Department of City and Regional Planning Fall 2022
University of Pennsylvania Xinge Zhang, Yiping Ying

**E-scooter usage analysis: A comparison study of Austin, TX and Chicago, IL**

## Abstract:

Governments around the world have taken different approaches to regulate e-scooter sharing. Some have embraced it to reduce traffic congestion and improve air quality, while others have been more hesitant to allow e-scooter sharing due to concerns about safety and liability. In 2018 and 2019, the governments of Austin and Chicago began programs for e-scooter pilots, respectively. From the previous study, we can see that the use of scooters is greatly limited by the location of the pilot arranged by the government, therefore, to facilitate the government to better open the pilot sites for citizens. We will build models to predict e-scooter usage in Austin or Chicago separately to help planners further understand the motivation and behavioral characteristics of people using shared scooters.

### 1. Introduction:

E-scooters have generally become a popular transportation option in many cities around the world in recent years. They offer a convenient, affordable, and environmentally friendly way to get around for short trips. The use of electric scooters can be influenced by various factors such as the availability of electric scooter rental services, the cost of using electric scooters, the availability of alternative transportation options, and local laws and regulations. From June to October 2019, the City of Chicago opened to set up a large number of scooter sites and collect trip data. The government sought to collate and analyze the data to help select the location of future scooter sites, and to explore the number of scooter placements at different sites and how they are regulated. Based on this, we tried to set up a predictive comparison model of scooters across different cities. Through analysis of trip data and feature importance in the model to help decision-makers better understand the motivation and behavioral characteristics of people using shared scooters.

In the comparative analysis of modeling, we chose to use Austin's scooter use as a control. Unlike Chicago, which just ran a scooter pilot, Austin's scooter program has been running longer and started earlier, and the long program cycle means that local people's awareness of scooter use has been more mature, and the coefficient of the different predictors in the modeling model may more accurately reflect the determinants of scooter use. In addition, the two cities are similar in terms of household size and median age, which means that the two cities can derive a more universal value of urban environment determinants of scooter use through the cross-sectional comparison analysis of the same elements between different models, which is useful for establishing similar sites for scooter sharing services and evaluating scooter shared services are valuable.

### 2. Literature Review:

Micro-mobility transportation develops at an ever-accelerating rate. Generally, the industry refers to such an intertwined network in which users can fulfill their short-distance travel demands by renting one of many shared, small vehicles such as bikes and scooters as shared active transportation (National Association of City Transportation Officials (NACTO), 2018). Related research could hardly catch up with the speed, the definition of such transportation mode varies due to different criteria, while at present this new kind of travel mode is open to multiple criteria that characterize small-sized vehicles, including weight (less than 500 kg), passenger capacity (mostly one passenger only), maximum travel duration/distance, and so on. (Zarif et al., 2019).

Before the advent of e-scooters, shared bicycles were the most representative tools in micro-mobility. Most of the research literature revolves around shared bicycles, which mainly focuses on the following aspects: comparison between docked/dock less bike sharing systems; spatial/temporal patterns (Zhou et al., 2020); social-economic factors (Gu et al., 2019) and the linkages between the built environment and travel behaviors (Ewing and Cervero, 2010). The most common topics in this area are those exploring the relationship between various socioeconomic indicators and shared bikes. Similar research was conducted in different cities, such as Seattle, US (Mooney et al., 2019), and Singapore (Xu et al., 2019). The main conclusions are that higher income/ highly educated and young neighborhoods tend to have higher availability of shared bikes.

Though e-scooters prosper, there is still not much literature on this topic. The current articles are mainly comparing e-scooters with shared bicycles.  Among the existing e-scooter studies, McKenzie compared the spatiotemporal usage patterns of e-scooters with the station-based bike-share system in Washington D.C. and found that e-scooter trips were more similar to casual bike-share trips in terms of the time of use and were dissimilar spatially (McKenzie, 2019). Another study in Indianapolis, IN, reported that downtown and universities showed heavy e-scooter traffic, and the usage peaked in the afternoon rather than in the morning, indicating citizens in Indianapolis seldom used e-scooters for their morning commute (Mathew et al., 2019).

In 2020, Bai and Jiao investigated e-scooter ridership in Austin and Minneapolis using GIS hotspot spatial analysis and negative binomial regression models. The spatial analysis results showed that the densest e-scooter usage happened in downtown areas and university campuses in both cities. In Austin, afternoons, and weekends experienced greater e-scooter traffic. The regression also indicated that proximity to the city center, better access to transit, and greater land-use diversity positively correlated with higher e-scooter ridership (Bai & Jiao, 2020). This study supplemented the literature pool with rigorous empirical evidence of e-scooter operations and provided a good case study in Austin.

More and more cities in the U.S. are planning to launch e-scooters. It is of great importance and urgent for U.S. cities to understand and predict the usage after having e-scooters running on the roads. Due to a lack of experience, many predictions are made based on the study of shared bicycles. However, the major problem of this method lies in the significant difference between e-scooters and shared bicycles. Shared bicycles are used to commute in general, while e-scooters are more often applied for leisure use (Zhou et al., 2020). Most of the users differ as well, and the usage of e-scooters is more constrained to the urban terrain.

Based on these observations, we decide to create an analytical model that can better determine which areas in the same city will better take scooter potential and whether there is a difference in user population between cities. Through such a census tract-based analysis model, we will be able to help decision-makers to identify more valuable investment areas, and help planners further understand the motivation and behavioral characteristics of people using shared scooters.

