A Work Project, presented as part of the requirements for the Award of a Master's degree in Management from the Nova School of Business and Economics

EVALUATING THE IMPACT OF BIKE-SHARING SERVICES ON UBER - EVIDENCE FROM NEW YORK CITY

DAVID AUTZ

Work project carried out under the supervision of:

Prof. Bernardo Melo Pimentel

17-12-2021

Abstract:

This paper analyses the impact of bike-sharing services on ride-sharing services by using trip data in New York City during 2014 and 2015. The study works with datasets of Uber and Citi Bike, the two largest operators in their respective sectors. Using multiple linear regressions, the results show that both services can be seen as substitutes on a trip-by-trip basis. However, when widening the observational period and looking at the complete customer journey, this relationship is no longer significant. Moreover, trip duration and customer types play an essential role when quantifying the impact.

Keywords:

Bike-sharing services, Ride-sharing services, Sharing economy, Peer-to-peer platforms, Uber, Citi Bike

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

The views expressed in this paper are solely the author's and do not necessarily reflect the views of Uber Technologies, Inc. David Autz was employed by Uber Technologies, Inc. for part of the time in which this paper was written.

Table of Contents

1 Introduction

The rise of the "sharing economy" has influenced the lives of billions of people (Yaraghi and Ravi 2017). Peer-to-peer platforms such as Airbnb and Uber allow consumers to share underutilized inventory on virtual marketplaces. The rapid growth of these platforms was enabled primarily through technological innovation, namely the widespread distribution of reliable internet connections and the development of smartphones that allow users to access these platforms conveniently from anywhere. Platforms operating in the sharing economy have been proven to be particularly effective in industries that show high demand volatility while traditionally having fixed-capacity supply. With the entry of companies like Airbnb or Uber, they have changed the industry's competitive landscape by building platforms that introduce flexible capacity (private cars in the case of Uber; private rooms in the case of Airbnb) to accommodate the volatile demand.

The present paper focuses on these platforms by providing information on the impact that associated industries have on peer-to-peer marketplaces. Specifically, this study aims to provide additional insights into the effects of ride-sharing services by estimating the impact of bike-sharing services on Uber in the short run. The empirical analysis focuses on New York City during 2014 and 2015. The underlying hypothesis throughout this paper is that some bike-sharing rides serve as a substitute for rides taken with Uber's platform. It is claimed that an increase in bike-sharing usage negatively affects the number of Uber rides.

The short-run impact is quantified using multiple linear regressions. The results show that Uber's car-sharing service and Citi Bikes can be considered substitutes when looking at a single trip. However, people might use different modes of transportation throughout the day. By widening the observation period from hourly to daily, Uber rides and rides with Citi Bikes are no longer substitutes. The analysis also reveals that user behaviour is different on weekdays and weekends. The substitutive relationship between both modes is stronger on weekends.

Further analysis shows that Citi Bike rides with lower trip durations have a lower impact on Uber ridership and users with an annual subscription for bike-sharing services have a higher impact.

Different modes of transportation offer different benefits and costs. In some cases, consumers combine different modes of transportation and in other cases they need to choose between modes. This paper expands the research on this topic by quantifying whether two specific modes are considered substitutes or complements. While literature on this research question exists it vastly focuses on the main competitors of Uber, namely taxi services and public transportation. The present study extends the current research by focusing on the relationship between Uber and a different mode of transportation that is getting increasing attention in large cities due to its sustainability and low cost for users: bike-sharing.

An analysis of bike-sharing services on Uber ridership is interesting for two main stakeholders, policymakers and companies within the mobility industry. As many local and federal governments are considering limiting Uber's operations within cities, they must have clear insights into the direct and indirect effects of these interferences. User behaviour on modal transportation is highly complex, and a ban of one mode might have severe consequences on other modes. In some cities, trips require the use of multiple modes of transportation to get to the destination. A ban on one mode of transportation could also reduce the usage of other modes. This paper aims to provide new insights into the complex user behaviour and give policymakers more quantifiable information to make strategic decisions. These insights are also valuable for organisations within the mobility industry. Private companies in this sector are consistently looking to expand their services to capitalize on the complete user journey. As the industry is highly competitive and profit margins are generally relatively low, economies of scale and network effects play a critical role. Ride-sharing platforms are especially aggressive in their actions. Throughout the years, Uber and its competitors have tried many different strategies to increase their market share. Investments in bike-sharing services have played a significant role. In 2018 Uber acquired a bike-sharing service called "JUMP bikes". After two years of operation, it divested this project. Lyft, one of the main competitors of Uber, acquired Citi Bikes, the largest bike-sharing service within the US. It is still the owner and responsible for its operations. For ride-sharing companies, it is crucial to understand whether bike-sharing ridership and car-sharing ridership have a substitutive or complemental relationship. In case of a substitutive relationship, companies might risk cannibalism of sales, whereas a complemental relationship enables them to capitalize on the whole user journey. Then, an investment into bike-sharing services might be an excellent opportunity to leverage network effects and to build up an ecosystem of transportation.

The present study consists of four sections. Section one provides an overview on existing literature in the area of ride-sharing and bike-sharing platforms. Section two provides necessary background information on Uber and Citi Bike, details on the data sets used and explanations for the applied methodology. Section three presents the main results that proof the hypothesis and provide additional insights. Section four concludes the identified results and acknowledges limitations of the study.

