity densities and incomes. Cluster A's activity density (12,325 jobs + population per square mile) is the lowest of all groups, while Cluster B's median income of $25,024 is lower than all others. These two clusters have similar educational attainment (low percentage of people 25 and older with bachelor's degree), the highest unemployment of all cluster groups (21% and 16%, respectively), and high percentages of people without cars. They both have above-the-average bus availability, with Cluster B having a reasonably relatively high rail density of 1.31. Cluster B has the highest percentage of households that are car-free out of all other cluster groups, with an average of 46.59% per tract. Both of these clusters have a majority African American population (93% and 88%, respectively). Therefore, cluster A is labeled as *Medium Transit, Low Activity, Low Income, Majority Black*, and cluster B is labeled as *High Transit, Low Activity, Lowest Income, Majority Black.*

Clusters C and D both have incomes of slightly above $45,000 (below the clusters' average of $ 54K) and activity densities of about 22,000 (also below the average of 24,000). Cluster C has the highest number of households with children under 18 (42.32%), the lowest percentage of people living alone (21.08%), the highest household size (3.43), and a majority Latino population (77.87%). Cluster D is very racially diverse and has the highest percentage of Asian population compared to other clusters (14.99%). Both have an unemployment rate of around 10% and have below the average percentage of people (25 and older) with bachelor's degrees.

Table 2. Mean Values of Characteristics Describing Each Neighborhood Type

	Cluster							
Variables	A	B	C	D	E	F	G	H
Transportation-related variables								
Rail Stops/square mile	0.44	1.31	0.3	0.5	0.12	1.47	3.17	0.5
Bus Stops/square mile	54.41	64.52	52.6	54.3	37.32	63.74	83.43	12.24
% Car-Free Households	34.26	46.59	17.79	23.89	10.56	21.19	43.93	-
Population/square mile	11,092	12,676	19,697	18,547	11,850	22,750	37,060	0
Land-use mix and intensity-related variables								
Jobs/square mile	1,233	2,372	2,773	3,470	2,672	9,223	37,148	5170.31
Jobs to population ratio	0.15	0.29	0.23	0.38	0.28	0.62	1.56	-

Activity density: (pop+job)/square mile	12,325	13,502	22,470	22,017	14,523	33,575	74,208	5170.31
Social, economic, and demographic variables								
Median Income	32,284	25,024	45,144	45,587	74,597	99,491	60,005	-
Average HH Size	2.92	2.19	3.43	3.36	2.71	2.31	1.86	-
% HH with people < 18 years old	34.85	20.33	42.32	32.39	30.35	20.38	9.98	-
% HH with people living alone	34.15	54.59	21.08	30.16	27.9	35.28	56.69	-
% people > 25 with bachelor's degree	10.06	12.65	10.96	17.93	23.81	43.27	35.15	-
% unemployed	21.08	15.88	9.37	10.93	4.95	3.42	5.24	-
% White	3.32	6.58	58.12	37.34	79.57	81.45	67.51	-
% Black	92.98	87.75	4.36	32.25	3.57	5.1	15.22	-
% Latino	3.88	4.29	77.87	29.46	23.68	16.07	10.4	-
% Asian	0.3	2.39	2.43	14.99	7.09	6.5	11.49	-
% Older adults	14.02	17.32	8.98	10.49	14.53	6.69	15.12	-

Cluster E represents tracts that lie in the northwest part of Chicago and on the outskirts of the city. The tracts have a median income of $74,597. This cluster has a white majority and is mostly low-density residential land use. Cluster E also has very limited transit, with the lowest rail and bus density of all groups (0.12/mile and 37.32/mile, respectively). This cluster also has the highest car ownership of any of the other clusters, with only 10.56% of households being car-free. Cluster E is mostly White (79.6%) with a low unemployment rate (below 5%) and a percentage of population with bachelor's degrees slightly above the clusters' average. Therefore, this cluster was labeled as *Low Transit, Low Activity, High Income, and Majority White.*

Clusters F and G are, for the most part, contained to the northeast side of Chicago and represented higher-income tracts as well. These Cluster groups are the predominant typology of the downtown tracts, with Cluster G containing the "Loop." These are the tracts with the highest activity densities, with cluster G having the maximum number (74,208). Both are transit-rich, with Cluster G having a high car-free percentage of 43.93%. Cluster F has the highest median income of all groups, with an average of $99,491 per household, and it has the lowest percentage of older adults, 6.69%. Cluster G appears to have fewer families, with a household size of only 1.86 and high percentages of older adults, households where the individual lives alone, and households without children under 18. Clusters F and G also have a majority white population and are the tracts with the highest percentage of people with bachelor's degrees and the lowest unemployment rates. These clusters are called *High Transit, Medium/High Activity, Highest Income, Majority White*, and *Highest Transit, Highest Activity, Medium Income, and Majority White*, respectively.

Cluster H is made up of the two tracts with Chicago's airports, O'Hare International and Midway International. These do not have any residents living there, and thus many of the demographic variables used to describe the other clusters are absent. The airports are major job centers, with 34,745 reported in the O'Hare tract and 6,689 in the Midway tract.