## 3. Materials & Methods

### 3.1 Research design:

This study uses individual trip data from Austin and Chicago to examine the association between electric scooter use and different types of land use and the complexity of land use within the region. In the first part of the study, we compare Austin and Chicago by analyzing the temporal distribution of trips and urban built environment indicators to identify similarities and differences between Austin and Chicago. In the second part, we use the random forest and multiple linear regression models to predict scooter usage and find the importance of the same built environment factors in different cities' scooter usage. In the third part, we compare scooter policies in Austin and Chicago and use a word cloud to analyze the public response to government policies. This will help planners better understand the significance of share micro-mobility for citizen travel and provide inspiration for better policy formulation and scooter site planning.

### 3.2 Study area and data：

The study area is in Austin, TX, and Chicago, IL. Although the two cities are very different in size, they also have some commonalities. (Table 1) For example, both cities have relatively densely populated areas for scooter pilot programs in 2019, including recreational and tourist areas. The two cities are similar in population age structure as well as household structure. If we lower the geographic unit of the model regression to control at the census tract level, we will make the results of the two cities comparable.

Table 1:
Land Area and Demographic Features of Austin and Chicago. (2019)

| | Austin | Chicago |
|---|---|---|
| Land Area ($mi^2$ ) | 319.94 | 227.37 |
| 2019 Population | 950,807 | 2,709,534 |
| Population Density(people/$mi^2$) | 2,971.84 | 11,916.92 |
| 2019 Median Household Income | 43,043 | 58,247 |
| Average Household Size | 2.44 | 2.48 |
| Median Age | 33.3 | 34.6 |

* Source: The United States Census Bureau

As shown in Table 2 below, the two datasets used in this study are individual trip records officially posted by local transportation departments on shared data sites. As part of their contracts with municipalities, licensed service providers are obligated to share their operational histories, including trip records, with cities as they deploy and operate their fleets. The logs record the defining characteristics of each motorized scooter trip, including origin-destination (OD) location, date, start and end times, trip duration, and distance. To protect privacy, the OD coordinates in the Austin and Chicago datasets are aggregated at the center of the census tract, and the trip start and end times are regressed to hourly units.

Table 2:E-scooter Datasets Published by Austin and Chicago.

| | Austin | Chicago |
|---|---|---|
| Data Time | June 15 – October 15, 2019 | |
| Data Format | One observation represents one individual trip | |
| Number of Records | 2011k | 711k |
| Geographic Unit | Aggregated to 500ft hexagonal grid | Aggregated to census tract centroid |
| Key Attributes | Date, start/end time, duration, distance, OD hexagon coordinates | Date, start/end time, duration, distance, start centroid latitude/ longitude |

The data cleaning process in this study was as follows. First, the study was analytically meaningful by excluding trips that lasted less than the extreme outlier*in trip duration and trip distance, then we

excluded trips that lasted less than 60 seconds and limited the revelation time of the trip to 6-22 points to make the study data consistent with people's regular travel characteristics.
(extreme outliers are calculated by following: q1= quantile(0.25) q3= quantile(0.75) IQR=q3-q1
noutliers = df[((df>=(q1-1.5*IQR)) & (df<=(q3+2*IQR)))])

Due to the lack of user surveys, the actual demographics and socioeconomic status of e-scooter users in this study are uncertain, and this study refers to model specifications from previous studies. We obtained demographic information from the U.S. Census Bureau for each city, including population, age, gender, educational status, number of students at the census tract level, number of two people family households, number of nonfamily households, number of educated people, and number of employed people, in addition to the number of users in the two cities with different transit We also obtained the number of users of different transit modes in two cities. The effect of the total population size is overcome by calculating the ratio. The next section describes in detail how to present these data.

We also considered the impact of the built environment: accessibility as represented by distance to city bus stops and subway stations, and accessibility to potential users as represented by distance to universities, schools, restaurants, and urban land use diversity (the level of land use mix in the area).

In addition, we have seen a lot of concern and reports on the safety of scooter use in government reports and the public media, so we deliberately included statistics on traffic crashes in the study.

Table 3:
Descriptive Statistics of Dependent Variable (DV) and Independent Variables (IV) for Models

| Category | Unit | Austin | | | | Chicago | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | *Mean* | *SD* | *Min* | *Max* | *Mean* | *SD* | *Min* | *Max* |
| Average hourly trip (DV) | Count | 40.075 | 323.753 | 1 | 11801 | 173.304 | 540.149 | 1 | 9181 |
| Income level(IV) | Dollar | 41286.528 | 20028.624 | 3720 | 89455 | 35862.32 | 18707.663 | 9703 | 92340 |
| Median Age (IV) | Year | 32.869 | 4.945 | 20.3 | 50.5 | 36.433 | 3.855 | 28 | 46.2 |
| Gender ratio (IV) | Ratio | 1.272 | 0.331 | 0.787 | 3.037 | 1.135 | 0.319 | 0.325 | 2.693 |
| Residential Type (IV) | Ratio | 0.852 | 0.095 | 0.375 | 0.989 | 0.645 | 0.147 | 0.246 | 0.956 |
| Student Ratio (IV) | Ratio | 0.029 | 0.024 | 0.001 | 0.119 | 0.027 | 0.018 | 0 | 0.129 |
| High Education Ratio (IV) | Ratio | 0.068 | 0.036 | 0.004 | 0.161 | 0.071 | 0.031 | 0.007 | 0.196 |
| Employment Ratio(IV) | Ratio | 0.639 | 0.128 | 0.295 | 0.849 | 0.538 | 0.141 | 0.044 | 0.829 |
| Nondriving Commute(IV) | Ratio | 0.594 | 0.104 | 0.329 | 0.872 | 0.755 | 0.074 | 0.595 | 0.974 |
| Population density (IV) | *1000 per square mile | 7 | 7 | 0 | 27 | 0.15 | 6 | 0 | 50 |
| Land use diversity(IV) | Ratio | 0.074 | 0.128 | 0 | 0.556 | 0.209 | 0.112 | 0.05 | 1 |
| Traffic Crash (IV) | Count | 0.518 | 0.85 | 0 | 4 | 165.1 | 136.75 | 20 | 1101 |
| Trip Start Time (IV) | Hour of Day | 15.684 | 4.73 | 6 | 22 | 15.362 | 4.226 | 6 | 22 |
| Trip Duration (IV) | s | 617.789 | 311.93 | 64 | 1608 | 562.917 | 275.141 | 64 | 2083 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Distance to bus stops (IV) | Meter | 0.072 | 0.031 | 0.021 | 0.183 | 10.633 | 3.025 | 3.103 | 15.453 |
| Distance to universities (IV) | Meter | 0.193 | 0.063 | 0.066 | 0.378 | 36.976 | 9.073 | 5.975 | 56.793 |
| Distance to schools (IV) | Meter | 0.126 | 0.039 | 0.041 | 0.324 | 16.801 | 4.904 | 5.573 | 29.205 |
| Distance to restaurants (IV) | Meter | 0.083 | 0.032 | 0.031 | 0.209 | 15.424 | 5.873 | 3.49 | 34.346 |