2 Literature Review

While Uber and its impact have been discussed widely by the media and policymakers, the discussions are almost exclusively centred around Uber's impact on the taxi industry. Research regarding Uber's effects on other modes of transportation is scarce. The most noticeable insights are provided by Jonathan Hall, Craig Palsson and Joseph Price (2018), who analyzed whether Uber is a substitute or complement for public transit (not including bike-sharing services). Employing a difference-in-difference model, they conclude that "Uber provides a complement for the average transit agency" (Hall, Pallson and Price 2018), as the platform increases public transport ridership by 5% after two years. They also show that Uber has a negative effect on bus ridership and a large complementary effect on rail ridership.

The impact of bike-sharing services on different modes of transportation is analyzed more closely. Campbell and Brakewood (2017) show that bike-sharing has a negative impact on bus ridership in New York City. Using a difference-in-difference design, they conclude that 1.000 bike-sharing stations along a bus route lead to a 2.42% reduction in bus trips. Multiple surveys support this result (Buck, et al. 2013) (Murphy and Usher 2015). Ma, Liu and Erdogan (2015) analyze the effect of bike-sharing on metrorail ridership in Washington DC. Using a linear regression analysis, they conclude that a 10% increase in annual bike-sharing ridership is associated with a 2.8% increase in daily metrorail ridership. Zhou, Wang and Li (2019) model the travel choices between taxis and bike-sharing services in Chicago using machine learning. Their results show that bike-sharing services remain a competitive mode of transportation for short travel distances of up to 6-8km, implying that trips with both modes are likely to be substitutes within this range.

Until now, the effect between Uber and bike-sharing services has not been analyzed thoroughly. Bakó, et al. (2020) assessed the impact of Uber on bike-sharing systems in Budapest. A market exit of Uber enabled them to quantify the impact using a difference-in-difference model. For users with subscriptions to bike-sharing services, the exit of Uber resulted in a 6.5% decrease in usage of bike-sharing services on weekdays. For ad-hoc users, the exit of Uber in the Budapest market resulted in a 23% increase in bike sharing usage on weekends. Overall, the exit of Uber resulted in a decrease in the usage of bike-sharing services. However, Péter Bucsky (2020) notes that the results are questionable as the dataset is too small to draw meaningful conclusions. In fact, the sample used in the analysis consists of six months, with only one month covering the timespan after Uber exited the market. Péter Bucsky also mentions that the market exit of Uber in Budapest was not forced by legislators but was rather a strategic decision made by Uber itself. While Uber left the market, other ride-sharing operators gained market share. Many former Uber driver-partners just migrated to different platforms. Therefore, using a difference-in-difference design to identify the impact of bike-sharing services on ride-sharing platforms is not suitable in this case.

3 Data and Methods

The analysis in this study uses data from various sources from April 2014 to June 2015, with a gap between September 2014 to December 2014. The gap is due to unavailable data from Uber. Observations were conducted on an hourly and daily basis resulting in 7967 observations for the hourly analysis and 330 observations for the daily analysis.

Data on over 17 million Uber trips are taken from kaggle.com, which obtained the data from the NYC Taxi & Limousine Commission (TLC) by submitting a Freedom of Information Law request in July 2015. For data on bike-sharing services, this analysis relies solely on Citi Bike, New York City's largest operator. The company publishes its system data directly on the website. Data on NYC yellow cab taxis are provided by TLC. Other sources for data are the Port Authority of New York and New Jersey and World Weather Online.[1]

3.1 Background Information on Uber

Uber was founded in 2009 as an app-based transportation service that allows users to book private or shared car rides in exchange for money. The company started operating in San Francisco, and ten years later, it has expanded into over 900 cities and provided 7 billion trips per year (Uber 2020). However, Uber's activities have repeatedly become a concern leading to

[1] For more detailed information on sources see the appendix.

some policymakers introducing strict regulations or even banning operations entirely. Key stakeholders in this debate are drivers, incumbent firms and governments.

Uber generally treats driver-partners as independent contractors. Only in a few countries the company is required to formally employ its drivers (Davidov 2016). This means its driver-partners are usually not entitled to minimum wage, overtime pay or the right to unionize (Fishman 2021). In many instances drivers have complained about low payouts for their work. Incumbent firms see Uber as an unfair competitor. Drivers of limousines and licensed taxis are subject to strict insurance and safety requirements. In many cities Uber avoids these regulations, and thus costs, by not being a formal taxi company and by not formally employing its driver-partners. Governments are also concerned that Uber increases congestion in city centres (Erhardt, et al. 2019) and the lack of safety requirements for Uber drivers. In London Uber was declined operating due to "[…] several breaches that placed passengers and their safety at risk" (TfL 2019). Prior it was discovered that Uber's driver accounts had been used by unauthorized drivers in over 14.000 trips (TfL 2019).