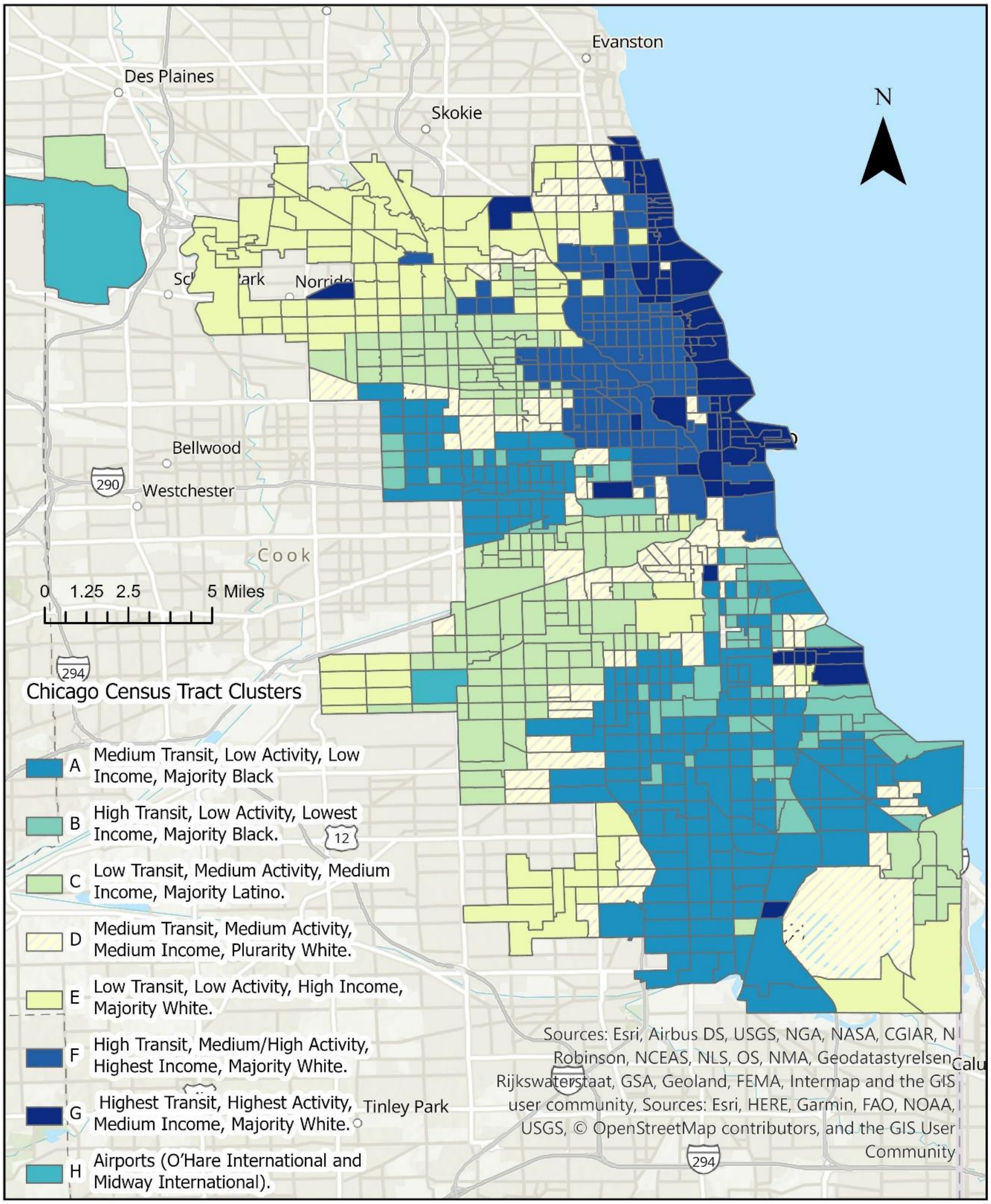

Figure 2. Types of Neighborhoods Across Chicago

5.2 Analysis of Trip Data Across Cluster Types

The trip dataset was analyzed to explore temporal and spatial differences in pooling behavior and trip characteristics across socio-demographic groups. Three key factors were evaluated in this process. The percentage of trips requested as pooled, interpreted as a willingness to use pooled rides, sensitivity to price changes, and pooling success rates. Note that the number of requested pooled trips and successfully pooled trips declined starting in April 2019.

5.3 Willingness to Pool

The percentage of a census tract's trips that are *requested as pooled* varies significantly amongst the identified cluster groups. Assuming that the *pooled* option is available in all tracts, this section explores that variation and its associated factors.

The lowest percentage is in Cluster H (the airports), which average an authorization rate of 8.94%, while tracts in Clusters A and B (low-income, majority Black) have a rate of pooled trip requests of around 40%. Median income appears to play a very large role in this rate, consistent with the findings of Hou et. al. (2020). As can be seen in Figure 2, which depicts the individual census tracts as blue dots and the clusters as red stars. The percentage of *trips requested as pooled* decreases as the median income of a tract rises. The effect is more drastic for incomes below $80,000, with the data exhibiting a plateauing effect after this threshold. As a result, the rate of trips requested as pooled in low-income clusters (A and B) is 2 to 2.7 times the rate of requested trips in high-income clusters (F and E).

Similarly, clusters with middle-income (C and D) show rates between 1.5 and 2 higher than clusters F and E. This finding highlights the importance of these *pooled s*ervices in low-income areas as they are more affordable than single rides. However, it should be noted that transportation costs using *pooled* services could still represent significant expenditure for households in this area.