*Notice: Here is the original data. When modeling, all the variables here are log-transformed and standard scaled to be modeled to get rid of the skew distribution and get a more accurate analysis.

### 3.3 Measure：

The dependent variable in this study is the total number of scooter trips per hour originating from each census tract, and the total number of trips per quarter is calculated using the following equation:

$$y = \sum_{Jun}^{Oct} n_{outflow\ per\ each\ hour}$$

Here $n_{outflow\ per\ each\ hour}$ represents the total trip count per each hour for every census tract in the whole season.

The independent variables for the census tract level included **8 SES(social economic status) variables**, **5 BE(building environment) variables, and 3 trip-related variables**. The population density was calculated by dividing the number of people in thousands by the area in square miles. Gender is represented by the proportion of males and females. Since electric scooters are most likely to be ridden by young people, we focus on indicators related to schools and students, including the number of students in school as a percentage of the total population, etc.

For this study, the proportion of the population with an educational background was chosen as an indicator of the educational status of a region. In previous studies, we have seen that some people choose to travel by e-scooter as an alternative mode of transportation to work, so we counted the number of employed people in the census tract as a percentage of the total population. We also measure the household structure of the area by counting the proportion of 2-person households and non-households to the total number of households. In addition, we also consider the effect of people's usual mode of transportation on their choice of e-scooter riding. We consider that people who are not originally used to driving are more likely to use scooters to travel, so we measure the willingness of people to use e-scooters when choosing their mode of travel by counting the proportion of the population that does not drive.

The 5 BE variables include accessibility represented by distance to city bus stops and distance to university, school, restaurants to reflect the urban land distribution of possible scooter demand. The distance to the bus stop, university, school, and restaurant, is calculated based on the Euclidean distance from the center of mass of the census tract to the location of the corresponding points of amenities published in the open street map.

Land use data can be found on the city's official website. However, it is difficult to make one city's land use classification code exactly match another city's land use classification, so we measured site complexity by calculating the number of land use types per census tract as a proportion of the overall zoning plan parcel count.

Regarding the traffic crash data, for statistical convenience, we used the corresponding crash data for the whole year of 2019 and aggregated it to the census tract to count the number of crashes. Similarly, we add the duration and the start time of the trip in the trip data to the regression equation to consider and examine the impact of people's scooter-related behavioral characteristics on the final scooter use.

### 3.4 Analysis:

The study first conducted a trip analysis to visualize the spatial and temporal distribution of e-scooter trips in the two cities and some related economic and demographic indicators. The possible determinants of scooter use by commuters were obtained by comparing the spatial and temporal distribution of e-scooter trips in hourly, daily, and weekly usage of the two cities and the locations where e-scooter use is concentrated.

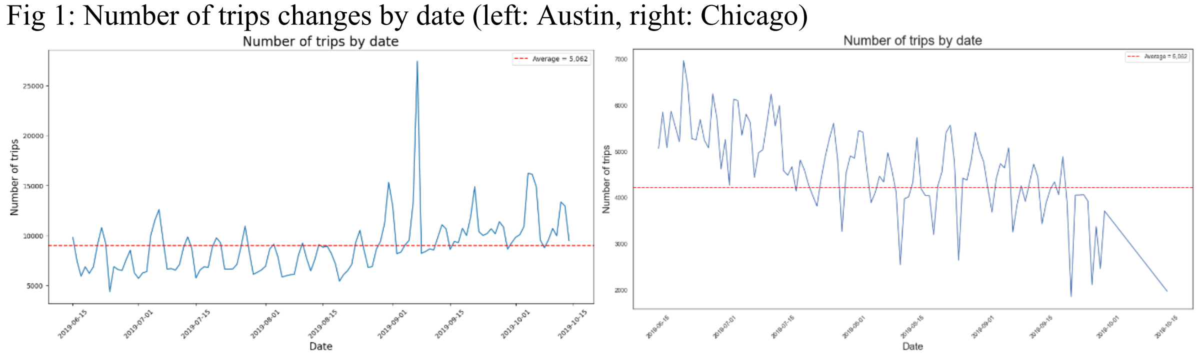
To examine the association between e-scooter usage and built environment, demographic characteristics, traffic behavior characteristics, etc., we use the random forest and multi-linear factor regression for model prediction and compare indicators in different cities by the rank of coefficient rate values.