However, ride-sharing apps like Uber bring technological advancement to an inefficient industry. Standard taxi regulations lead to significant inefficiencies due to search frictions and misallocation. Prior to Uber, the taxi industry in large cities was often named as one of the primary examples used by economists to illustrate the inefficiency of governmental regulations (Cairns and Liston-Heyes 1994). In many cities, rates for taxis are fixed within a specific range. This enforcement of price stability is aimed at protecting customers, but is not always helpful. According to Buchholz (2019), a more flexible tariff pricing system for yellow cabs could provide up to 12% more welfare and 25% more trips in New York City. Fares booked through Uber's platform are calculated by a "Surge-Pricing" algorithm that provides a price based on current and local demand and supply. This variable pricing strategy can minimize inefficiencies in the taxi sector. Besides, Uber's platform enables consumers to pay cashless and order rides conveniently from their phones. Through all these benefits, Uber creates \$1.60 of consumer surplus for every dollar spent, resulting in an overall estimated consumer surplus of \$6.8 billion in the US in 2015 (Cohen, et al. 2016). On the driver side, it is estimated that Uber creates \$150 of driver labour surplus per week for each driver-partner. A large part of this surplus is made due to the driver's flexibility of working hours (Chen, et al. 2017). As drivers can log in and out of the platform anytime, they are able to schedule their working time to their convenience. Research by Uber shows that this flexibility is particularly important for drivers as many of them have other full- or part-time employment positions and use Uber to smooth fluctuations in their income (Hall and Krueger 2015). The platform also influences consumer behaviour that leads to various indirect benefits. For example, it has been shown that Uber reduces the amount of alcohol-related traffic fatalities (Anderson and Davis 2021). Moreover, Uber can boost tourism in cities (Park, Kim and Pan 2020). The platform provides tourists with a user-friendly mobility option that displays the price and route even for people not able to speak the local language, thereby avoiding scams and tourist exploitation.

3.2 Uber Business Model

Uber's business model can be described as a two-sided marketplace. On one side, there are providers of services (drivers), and on the other side, there are consumers of services (riders). Uber organizes and enables communication and interactions between both sides to exchange services and money. The vital force of this model are cross sided network effects making the platform more valuable for both sides when more people join. As the operator of this marketplace, Uber creates revenue by charging riders roughly 20% of the total trip fare. The costs of Uber are mainly driven by the development and maintenance of its platform, insurances for riders and drivers and credit card processing fees. Throughout the years, Uber has entered different industries and has built up similar marketplaces such as Uber Eats (food delivery) and Uber Freight (logistics).

3.3 Surge Pricing

A key factor to Uber's success is the introduction of a flexible pricing system. As demand is highly volatile and localized for passenger transportation, it is crucial to have a system in place that can meet this demand at any given time. Uber does not formally employ its driver-partners, so the company is not able to set working hours. Every driver can log in and out of the platform as they wish. To meet this localized and volatile demand Uber incentivizes its drivers to work during peak hours through higher payouts. In cases of high demand, prices for a ride may increase so that users who need a ride urgently can get one while users that are more flexible will wait or choose alternative modes of transportation. Uber calls this system "Surge Pricing", and it ensures that supply and demand are met at a price that maximizes utility for both sides.

As the price is a crucial factor for the amount of Uber rides taken, it needs to be included as an explanatory variable in the regression analysis. Uber itself does not publish detailed data on its surge pricing algorithm, but a study by Cohen, et al. (2016) includes a heat map showing the percentage of sessions with surge pricing by hour of the week (see Table 1). The table uses a sample of nearly 50 million UberX sessions in 2015 from Uber's four biggest US markets which include New York City. While the data provided by this graph is only providing the likelihood of surge pricing being enabled for the average hour per week, it is the best publicly available information. Still, it offers a good proxy for the price of an Uber ride.

3.4 Bike-Sharing in New York City

Citi Bike is the only official bike-share service licensed by New York City's Department of Transportation. It started operating in May 2013 in the districts of Manhattan and Brooklyn and later expanded into Queens. Since its early days, Citi Bike has been the largest bike-sharing system in the US, with 8 million recorded trips in 2014. The system is run through a public-private partnership agreement between New York City and Motivate (acquired by Lyft in 2018). It operates as a station-based bike-sharing service. Riders can unlock the bikes from one station and return the bike to any other station in the system. Citi Bike does not receive any subsidies from the city or government and is funded entirely through private capital, sponsorship agreements, and revenues from memberships.

During the period of this analysis, around 330 stations were in operation. While some stations were closed and new ones were established, the absolute number of stations remained relatively constant. At that time, stations were located in the city centre of New York City, namely Manhattan and the northern area of Brooklyn (see Figure 1 for a map of all stations).

3.5 Uber Trips and Citi Bike Trips in New York City

The total observational period is 11 months. During that time, 17.6 Mio individual Uber trips and 7.8 Mio Citi Bike trips were recorded. For the purpose of this study, these trips were aggregated on an hourly and daily basis. Figure 2 shows the weekly number of trips with both modes of transportation. In the first observational period (April 14 – August 14), more rides with Citi Bikes were taken than through Uber's platform. On average, for every Uber ride 1,25 Citi Bike rides were taken. The second observational period (January 15 – June 15) looks substantially different. More than four times more Uber rides were taken compared to Citi Bike rides. Looking at both periods, Uber demonstrates a steep incline in usage, whereas the use of bike-sharing services remains relatively constant on a year-on-year comparison. It is to be noted that there is a decline in bike-sharing services from January 15 to April 15, which is likely due to its seasonal weather conditions making bike riding an unpleasant mode of transportation. In that period, the average outside temperature in New York City was 1°C. Overall, the weekly usage of both modes of transportation is positively correlated with $r = 0{,}1632$. Taking the natural logarithm of both variables yields an even higher positive correlation coefficient of $r = 0{,}4128$.