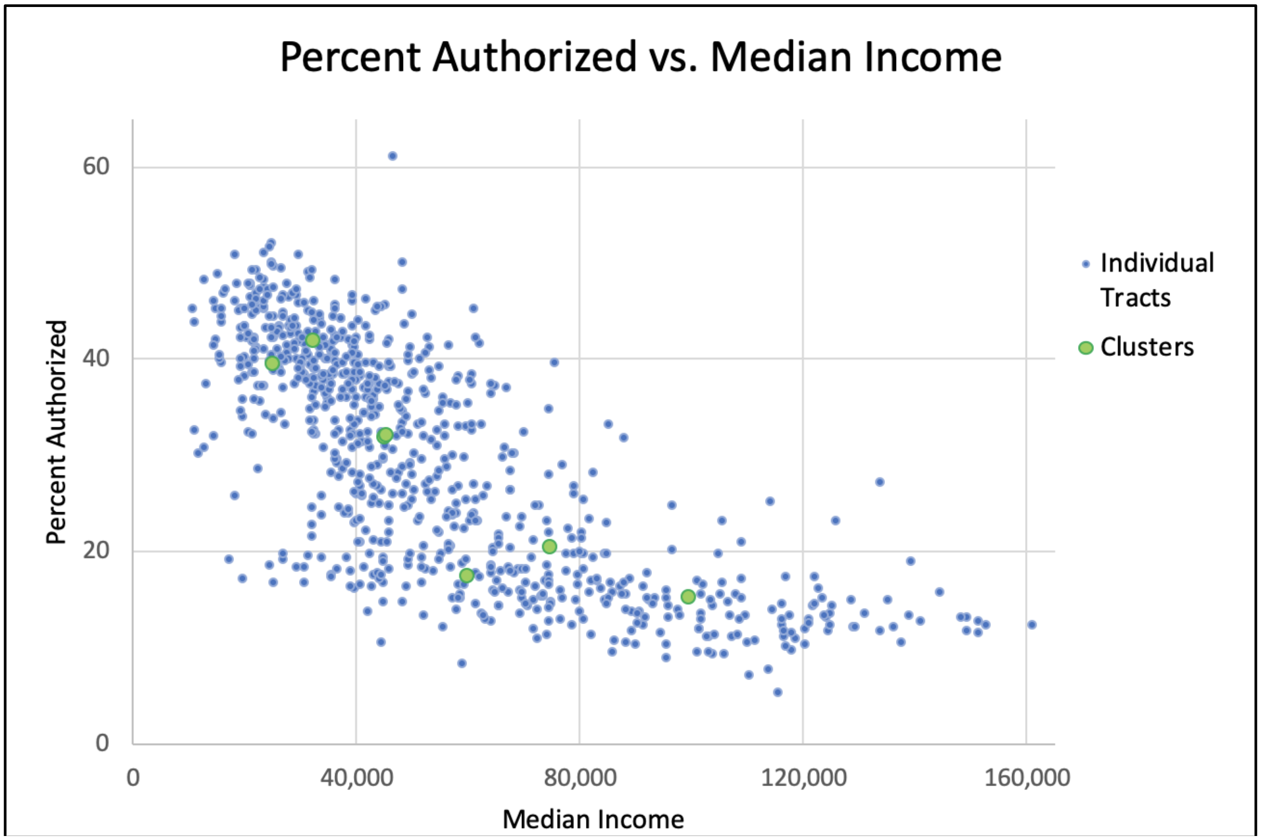

Figure 3. Relationship Between Income and the Percentage of Shared Trip Authorized Rides

Note: Cluster H is not shown because Median Income is not available

Willingness to pool, represented by the percent of trips *requested as pooled*, is also associated with the length of the journey. The average trip mileage in a tract and the percent of trips in a tract that are *requested as pooled* are positively correlated with a Pearson coefficient of 0.35. The trip length seems to also play a role in pooling success, consistent with findings by Kucharski and Cats (2020), who noted that finding matches are easier when trips are long. Amongst trips *requested as pooled*, longer trips make up a larger share of the successfully pooled (i.e., *requested as pooled and successfully pooled)* trips than the unsuccessfully pooled trips. The mean distance for successfully pooled trips is 5.90 miles, while unsuccessful trips are, on average, 3.38 miles. **Figure 4** shows the upward trend in the share of rides that are successfully pooled as mileage increases, with the highest percentage being 89.7% between 8-10 miles. Increasing willingness to pool longer trips is likely due to the incentive to reduce costs on longer, more expensive trips. This may be especially true given that some of the cluster groups that have longer commutes (see **Table 3**) are also those whose neighborhoods have fewer financial resources. This correlation with distance may also be influenced by the availability of other modes, such as transit, cost, travel time, and destinations served, among other factors (See, for example, Schwieterman and Livingston, 2018).