In terms of model training method, we build the random forest and multi-linear regression models. By building a random forest model, we can get the general importance of different predictors by feature importance. The advantage of using multi-linear factor regression is that the model is highly interpretable and can be used to make cross-sectional comparisons of different factors by putting all relevant factors into the same model. We can determine the influence of predictors on e-scooter usage by the magnitude of the coefficient rate in the model and the positive and negative. In linear regression, to avoid the problem of difficult fitting due to scattered data and large variance, we performed standard scale and log transform in the data processing, which may make the direct interpretation of the coefficient rate more difficult. However, since we expect to make relative comparisons among variables without discussing absolute value differences in changes in independent variables due to unit-dependent variables, we believe that it is feasible to use this calculation.

## 4. Result:

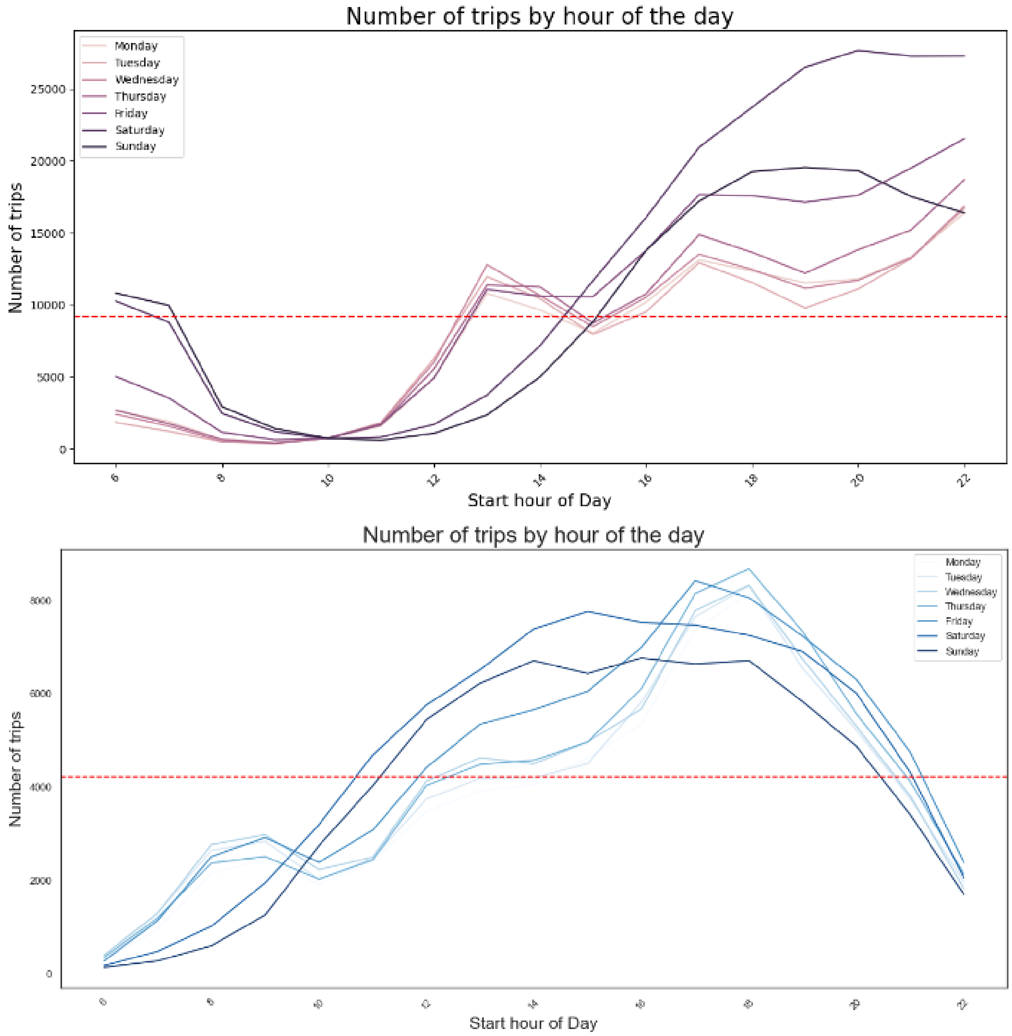
### 4.1 Comparison of e-scooter usage between two cities:

Fig 1: Number of trips changes by date (left: Austin, right: Chicago)




As shown in Figure 1, between June 15 and October 15, 2019, Austin's trip count basically remained above and below the median, and it is especially noteworthy that the trip count peaked around September 7, which may be related to the city's College GameDay event, during which scooter use was nearly five times the daily use. This indicates that most tourists and people associated with the event choose to use scooters to travel during this period. In Chicago, the use of scooters is not optimistic, and the number of trips continues to decline during the pilot period. In Chicago, people's use of scooters is divided into three phases: the peak period of use from June 15 to July 15, the stable period from July 15 to September 1, and the decline period from September 1 to October 15.

Fig 2: Number of trips changes by the hour of the day (left: Austin, right: Chicago)