Figure 3 illustrates the average number of rides taken during the week for both modes of transportation. The graph shows that both options are used at the same time. On weekdays there is the first peak in usage for Citi Bikes and Uber car rides during 8 am and a second higher peak around 6 pm. These peaks correspond largely to "Rush hour" in New York City, meaning both forms of transportation are used to get to work in the morning and get home after work. On Saturdays, the peak of Uber usage moves towards 11 pm. The peak usage of Citi Bikes on Saturdays is during the middle of the day (1 pm – 3 pm). On Sundays, both modes of transportation report the lowest ridership. Again, the peak for the usage of bike-sharing services is earlier than that of Uber. As ridership for both modes are substantially different on weekdays and weekends, it makes sense to perform an impact analysis while separating these two groups. Overall, both graphs are very similar in shape and show a high correlation of $r = 0{,}6165$ (correlation with natural logarithm is $r = 0{,}8412$).

3.6 Supplementary Data

Apart from data on Uber rides and Citi Bike rides, several other variables are used in the regression analysis to avoid an omitted variable bias and account for external effects. As taxicabs are widely regarded as one of the main competitors to Uber's offerings and have been proven to impact the ridership significantly (Chang 2017), the taxicabs are included as a supplementary variable. Yellow Cabs are the official taxicabs that require a medallion to operate and are licensed by the New York City Taxi and Limousine Commission.

Monthly data on airline passengers is obtained from the Port Authority of New York and New Jersey and serves as a proxy for tourism and business travel inflow. This variable is included with the assumption that tourism and business travel produce a significant amount of ridership. The more incoming recreational and business travel there is, the more rides there are through Uber's platform.

The regression analysis also covers a set of variables from World Weather Online. It is assumed that weather plays an important role when deciding which mode of transportation to use. In fact, temperature, precipitation and wind speed correlate with at least one mode of transportation. Figure 4 shows that ridership through Uber's platform decreases steadily with temperature rises, whereas Citi Bike ridership increases with temperature up until its peak at 25 °C. For temperatures exceeding this peak, the bike sharing system tends to be less favourable. Citi Bike ridership is also negatively correlated with precipitation. There is no visible correlation found between precipitation and Uber ridership.

3.7 Hypotheses

The overall hypothesis is that bike-sharing services and ride-sharing services are substitutes. An increase in bike-sharing usage will lead to a decrease in rides taken through platforms like Uber.

There are justifiable reasons to believe that this hypothesis might be true or false. Intuitively Uber and bike-sharing services are substitutes as both are within the mobility industry and satisfy the same underlying need: to move people from A to B. They are alternative modes of transportation and users might leave bike-sharing services for Uber rides and vice-versa.

However, there are also reasons to believe rides taken through Uber's platform and bike-sharing rides serve as complements. While Uber provides a relatively high-priced form of transportation, bike-sharing services are much cheaper to use. When used as part of a subscription plan, the marginal cost of bike-sharing services is even zero. Nevertheless, bike-sharing services that require their users to return the bike to a station limit their range of usage.

As stations are primarily located in city centres, bikes might be used there, and Uber rides might be used for transportation into suburbs and areas with no or few stations for bikes. Therefore, Uber rides and bike-sharing rides might be taken consecutively to reach the destination while saving money.

There is also the possibility that bike-sharing services and Uber rides are substitutes and complements depending on the time horizon. When looking at a single trip, both options might serve as a substitute to each other, but when looked at a longer time horizon (e.g., a day), they can serve as complements. Users might, for example, take a bike for a trip from home to the city centre during the day when it is sunny. On the way back, they might prefer to use a different form of transportation that suits their needs better as weather conditions have changed, they are too exhausted, they do not want to ride a bike during the night, or they are even under the influence of alcohol which prohibits them from riding a bike. The paper aims to bring new insights to this debate.

3.8 Methods

The impact of bike-sharing services on Uber ridership in New York City is quantified by applying a multiple linear regression model. All regressions use an ln-ln specification as it is assumed that the variables in these models are more likely to follow an exponential relationship rather than a linear relationship. Due to this specification, the results can be evaluated using percentage terms instead of unit terms ("A percentage change in X impacts Y by Z per cent."). All observations of variables requiring this ln specification need to be calculated with +1 as some variables have zeros in their dataset. Because $\ln(0)$ is not defined and $\ln(1)=0$, all observations have to be added by one unit not to lose any observations and to avoid a bias.

The impact is estimated using Equation 1 for the hourly analysis. Ln Uber Ridership is the dependent variable measuring the total amount of Uber rides taken within one hour. The explanatory variables are ln Citi Bike Ridership, ln Yellow Cab Ridership, Surge Pricing Possibility, ln Air Travel, Temperature, ln Windspeed and ln Precipitation. Day and Hour specific dummy variables are included as well. All variables are observed hourly except ln Air Travel, as this data is only available monthly.