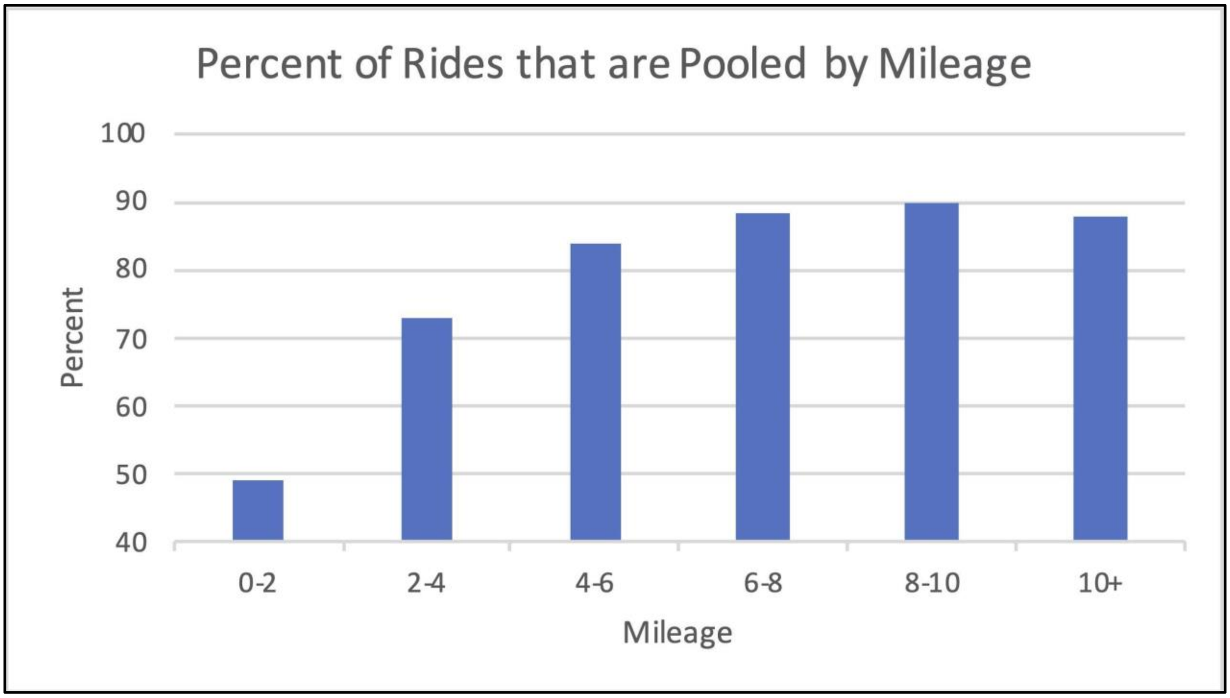

Figure 4. Ratio of Number of Pooled Rides to Number of Non-Pooled Rides by Mileage

Time of day also correlates with the willingness to pool. As depicted in **Figure 5**, trips *requested as pooled* are highest during the peak commuting hours of the day. Willingness to pool seems highest during the peak commuting hours of the day. Without considering airports, the highest demand occurs between 6-7 a.m. and 4-5 p.m. This suggests that ride-hailing, and specifically pooled rides, are being used to commute to and from work. Trips going to and from the airports generally have a much lower willingness to pool, but the percentages spike at night. It is unclear what the reasons for this trend are, but it could be related to less time sensitivity at night when traveling to or from an airport than during the peak day times, safety concerns, price surging at night, or airport employees utilizing shared rides. Pooling success follows a similar pattern, indicating a correlation between as more riders are willing to pool, there is a higher likelihood of rides being pooled. Note that the willingness to pool could be a reaction to surge pricing (higher fares) during peak hours.