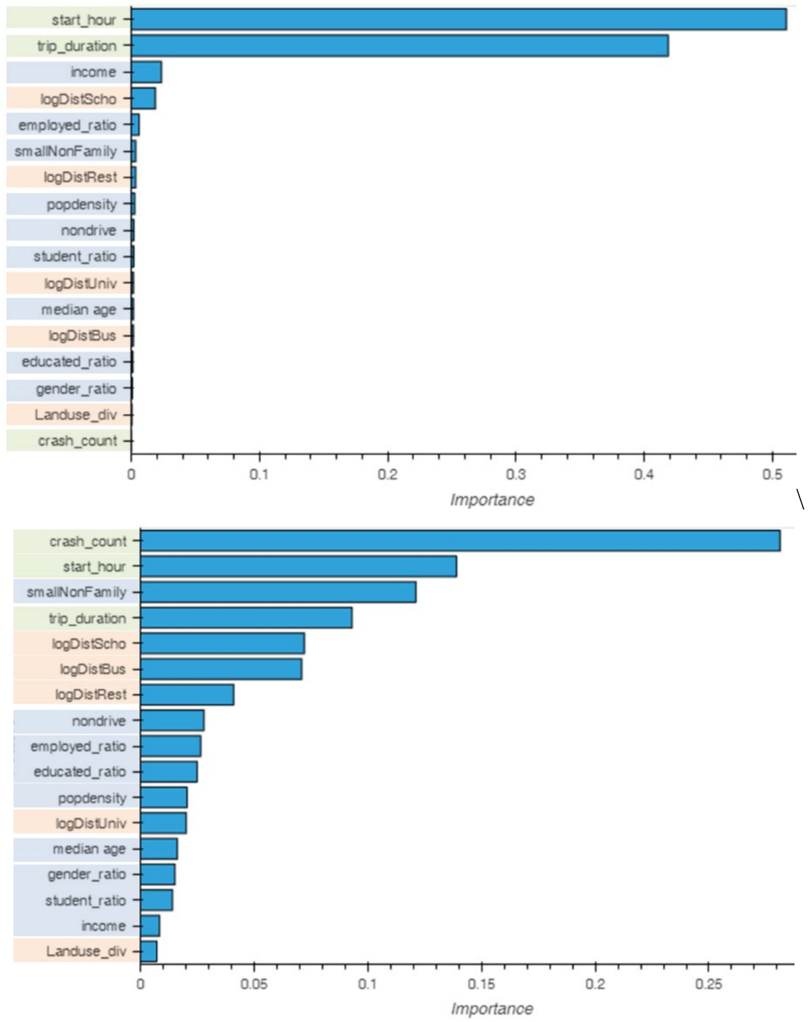

From the weekly scooter usage for both cities, people’s scooter usage patterns on weekends and weekdays are distinctly different. There are two peak periods of scooter usage on weekdays depending on the commuting time of local people, while weekend scooter usage is concentrated during the day, but there is no significant peak hours. Generally speaking, the peak period of scooter use is located between 14:00-21:00 every day.

Unlike Chicago, Austin scooter's weekend use is more than during weekdays, and the peak use of scooters in a day is nearly the same, while in Chicago scooter's hourly use fluctuates, but is mainly concentrated at around 18:00. This indicates that people in Chicago choose to use the scooter to commute in the evening, while people in Austin are likely to choose to use the scooter to commute in the evening and morning.

### 4.2 Model results and analysis:

#### 4.2.1 Random Forest result analysis:

Figure 3: Comparison of feature importance in models（top: Austin, down: Chicago）




In the modeling, we first used the random forest model. We conducted iteration by grid search and derived the final feature importance results from the best-performing model. From the results of the two models, we can see that for Austin, the most important predictors include: start hour, trip duration, income, distance to school, and employed ratio, and for Chicago, the most important predictor variables include: crash count, start hour, family type, trip duration, distance to school. Among all the predictors, people's use of scooters is most likely to be related to the length of the trip and travel time. It is worth noting that in Chicago, people are very concerned about whether it is safe to use a scooter in the place of travel, and the importance of this feature in Austin is not particularly important, which may be due to the better traffic environment in Austin than in Chicago, or the relatively uniform distribution of crashes, but this point does not prove the relationship between crash counts and people's scooter use. We need to further explore the coefficient rate of different factors in the OLS model. (We also tried to use the Chicago prediction model with the highest score to predict the trip count in Austin, but the model scored very badly, which means that the prediction of trip count depends on the environment of the city itself, and if in the future government wants to build a prediction model, it is better to regress the historical data of the city itself, and the cross-city can refer to the modeling method, but direct model migration prediction may not be meaningful)

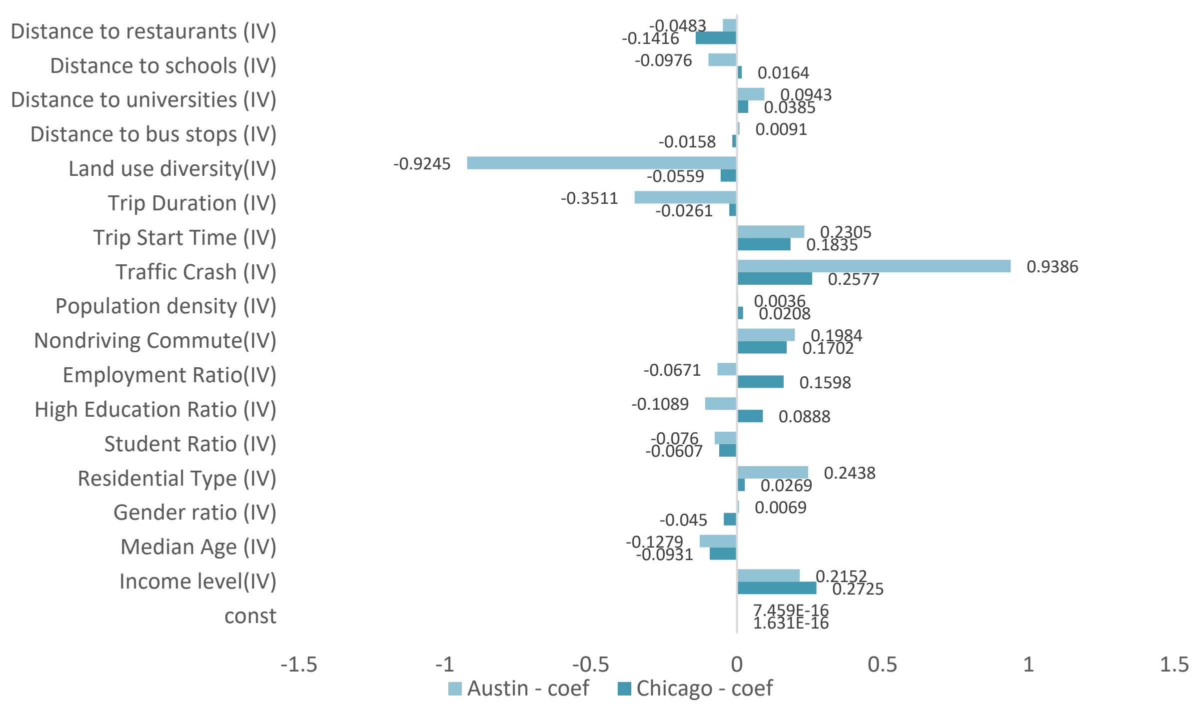
#### 4.2.2 OLS result analysis:

Table 6: OLS Model results

| | Chicago OLS Regression Results | | | Austin OLS Regression Results | | |
|---|---|---|---|---|---|---|
| | Chicago - coef | std err | P>\|t\| | Austin - coef | std err | P>\|t\| |
| const | 1.631E-16 | 0.015 | 1 | 7.459E-16 | 0.009 | 1 |
| Income level(IV) | 0.2725 | 0.032 | 0 | 0.2152 | 0.038 | 0 |
| Median Age (IV) | -0.0931 | 0.023 | 0 | -0.1279 | 0.026 | 0 |
| Gender ratio (IV) | -0.045 | 0.018 | 0.013 | 0.0069 | 0.011 | 0.527 |
| Residential Type (IV) | 0.0269 | 0.025 | 0.277 | 0.2438 | 0.017 | 0 |
| Student Ratio (IV) | -0.0607 | 0.017 | 0 | -0.076 | 0.012 | 0 |
| High Education Ratio (IV) | 0.0888 | 0.018 | 0 | -0.1089 | 0.017 | 0 |
| Employment Ratio(IV) | 0.1598 | 0.034 | 0 | -0.0671 | 0.036 | 0.062 |
| Nondriving Commute(IV) | 0.1702 | 0.025 | 0 | 0.1984 | 0.025 | 0 |
| Population density (IV) | 0.0208 | 0.019 | 0.278 | 0.0036 | 0.015 | 0.808 |
| Traffic Crash (IV) | 0.2577 | 0.021 | 0 | 0.9386 | 0.127 | 0 |
| Trip Start Time (IV) | 0.1835 | 0.015 | 0 | 0.2305 | 0.009 | 0 |
| Trip Duration (IV) | -0.0261 | 0.016 | 0.096 | -0.3511 | 0.009 | 0 |
| Land use diversity(IV) | -0.0559 | 0.021 | 0.008 | -0.9245 | 0.126 | 0 |
| Distance to bus stops (IV) | -0.0158 | 0.017 | 0.349 | 0.0091 | 0.013 | 0.483 |
| Distance to universities (IV) | 0.0385 | 0.019 | 0.045 | 0.0943 | 0.015 | 0 |
| Distance to schools (IV) | 0.0164 | 0.016 | 0.307 | -0.0976 | 0.016 | 0 |
| Distance to restaurants (IV) | -0.1416 | 0.021 | 0 | -0.0483 | 0.015 | 0.001 |

*Blue color represents 8 SES variables, orange color represents 5 BE variables, and green color represents 3 trip-related variables

Figure 4: Comparison of coefficients differences in models



From the results, we can see that like the feature importance, the coefficient rates of features in Austin are more polarized with a few features having a coefficient rate that is relatively far from zero, while the coefficient rate of features in Chicago is not that polarized.

In terms of positive and negative coefficient rates, most of the positive and negative correlations between features and trip count are the same between the two cities. Only a few indicators are opposite, including distance to the bus stop, employment ratio, educated ratio, and gender ratio. This may reflect the different characters of people using scooters in Chicago and Austin. In Chicago, most e-scooter users are likely to be educated and employed, while in Austin, e-scooter users are likely to be non-employed individuals without higher education. It is interesting to note that the gender ratio (male: female) has a completely different effect on the trip count in the two cities. And it is noteworthy that the coefficient rate is negative in Chicago, although in previous studies we can see that it is indeed mostly men who use e-scooters.

Among the features with the same positive and negative coefficient rate, we can also note that for land use diversity, trip duration, traffic crash, and residential type, their coefficients in Austin are significantly larger than those in Chicago, which may be related to the amount of data or the original data range. For example, the range of land use diversity in Austin is 0-0.556, and the range of land use diversity in Chicago is 0.05-1. Although the standard deviation is similar between the two, the difference in the size of the data range leads to a huge difference in the coefficient rate.

Figure 5: Comparison of coefficients rankings in OLS models

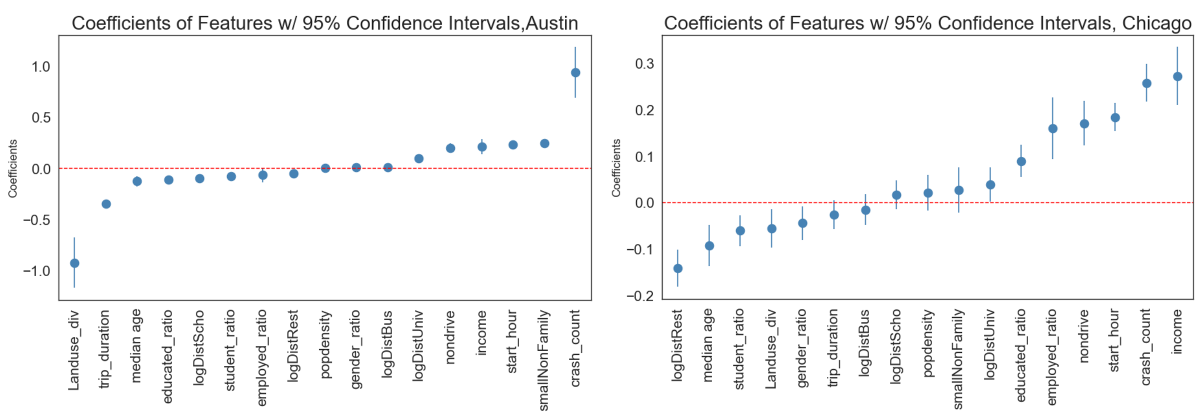



Table 7: Summary of predictors by coefficient rates ranking in Austin and Chicago.