Equation 2 follows a very similar structure to analyze the same impact on a daily basis. Notably, the variable for Surge Pricing Possibility is missing. This variable must be excluded to avoid problems of multicollinearity with the day-specific dummy variables.

The coefficient of interest in both cases is β which measures the elasticity. As stated previously, a negative β is interpreted as an indication that some bike-sharing rides serve as a substitute to car-sharing rides through Uber's platform. Consequently, it means that an increase in Citi Bike ridership will reduce Uber ridership. A positive β indicates a complementary relationship between both modes of transportation.

4 Results

4.1 Impact of Bike-Sharing Services on Uber Ridership

Table 2 reports the results of the regression analysis. The first three columns show the results using the hourly analysis. As formulated in the hypothesis, bike-sharing ridership is regarded as a substitute for Uber ridership. A 10% increase in Citi Bike ridership leads to a highly significant decrease of 0.51% in Uber ridership. The impact is significantly larger on weekends, leading to a 2.74% decrease in Uber ridership. During the week, no significant effect is recorded. Looking at the daily analysis, the results show substantial differences. The total impact of Citi Bike ridership on Uber ridership turns positive. However, the effect is not significant. Like in the hourly analysis, ridership between both modes has a more complementary relationship during the week. In fact, on weekdays, a 10% increase in Citi Bike ridership leads to a significant 3.61% increase in Uber ridership. Overall, this analysis provides two quantifiable insights:

(1) User regard Uber's car-sharing service and Citi Bikes as substitutes when looking at a single trip. However, they might use different modes of transportation throughout the day. By widening the observation period from hourly to daily, Uber rides and rides with Citi Bikes are no longer substitutes.

(2) User behaviour is different on weekdays and weekends. The substitutive relationship between both modes is stronger on weekends.

Looking at the other six explanatory variables, they mostly fit the expectations as well. Ln Yellow Cab Ridership is the only variable that records substantial differences in its coefficient when looking at the different models. In some cases, it has positive coefficients, and in some cases, it has negative coefficients. This high variance might be due to an omitted variable bias. As the variable is not particularly robust, it does not make much sense to interpret its results. However, the inclusion of this variable proves to be important as it increases the model fit considerably (measured with adjusted R^2).

The variable Surge Pricing Possibility has a positive coefficient in all cases. This result is expectable since surge pricing is only activated when demand outstrips supply. Naturally, that is more likely to happen during hours of peak demand, as Table 1 shows.

The coefficients of Ln Air Travel are positive in all six models, meaning an increase in tourism and business travel will positively affect Uber ridership. This result is also not surprising as these people likely do not have a private car available and like to stick to modes of transportation they are used to. Uber being a platform offering its services all over the world, provides a viable option to them.

Coefficients for weather-related variables perform mostly as expected. A decrease in temperature is associated with an increase in Uber ridership. This is logical as the climate in

New York City is relatively mild throughout the year. Lower temperatures will draw users away from modes of transportation that expose them to cold weather (like riding bikes, walking, etc.) while drawing them towards modes that offer more comfort. A similar observation is made when looking at the variable ln Windspeed. Again, the more wind there is, the more users tend to take Uber rides as they offer more comfort and shelter from bad weather conditions. When looking at precipitation in all but one case, the coefficients are positive. However, it is surprising that the magnitude of this coefficient is usually lower and less significant than other included weather-related variables. Temperature and wind speed tend to be more important factors when deciding for car-sharing services through Uber's platform.

4.2 Detailed Analysis: Trip Duration

After analyzing the overall impact, this section provides a more detailed look into the relationship between bike-sharing services and Uber ridership. Not all trips taken with Citi Bikes likely follow the substitutive relationship that was established above. As bike-sharing services are used for various reasons, some trips might have a stronger impact on Uber ridership than others. It can be expected that longer trip durations with Citi Bikes have a larger negative effect on Uber ridership. While bikes are usually used to cover short distances, car rides are quicker and more efficient for longer distances. Shorter bike rides will probably have a less pronounced effect on Uber ridership than longer rides. It could even be the case that for short bike rides, Citi Bikes and Uber rides are complements. As the bike-sharing service is exclusively offered in the heavily congested city centre of NYC, bikes might be used for short trips within the city centre. In contrast, Uber's car-sharing service might be used to get from the city centre to the suburbs.

To analyze this hypothesis, Equation 1 was used while filtering Citi Bike ridership for specific times. To ensure there are sufficient observations in each group, a period of three minutes was chosen. Figure 5 illustrates the results. In fact, Citi Bike rides with lower trip durations have a lower impact on Uber ridership. The coefficient is even positive for very short bike rides, meaning bike ridership and car-sharing ridership serve as complements. All other explanatory variables are very similar to the coefficients recorded in Table 2, indicating high robustness of the results.

4.3 Detailed Analysis: User Types

Citi Bike offers three different payment models for users: an annual subscription, a 3-day pass and a 24h-pass. The company thereby allows grouping its users into frequent users and occasional users. In the system data Citi Bike refers to users with an annual subscription as "Subscribers" and to users with 24h-passes or 3-day passes as "Customers".