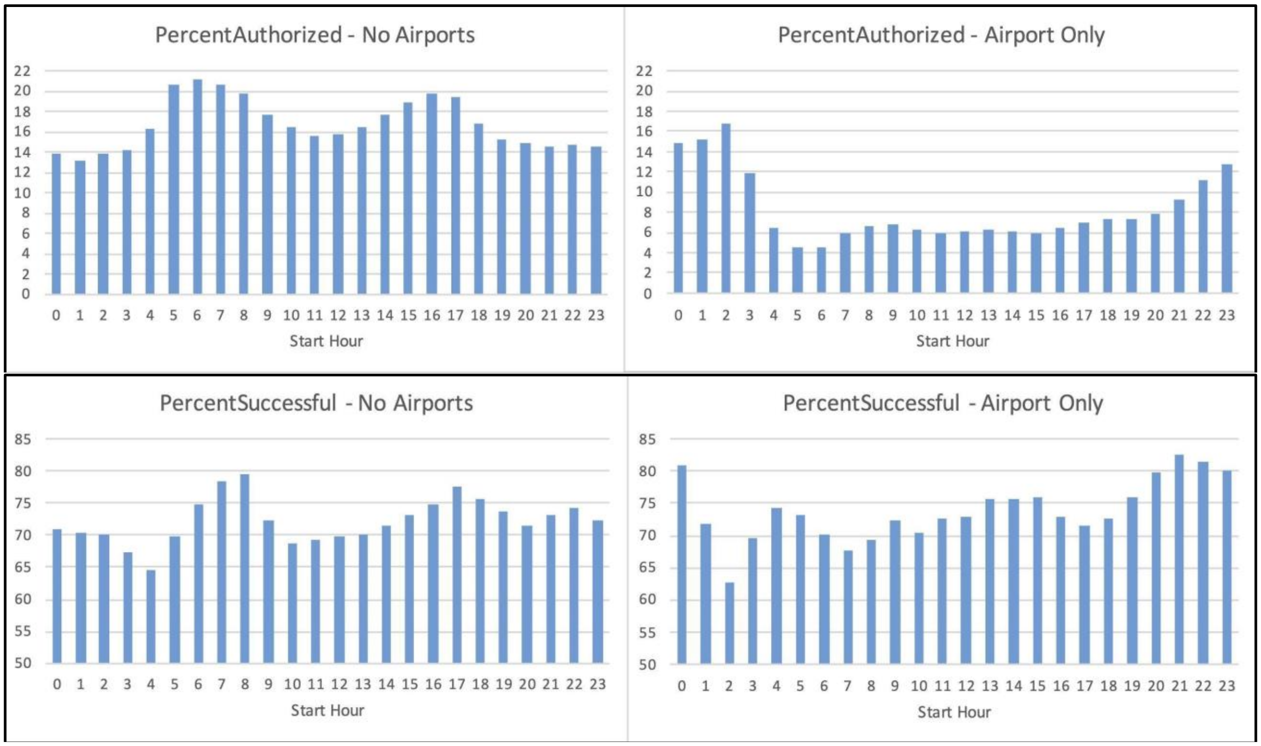

Figure 5. Percentage of Rides that were Requested As Pooled (Shared Trip Authorized) Per Hour for all tracts except airport locations.

5.4 Sensitivity to Fare Change

It was also observed that the number of trips *requested as pooled* began steadily declining after March 2019. Between January and March 2019, there were 4.4M trips requested as pooled, and this number declined to 1.9M trips between October to December (a decline of about 55%). The effects were most pronounced in the downtown high-income tracts, which also experienced the most drastic decrease in demand for pooled trips. During this period, a steady price increase also occurred (see Figure 5), consistent with a Reuters analysis from November 2019 (Bellon 2019), while the fare per second for solo rides remained consistent. Figure 5 shows Single Ride and non-Pooled Authorized trendlines. The fare/second for trips requested as pooled shows an upward trend. From January to December, the fare per second for pooled trips rose nearly 30%. The reasons behind this fare increase are unclear but point to a new pricing strategy by Uber, which has around 72% of the market share in Chicago (Bellon 2019).

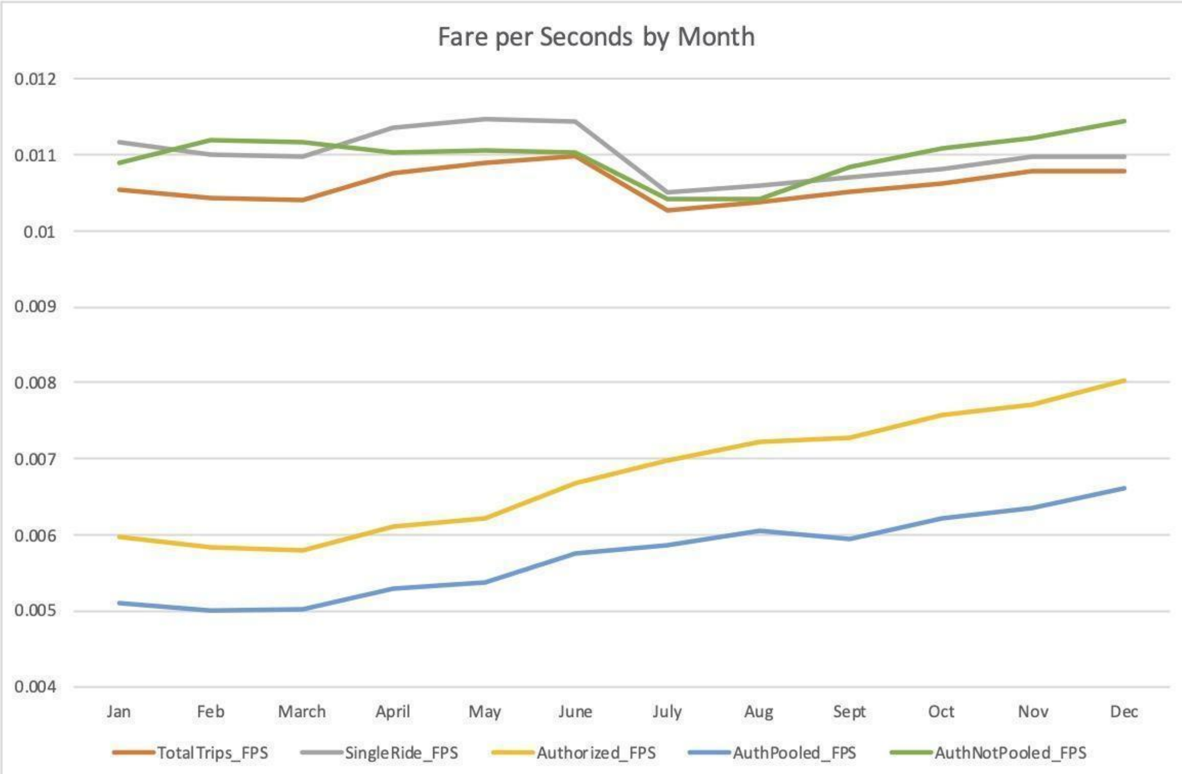

Figure 6. Fare per second by Trip Type

The rising cost of pooled rides may be a significant cause of the decline in trips *requested as pooled*, which are expected to be very price-sensitive due to the time and cost tradeoff, which acts as a deterrent (Alonso-González et. al., 2020). However, the level of sensitivity to price changes varies across cluster groups, as shown in Table 4. The price and cross-price elasticity of demand for each of the cluster groups is assessed using Equations 1 and 2.