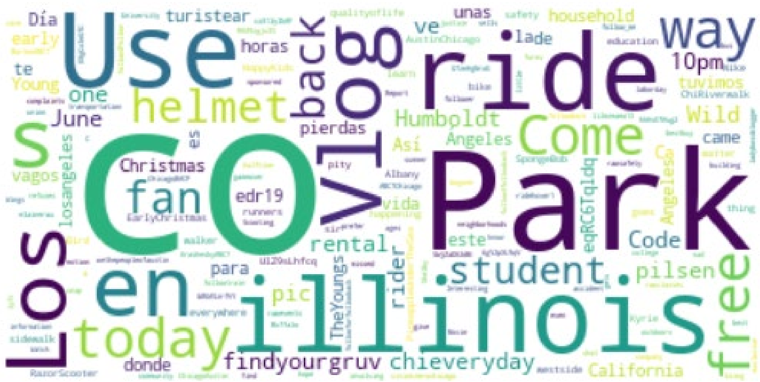
| | Positive | Negative |
|---|---|---|
| Austin | crash count, family type, start hour, income, and nondriving ratio | land use diversity, trip duration, the median age |
| Chicago | income, crash count, start hour, and nondriving ratio | distance to the restaurant, and median age |

In the coefficient rate ranking, it is interesting to note that crash count shows a positive correlation with the trip count in Austin and Chicago, which is contrary to our intuition, but we can also understand that the distribution points of scooters are mostly concentrated in the dense traffic areas where traffic crashes are likely to occur. The indicator here is not specific to the type of traffic crash, and it is more about the traffic access condition rather than the crash itself.

Another point is that the relationship between land use diversity and trip count in both city models shows a negative correlation, which indicates that people prefer to use e-scooter in a relatively single land use environment than a relatively complex and diverse urban environment, which is probably due to people's consideration for the safety of using the scooter.

### 4.3 Word cloud analysis.

Figure 5: Keywords of e-scooter usage in Chicago, 2019

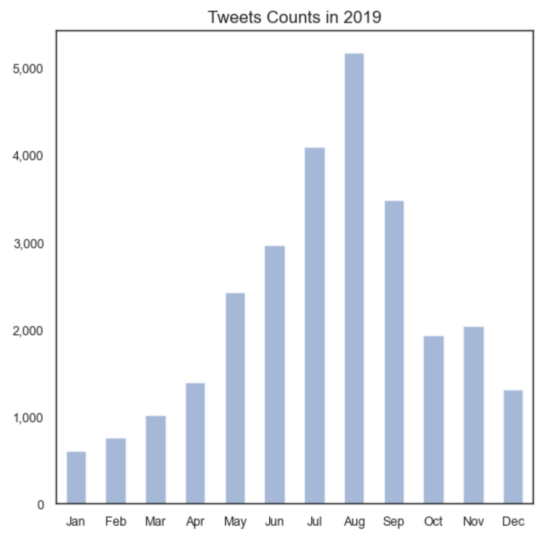



Through word cloud, we can see the main concern of people in this information lies in helmets, parking, use, and other topics, which coincides with the Chicago government's pilot test report released in 2020, in the actual e- scooter use, people are mostly concerned about the safety of scooter travel, and a lot of complaints about scooter parking, including parked scooters occupy bike parking spaces, sidewalk space and so on.

Figure 6: Keywords of e-scooter usage in Austin, 2019


Among Austin's Twitter keywords, we can see a recurring keyword: SXSW, which is an annual conglomeration of parallel film, interactive media, and music festivals and conferences organized jointly that take place in mid-March in Austin. Although this event was not covered in our study time, we can clearly see the relationship between scooter use and special city events through a word cloud. We can also see the discussion of scooter providers in the tweet buzzwords, including bird, Lyft, lime, etc., which may indicate that people are very concerned about the price of scooter trips. the keyword Everywhere may show the difficulty of parking scooters, etc.

Figure 7: Tweets Counts in 2019




In 2019, many cities launched e-scooter rental programs or shared mobility services that allow people to easily rent an e-scooter for a short period of time, making it easier for people to try out e-scooters and see if they are a good transportation option for them. It can be seen from May, there was a wave of tweeting on Twitter which reached a peak around August, and then gradually decreased, which may be related to people's reaction to the government's policy, but also may be due to the July-September climate is more suitable for scooter riding use.

## 5. Discussion & Conclusion

In this paper, we conduct random forest and multilinear regression prediction analysis for e-scooter quarterly trip counts in Chicago and Austin. The feature importance of predictors and the coefficient rate in the models were used to infer and analyze the impact of social economic factors, building environment factors, and trip-related features on e-scooter trip counts in the cities.

In both cases, trip-related features are more important for the impact of the e-scooter. However, in the context of many cities that just started the pilot project of scooter sharing, summarizing social economic factors and building environment factors is also important for the location of the city's e-scooter pilot project. From the regression analysis results of the two cities, we can see that neighborhoods that are younger, with higher incomes, and without previous driving habits will have more scooter usage.  Considering the building environment factors, more homogeneous neighborhoods will have more people using e-scooters.