When evaluating the impact of bike-sharing services on Uber, it is important to take a deeper look and differentiate between these two user types. Unlike Customers, Subscribers pay an annual membership which results in theoretically no marginal costs of taking a Citi Bike. Therefore, Uber rides will appear relatively more expensive to this user group. They are also more likely to be frequent users of this service and value Citi Bike's benefits, making them loyal users. Customers, on the other hand, are likely to consider other forms of transportation more. Their trade-off between Citi Bikes and car-sharing rides should appear to be less pronounced, yielding a coefficient that will be lower than that of Subscribers. Indeed, Table 3 proves these assumptions to be correct. Using Equation 1 while filtering for Subscribers and Customers, respectively, it shows that users with an annual subscription have a higher impact. A 10% increase in Citi Bike ridership is associated with a highly significant 0,64% reduction in Uber ridership for this user type. Surprisingly, the effect is less severe than one could expect. For the Customer user type, a 10% increase in Citi Bike ridership still leads to a highly significant 0,58% reduction in Uber ridership.

5 Conclusion

This paper uses a short-run analysis to quantify the impact of bike-sharing services on Uber ridership in New York City. The research is performed using a multiple linear regression model with system data provided by Uber and Citi Bikes. The results indicate that on a trip-by-trip basis, bike-sharing services and car-sharing rides through Uber's platform can be viewed as substitutes. When widening the observation period and looking at user behaviour throughout the day, it is shown that users may change between both modes of transportation. Bike-sharing services might be used for one trip to the destination, and Uber's car-sharing service might be used for a trip back home. Within an observation period of 24 hours, the substitutive relationship is no longer observable. The paper also points out that there are significant differences for different trip durations and user types. An increase in trip duration with bike-sharing services is associated with an increase in substitutive relationship towards Uber ridership. Users with a yearly subscription for bike-sharing services consider Uber's car-sharing service as a stronger substitute as users without these long-period subscriptions.

The reported findings are important primarily for the strategy development of car-sharing services and for local policymakers. For car-sharing services like Uber, Lyft, Bolt, etc., the results show that bike-sharing services should not automatically be seen as direct competitors. The relationship between these two modes is complex. The research in New York City indicates that a well-established bike-sharing service inside the city centre might increase ridership for their respective platforms. For policymakers, the analysis provides evidence that different modes of transportation within a city are connected. Not always will a ban or limitation of one service increase ridership of other modes of transportation.

The research for this project only uses publicly available information and data sets. This results in several limitations that this study faced as Uber only discloses minimal information. The dataset of Uber used in this research was published as it was legally required but only covers the very minimum necessary. Consequently, the analysis works with several approximations and restrictions.

The dataset has a four-month gap between September and December 2014 as no information on Uber ridership is available for that period. It is assumed that no extraordinary external effects occurred during that time that changed user behaviour substantially. Additionally, Uber does not provide any information on the pickup points or drop-off points, making it impossible to perform a more localized analysis. The study provides insights on the choice between different modes of transportation only on a city level. It is to be imagined that the location of Citi Bike stations plays an important role. If the next possibility of getting a Citi Bike requires a long walking distance, this might impact the trade-off between bike-sharing services and Uber rides. Also, the dataset does not provide any price-related data. This analysis refers to information provided by Cohen, et al. (2016), but the data is very basic for three reasons. First, it includes only information on the possibility of Surge Pricing being activated, not the actual level of Surge Pricing. Second, the data shows the weekly average, meaning it does not account for intertemporal changes. Third, the data includes the cities of Chicago, Los Angeles, and San Francisco besides New York. Nevertheless, the variable proves to be an excellent proxy for the average price of an Uber ride as it shows high significance in nearly all analyses.

The regression was performed using only six covariates to account for external shocks. There may be more unobservable variables correlated with both Uber ridership and Citi Bike ridership, which would cause reverse causality problems.

Lastly, this project provides a short-run analysis in New York City during 2014 and 2015. The regression accounts for the specific transportation situation in this city for that time frame. Bike-sharing services are likely to have a very different impact on other geographies. Differences in culture, economic situation, availability of other modes of transportation and other aspects will change the magnitude of the effect. A generalization of the finding would be inadequate.

However, it is not advised to look at the effects of bike-sharing services and Uber as a platform only through this lens. The impacts of these services are remarkably versatile. Uber affects many different sectors and provides opportunities to increase economic and social welfare on many frontiers. Many of these have not been investigated thoroughly and leave much room for further studies.

Appendix

A1 References

Anderson, Michael, and Lucas Davis. 2021. *Uber and Alcohol-related Traffic Fatalities.* Prod. National Bureau of Economic Research. July.

Bakó, Barna, Zombor Berezvai, Péter Isztin, and Enikö Vigh. 2020. "Does Uber affect bicycle-sharing usage? Evidence from a natural experiment in Budapest." *Transportation Research Part A* 290-302.

Buchholz, Nicholas. 2019. *Spatial Equilibrium, Search Frictions and Dynamic Efficiency in the Taxi Industry.* December 9.

Buck, Darren, Ralph Buehler, Patricia Happ, Bradley Rawls, Payton Chung, and Natalie Borecki. 2013. "Are Bikeshare Users Different from Regular Cyclists?" *Journal of the Transportation Research Board* (Transportation Research Board of the National Academies) (2387): 112-119.