The elasticities reflect the different rates at which the desire for each type of trip changed based on the fare increases. For nearly all clusters except Clusters A and H, the price elasticity of demand for Authorized trips had values less than -1, which is considered elastic. This means that demand falls more per unit than the fare per second increases.

On the other hand, most tracts had an inelastic or nearly unit-elastic cross-price elastic demand for individual rides. This means that when the price of the pooled rides was raised, in most cases, the demand for solo rides did not increase at the same rate (in fact, in most cases, it was less). The most elastic response (where individual ride requests outpaced the rate at which the pooled ride fare increased) was for cluster A.

Noticeably, the trends are somewhat opposite of each other – the tracts in which demand for pooled rides fell at a disproportionately large rate to the price increase experienced a disproportionately small increase in individual ride requests. Thus, it is unlikely that the demand shifted from pooled trips to individual trips. Based on the availability of cars to households in those tracts (Table 2), it is possible that trips shifted to personal automobiles, or in the case of a highly transit-rich tract like Tract G, transit.

Table 3. Elasticities and Change in Demand for Authorized Rides

Cluster	% Demand Change (Pooled Rides)	% Demand Change (Individual Rides)	% Pooled Fare Change	Price Elasticity	Cross Price Elasticity
A	-32%	38%	32%	-0.99	1.17
B	-48%	35%	34%	-1.41	1.04
C	-52%	30%	28%	-1.86	1.06
D	-55%	22%	28%	-1.95	0.78
E	-52%	25%	27%	-1.9	0.90
F	-59%	16%	28%	-2.08	0.56
G	-59%	22%	29%	-2.02	0.76
H	-7%	26%	19%	-0.34	1.34

Trip characteristics of pooled trips are shown in Table 4. Because detailed information on how these trips occur is pooled by each company making the information proprietary and not publicly available. However, understanding aggregated patterns of trips that have been *successfully pooled* can provide some insights.

The overall pooling success rate, which is the number of *successfully pooled* trips divided by the number of trips *requested as pooled*, was 73.57%. The success in pooling rides is also influenced by rides having common origins or destinations. Cluster H, the airports, has a success rate of 78.95%, which is likely indicative of the common origin that airport trips have. Cluster E, on the other hand, has a comparatively low success rate of 63.30%. Cluster E has the third lowest activity density (so there are fewer places where people are coming to and from), combined with the third lowest willingness to pool.

Additionally, Cluster E has the lowest transit availability of all groups. An analysis of the pooling success by drop off location appears to correlate with rail stops. Therefore, people may be using ride-sourcing services as a link to transit. With less transit available, this common destination is less prominent, which may explain comparatively lower success rates in the north and south of Chicago.

Pooling success declined within the same time period of the fare increase, potentially due to less availability of eligible trips or company policy changes to how pooling requests are satisfied. The pooling success dropped from 77% between January and March to 67% between October to December. The decline for most of the cluster groups was similar, with the percentage of successful trips dropping between 13%-15%. Clusters C and E were slightly above this, with 17% and 20% reductions, respectively. Cluster H, the airports, maintained a similar level of success as it did previously, with only a 3% reduction in successfully pooled trips.

Lastly, trips *requested as pooled but not successfully pooled* may more often be going away from the center (the Chicago Loop in this analysis). The mean distance from this point for all unmatched trips at the origin is 3.99 miles, while the destination is 4.17 miles away. For successful trips, the average origin is 3.77 miles away from the center, while the destination is 3.64 miles. Successfully pooled trips start farther from the center in the morning than they do in the evening. Trips that are unsuccessfully pooled tend to start closer to the center than successful trips in the morning. It seems that trips going to and from a centralized, common location, such as the downtown job center, may be more likely to get pooled than those going to less popular locations. Specifically, during morning commuting hours between 6–9 a.m., there is more success toward the center as it gets later during the commuting period and less overall success after 9 a.m. Opportunities to pool likely correspond to where most people are leaving from.

Table 4. Mean Values of Trip Characteristics by Neighborhood Type

	Cluster							
Variable	A	B	C	D	E	F	G	H
Trip Miles	6.39	5.99	6.08	5.81	6.95	4.06	4.99	15.53
Trip Seconds	1,161	1,128	1,240	1,163	1,258	944	1,012	2,142
Fare Pooled/mile	1.16	1.17	1.17	1.2	1.13	1.39	1.31	1.38
Fare Not Pooled/mile	2.68	2.7	2.62	2.69	2.38	2.72	2.81	1.62
Fare Pooled/seconds	0.0059	0.0057	0.0051	0.0053	0.0055	0.0054	0.0055	0.0093
Fare Not Pooled/seconds	0.0128	0.0125	0.0113	0.0117	0.0115	0.0104	0.0111	0.0127
Tip	0.18	0.24	0.36	0.37	0.66	0.52	0.55	1.99
Additional Charges	2.24	2.27	2.59	2.49	2.81	2.57	2.63	7.36
Trip Total	13.29	13.04	14.07	13.55	16.07	12.5	13.47	34.88
Total Count	10,804	27,571	20,183	39,307	21,108	234,753	392,306	1,340,602
% Authorized Per Tract	41.81	39.46	31.93	32.05	20.34	15.2	17.52	8.94
% Successfully Pooled Per Tract	70.32	73.05	73.34	72.73	63.3	72.37	73.64	78.95
Proportional Difference - Authorized	0.00026	0.00049	0.00023	0.00034	0.00009	-0.00045	-0.00105	-0.00905
Proportional Difference - Pooled	-0.00006	0.00003	0.00001	0	-0.00009	-0.00009	0.00031	0.00284