In data-driven analysis, most accurate models are built in the context of analysis with relatively stable variables. However, in urban environments, planners often need to face new project setups and decisions without historical data. This highlights the importance of forecasting and learning from other cities. From the prediction results, the contribution of the factors fluctuates greatly as the urban context changes. However, a few factors are in a relatively stable contribution ranking, indicating that for the same topic, there are always a few determinants that are not affected by the city environment, such as the income and age structure of the population, etc. For this kind of factor, it would be more effective to use more sophisticated GWR algorithms and to survey more cities to infer such correlations.

Also, we can see that cities need to make appropriate policy adjustments based on their own existing scooter trip feedback. For example, from the Twitter word cloud, we can clearly see that people are concerned about scooter parking and cycling safety. These are difficult to anticipate in the initial implementation of the project. Besides, as urban infrastructure, it may be more important to make more people benefit than to simply meet the needs of those who need it. Whether people's behavior in using scooters will in turn contribute to the improvement of urban form is also something that planners should consider. A common finding is that areas with more bicycle use help shape the urban form for car and public transportation use (Bento et al., 2005, Giuliano and Dargay, 2006, Guerra et al., 2018). From our study, we can see that for safety reasons, people are more likely to choose scooter travel in areas with relatively homogeneous land use

As a final point, we consider the purpose of people using e-scooter trips. From the difference in the time period distribution of scooter trips in Chicago and Austin, we can speculate that the purpose of people using scooters in the two places is extremely different, and the people using scooter trips in Austin in the early morning are likely to be engaged in manual labor related. However, in Chicago, most of the scooter trips are concentrated in the afternoon to evening, and in the analysis of the 2020 scooter hotspots, we can see that most of the hotspots are concentrated in tourist areas, and the users of these trips are most likely to be people who travel or engaged in recreational activities in Chicago. Setting up corresponding incentives for different usage contexts and places of use, such as setting up scooters near scenic spots, bike-first streets, providing dockless parking, etc., will create a better and more convenient travel environment and encourage the use of e-scooters.

**References:**


[1] Zarif, R., Pankratz, D., Kelman, B., 2019. Small is beautiful: making micromobility work for citizens, cities, and service providers. Deloitte Insights. https://www2.deloitte.com/us/en/insights/focus/future-of-mobility/micro-mobility-is-the-future-of urbantransportation.html?id=us:2ps:3gl:confidence:eng:cons:42319:nonem:na:nhRV7UOl:1149484916:344865936403:b:Future_of_Mobility:Micromobility_BMM:nb.

[2] National Association of City Transportation Officials (NACTO), 2018. Guidelines for the regulation and management of shared active transportation (No. 1). https://nacto.org/wp-content/uploads/2018/07/NACTO-Shared-Active-Transportation-Guidelines.pdf.

[3] Gu, T., Kim, I., Currie, G., 2019. To be or not to be dock less: Empirical analysis of dock less bikeshare development in China. Transportation Research Part A: Policy and Practice119, 122–147. https://doi.org/10.1016/j.tra.2018.11.007

[4] Ewing, R., Cervero, R., 2010. Travel and the built environment: a meta-analysis. J. Am. Planning Assoc. 76 (3), 265–294. https://doi.org/10.1080/01944361003766766.

[5] Mooney, S.J., Hosford, K., Howe, B., Yan, A., Winters, M., Bassok, A., Hirsch, J.A., 2019. Freedom from the station: spatial equity in access to dockless bike share. J. Transp. Geogr. 74, 91–96. https://doi.org/10.1016/j.jtrangeo.2018.11.009.

[6] Xu, Y., Chen, D., Zhang, X., Tu, W., Chen, Y., Shen, Y., Ratti, C., 2019. Unravel the landscape and pulses of cycling activities from a dockless bike-sharing system. Comput. Environ. Urban Syst. 75, 184–203. https://doi.org/10.1016/j.compenvurbsys.2019.02.002.

[7] Mathew, J.K., Liu, M., Seeder, S., Li, H., Bullock, D.M., 2019. Analysis of E-Scooter trips and their temporal usage patterns. Inst. Transp. Eng. J. 89 (6), 44–49

[8] Bai, Shunhua, and Junfeng Jiao. “Dockless E-Scooter Usage Patterns and Urban Built Environments: A Comparison Study of Austin, TX, and Minneapolis, MN.” Travel Behaviour and Society 20 (July 2020): 264–72. https://doi.org/10.1016/j.tbs.2020.04.005.

[9] Jiao, Junfeng, and Shunhua Bai. “Understanding the Shared E-Scooter Travels in Austin, TX.” ISPRS International Journal of Geo-Information 9, no. 2 (February 24, 2020): 135. https://doi.org/10.3390/ijgi9020135.

[10] Yang, Hongtai, Jinghai Huo, Yongxing Bao, Xuan Li, Linchuan Yang, and Christopher R. Cherry. “Impact of E-Scooter Sharing on Bike Sharing in Chicago.” Transportation Research Part A: Policy and Practice 154 (December 2021): 23–36. https://doi.org/10.1016/j.tra.2021.09.012.

[11] Yang, Hongtai, Rong Zheng, Xuan Li, Jinghai Huo, Linchuan Yang, and Tong Zhu. “Nonlinear and Threshold Effects of the Built Environment on E-Scooter Sharing Ridership.” Journal of Transport Geography 104 (October 2022): 103453. https://doi.org/10.1016/j.jtrangeo.2022.103453.

[12] Zhu, Rui, Xiaohu Zhang, Dániel Kondor, Paolo Santi, and Carlo Ratti. “Understanding Spatio-Temporal Heterogeneity of Bike-Sharing and Scooter-Sharing Mobility.” Computers, Environment and Urban Systems 81 (May 2020): 101483. https://doi.org/10.1016/j.compenvurbsys.2020.101483.