Bucsky, Péter. 2020. "Does Uber affect bicycle-sharing usage? Evidence from a natural experiment in Budapest: A comment." *Transportation Research Part A.*

Cairns, Robert, and Catherine Liston-Heyes. 1994. *Competition and Regulation in the Taxi Industry.* September.

Campbell, Kayleigh, and Candace Brakewood. 2017. *Sharing Riders: How Bikesharing Impacts Bus Ridership in New York City.* May 5.

Chang, Jung-Hao. 2017. "The Economic Effects of Uber on Taxi Drivers in Taiwan." *Journal of Competition Law & Economics* 1-26.

Chen, Keith, Judith Chevalier, Peter Rossi, and Emily Oehlsen. 2017. *The Value of Flexible Work: Evidence from Uber Drivers.* National Bureau of Economic Research. Cambridge, June.

Cohen, Peter, Robert Hahn, Jonathan Hall, Steven Levitt, and Robert Metcalfe. 2016. *Using Big Data to Estimate Consumer Surplus: The Case of Uber.* Cambridge: National Bureau of Economic Research, September.

Davidov, Guy. 2016. “The Status of Uber Drivers: A Purposive Approach.” *Hebrew University of Jerusalem Legal Studies Research Paper.*

Erhardt, Gregory, Sneha Roy, Drew Cooper, Bhargava Sana, Mei Chen, and Joe Castiglione. 2019. "Do transportation network companies decrease or increase congestion?" *Science Advances* (American Association for the Advancement of Science).

Fishman, Stephen. 2021. *Uber Drivers are Contractors Not Employees According to the NLRB.* Accessed August 27, 2021. https://www.nolo.com/legal-encyclopedia/uber-drivers-are-contractors-not-employees-according-to-the-nlrb.html.

Hall, Jonathan, and Alan Krueger. 2015. *An Analysis of the Labor Market for Uber’s Driver-Partners in the United States.* January 22.

Hall, Jonathan, Craig Pallson, and Joseph Price. 2018. *Is Uber a Substitute or Complement for Public Transit.* October 16.

Ma, Ting, Chao Liu, and Sevgi Erdogan. 2015. "Bicycle Sharing and Public Transit - Does Capital Bikeshare affect Metrorail ridership in washington, D.C.?" *Journal of the Transportation Research Board* (Transportation Research Board) (2534): 1-9.

Murphy, Enda, and Joe Usher. 2015. "The Role of Bicycle-sharing in the City: Analysis of the Irish Experience." *International Journal of Sustainable Transportation* (Taylor & Francis Group, LLC) 116-125.

Park, So Young, JunYun Kim, and Bing Pan. 2020. "The Influence of Uber on the Tourism Industry in Sub-Saharan Africa." *Journal of Travel Research.*

TfL. 2019. *Transport for London.* November 25. Accessed 08 29, 2021. https://tfl.gov.uk/info-for/media/press-releases/2019/november/uber-london-limited-found-to-be-not-fit-and-proper-to-hold-a-private-hire-operator-licence.

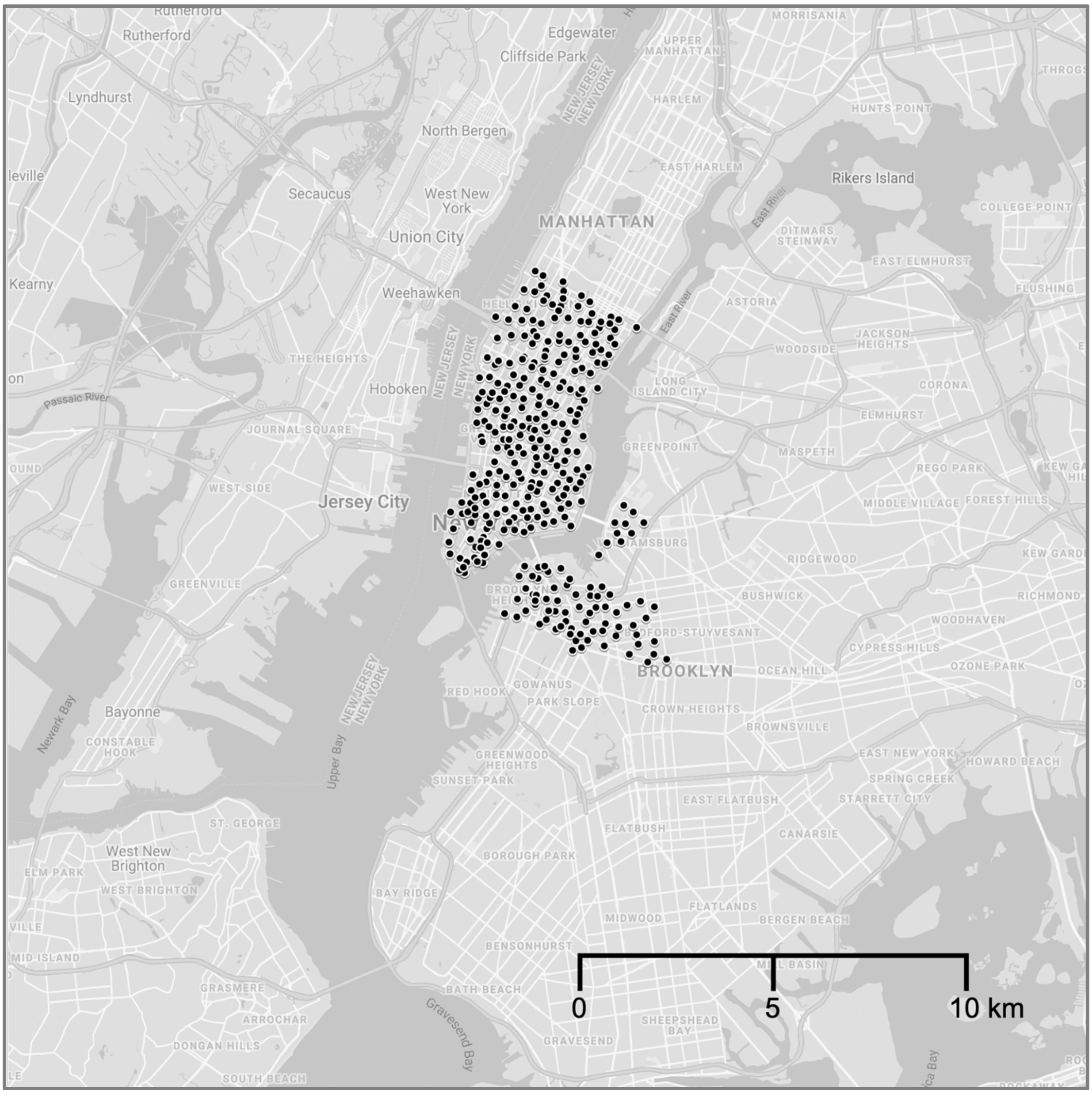
Uber. 2020. "2020 Investor Presentation." *Uber Investor Relations.* Accessed July 29, 2021. https://s23.q4cdn.com/407969754/files/doc_financials/2019/q4/InvestorPresentation_2020_Feb13.pdf.