Chapter 6. Discussion & Conclusions

6.1 Discussion

From this research, a number of equity and policy questions arise. A comparison of the clusters and pooled trip statistics demonstrated varying levels of reliance on pooled trips between different types of census tracts. Lower-income census tracts exhibit more reliance on ride pooling than tracts with higher-income residents, as exhibited by the negative correlation of a tract's income to authorization rate and the low-price elasticities for pooled rides in low-income clusters. This may be due to pooling's function as a cost-saving measure. Therefore, individuals who have a more limited income may be dependent on these rides to fill gaps in mobility. There are also significant spatiotemporal attributes behind the willingness of a person to pool a ride-sourcing ride and whether or not their request is met, which may affect the availability and affordability of these type of services. In that view, although offering significantly more flexibility in routes and schedules, the market-dependent costs (compared to bus fares, for example) could add a significant economic burden to low-income households. This possibility should be considered when public agencies consider providing transportation services through TNCs, and a mechanism to address those issues should be proposed.

Furthermore, the temporal analysis in 2019 showed that financial factors play an important role in the decision to pool a ride. As understood by the resulting decline in trips *requested as pooled* that occurred during the period when prices for pooled rides began to increase, ride pooling is an extremely price-sensitive activity. Because of the elastic nature of the demand for pooled rides, it is suggested that prices be kept as low as possible to incentivize consumers to choose that option. As Alonso-González et al. (2020) discussed, the time-cost tradeoff of pooled rides plays a major role in the desire to pool a ride, with greater cost benefits yielding higher ridership. This would also help keep costs low for people who are more reliant on these services because they are commuting from areas that are underserved by transit.

Without data such as occupancy numbers, exact fares, and specific pickup/dropoff information, it is difficult to characterize the ways in which pooled rides are being used, by whom, and the cost per person. Furthermore, the analysis herein only allows to draw conclusions about general trends rather than specific individuals' motives for using pooled rides. Further research could be undertaken using survey methods, which would allow for more detailed answers to be obtained. Additionally, a limitation of this data is that it was from 2019, before the COVID-19 pandemic. This means that the results could be different than the trends of the post-pandemic world, so while these are undoubtedly useful for informing policy, they should also be considered in tandem with data about how the pandemic impacted Chicago's communities.

6.2 Conclusions

This research used a sample of trips taken in Chicago during 2019 to explore what motivates a person's decision to use pooled ride-sourcing services and what impacts whether those rides are successfully matched with others. Financial factors and variables affecting the number of available pooling opportunities were the most significant elements that were analyzed in this study. The study used 74 M ride-sourcing trips, from which 15.94% were *requested as pooled* and 73.57% were *successfully pooled*. Eight clusters were developed to compare trip characteristics, and it was found that. While pooling success was consistent in all clusters, the willingness to pool ranged from 8.94% at airports to 41.81% in Cluster A. With a fare increase, the price elasticity of demand ranged from -0.34 to -2.08.

Encouraging pooling could unlock myriad benefits, and policymakers should consider this tactic alongside their policy objectives. Pooled rides have the potential to decrease VMT by combining trips on similar routes into one journey. However, as discussed by the Union of Concerned Scientists (Anair et al., 2020), pooled rides still surpass the emissions from mass transit, in part due to deadheading. Public policy could incentivize using ride-sourcing as a complement, rather than a substitute, to transit through partnerships between transit agencies and ride-sourcing companies (Jaffe 2015). Encouraging pooling would also have positive equity effects since there would be more opportunities for successful matches of rides, thus allowing lower fares to be attained by consumers who are reliant on ride-sourcing to get to work. Lower-income areas that have fewer mobility options are the most reliant on the low-fare options provided by ride-sourcing companies. Because our findings show that ride-sourcing has cost implications that are felt unequally by different neighborhoods, these must be considered carefully alongside pooling benefits. Our research aims to help inform policies that regulate ride-sourcing companies to ensure that equity is maintained in the legislative process. With a better understanding of how this transportation niche is functioning for users and the importance of its role in the overall context of their mobility options, there is a strong foundation for future research and policies.

References

Abouelela, M. (2020). *Why do we share our trips? Investigation of the factors that influence the acceptance, frequency of use and characteristics of organized pooled rides. The Jetty case study in Mexico* [Master's Thesis, Technical University of Munich]. https://mediatum.ub.tum.de/doc/1550475/542412.pdf

Alonso-González, M. J., Cats, O., van Oort, N., Hoogendoorn-Lanser, S., & Hoogendoorn, S. (2020). What are the determinants of the willingness to share rides in pooled on-demand services? Transportation. https://doi.org/10.1007/s11116-020-10110-2

Anair, D., Martin, J., Pinto de Moura, M., & Goldman, J. (2020). Ride-Hailing's Climate Risks: Steering a Growing Industry toward a Clean Transportation Future. https://www.ucsusa.org/resources/ride-hailing-climate-risks

Babar, Y., & Burtch, G. (2017). Examining the Heterogeneous Impact of Ridehailing Services on Public Transit Use. https://doi.org/10.2139/ssrn.3042805

Bellon, T. (2019, November 26). Uber drives up prices for shared rides, Chicago data reveals. Reuters. https://www.reuters.com/article/us-uber-pricing-chicago-idUSKBN1Y017V

Beojone, C. V., & Geroliminis, N. (2021). On the inefficiency of ride-sourcing services towards urban congestion. Transportation Research Part C: Emerging Technologies, 124, 102890. https://doi.org/10.1016/j.trc.2020.102890

Chen, X. (Michael), Zahiri, M., & Zhang, S. (2017). Understanding ridesplitting behavior of on-demand ride services: An ensemble learning approach. Transportation Research Part C: Emerging Technologies, 76, 51–70. https://doi.org/10.1016/j.trc.2016.12.018

Circella, G., Alemi, F., Tiedeman, K., Handy, S., & Mokhtarian, P. (2018). The Adoption of Shared Mobility in California and Its Relationship with Other Components of Travel Behavior. https://escholarship.org/uc/item/1kq5d07p

City of Chicago [dataset]. (2020a). Chicago Data Portal Shapefiles. Chicago. https://data.cityofchicago.org/browse?tags=shapefiles

City of Chicago [dataset]. (2020b). Transportation Network Providers—Trips. Chicago Data Portal. https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips/m6dm-c72p

Dong, Y., Wang, S., Li, L., & Zhang, Z. (2018). An empirical study on travel patterns of internet based ridesharing. Transportation Research Part C: Emerging Technologies, 86, 1–22. https://doi.org/10.1016/j.trc.2017.10.022

Ermagun, A., & Tilahun, N. (2020). Equity of transit accessibility across Chicago. Transportation Research Part D: Transport and Environment, 86, 102461. https://doi.org/10.1016/j.trd.2020.102461

Ferguson, E. (1997). The rise and fall of the American carpool: 1970-1990. Transportation, 24, 349-376. https://doi.org/10.1023/A:1004928012320

Grahn, R., Qian, S., Matthews, H. S., & Hendrickson, C. (2020). Are travelers substituting between transportation network companies (TNC) and public buses? A case study in Pittsburgh. Transportation. https://doi.org/10.1007/s11116-020-10081-4

Henao, A., & Marshall, W. E. (2019). The impact of ride-hailing on vehicle miles traveled. Transportation, 46(6), 2173–2194. https://doi.org/10.1007/s11116-018-9923-2

Hou, Y., Garikapati, V., Weigl, D., Henao, A., Moniot, M., & Sperling, J. (2020). Factors Influencing Willingness to Pool in Ride-Hailing Trips. Transportation Research Record, 2674(5), 419–429. https://doi.org/10.1177/0361198120915886

Jacob, J., & Roet-Green, R. (2018). Ride Solo or Pool: Designing Price-Service Menus for a Ridesharing Platform(SSRN Scholarly Paper ID 3008136). Social Science Research Network. https://doi.org/10.2139/ssrn.3008136

Jaffe, E. (2015, August 3). Can Uber and Public Transit Really Get Along? Bloomberg CityLab. https://www.bloomberg.com/news/articles/2015-08-03/uber-is-partnering-with-public-transit-agencies-for-first-last-mile-service

Kucharski, R., & Cats, O. (2020). Exact matching of attractive shared rides (ExMAS) for system-wide strategic evaluations. Transportation Research Part B: Methodological, 139, 285–310. https://doi.org/10.1016/j.trb.2020.06.006

Lavieri, P. S., & Bhat, C. R. (2019). Modeling individuals' willingness to share trips with strangers in an autonomous vehicle future. Transportation Research Part A: Policy and Practice, 124, 242–261. https://doi.org/10.1016/j.tra.2019.03.009

Movable Type Scripts. (2020). Calculate distance, bearing and more between Latitude/Longitude points. https://www.movable-type.co.uk/scripts/latlong.html

Saadi, I., Wong, M., Farooq, B., Teller, J., & Cools, M. (2017). An investigation into machine learning approaches for forecasting spatio-temporal demand in ride-hailing service. ArXiv:1703.02433 [Cs, Stat]. http://arxiv.org/abs/1703.02433

Schwieterman, J., & Livingston, M. (2018). Uber Economics: Evaluating the Monetary and Nonmonetary Tradeoffs of TNC and Transit Service in Chicago, Illinois. DePaul University: Chaddick Institute for Metropolitan Development. https://las.depaul.edu/centers-and-institutes/chaddick-institute-for-metropolitan-development/research-and-publications/Documents/Uber%20Economics_Live.pdf

U.S. Census Bureau [dataset]. (n.d.). Longitudinal Employer-Household Dynamics. Retrieved February 21, 2021, from https://lehd.ces.census.gov/data/

U.S. Census Bureau [dataset]. (2017). American Community Survey. https://data.census.gov/cedsci/

Xu, K., Sun, L., Liu, J. (2018). An empirical investigation of taxi driver response behavior to ride-hailing requests: A spatio-temporal perspective. PLoS ONE 13(6): e0198605. https://doi.org/10.1371/journal.pone.0198605

Yan, C., Zhu, H., Korolko, N., & Woodard, D. (2018). Dynamic Pricing and Matching in Ride-Hailing Platforms. Naval Research Logistics. http://dx.doi.org/10.2139/ssrn.3258234

Zhang, Y., & Zhang, Y. (2018). Exploring the Relationship between Ridesharing and Public Transit Use in the United States. International Journal of Environmental Research and Public Health, 15(8). https://doi.org/10.3390/ijerph15081763