Yaraghi, Niam, and Shamika Ravi. 2017. “The Current and Future State of the Sharing Economy.” *Brookings India IMPACT Series* (032017): 1-33.

Zhou, Xiaolu, Mingshu Wang, and Dongying Li. 2019. "Bike-sharing or taxi? Modeling the choices of travel mode in Chicago using machine learning." *Journal of Transport Geography* (79).

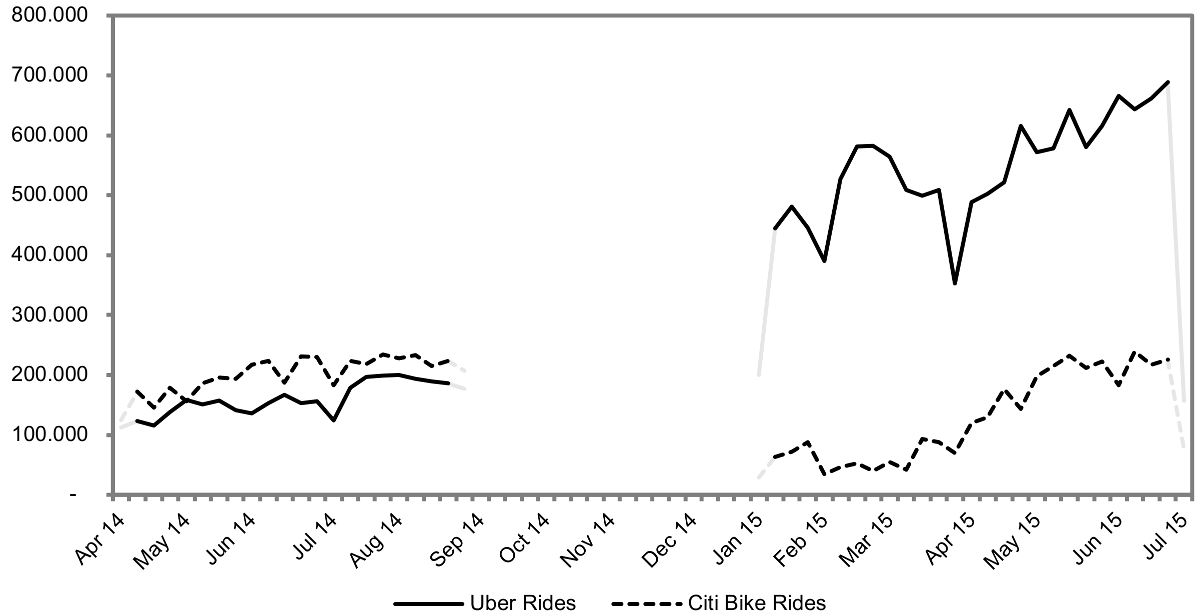
A2 Figures

Figure 1 - Location of Citi Bike Stations in New York City



Notes: The map shows all active Citi Bike stations as of June 2015. An interactive map is available under https://bit.ly/3FVznW3.

Figure 2 - Uber and Citi Bike Rides per Week



Notes: The light grey areas are not representative as they do not show data for a whole week due to observation starts and endings that are not at the beginning of a week. The drops for both modes around February 15 and April 15 are due to eliminated observations from both samples as they showed inconsistent data. The eliminated data are on 26.01.2015, 27.01.2015, 27.03.2015, 28.03.2015.

Figure 3 - Average Ridership throughout the Week

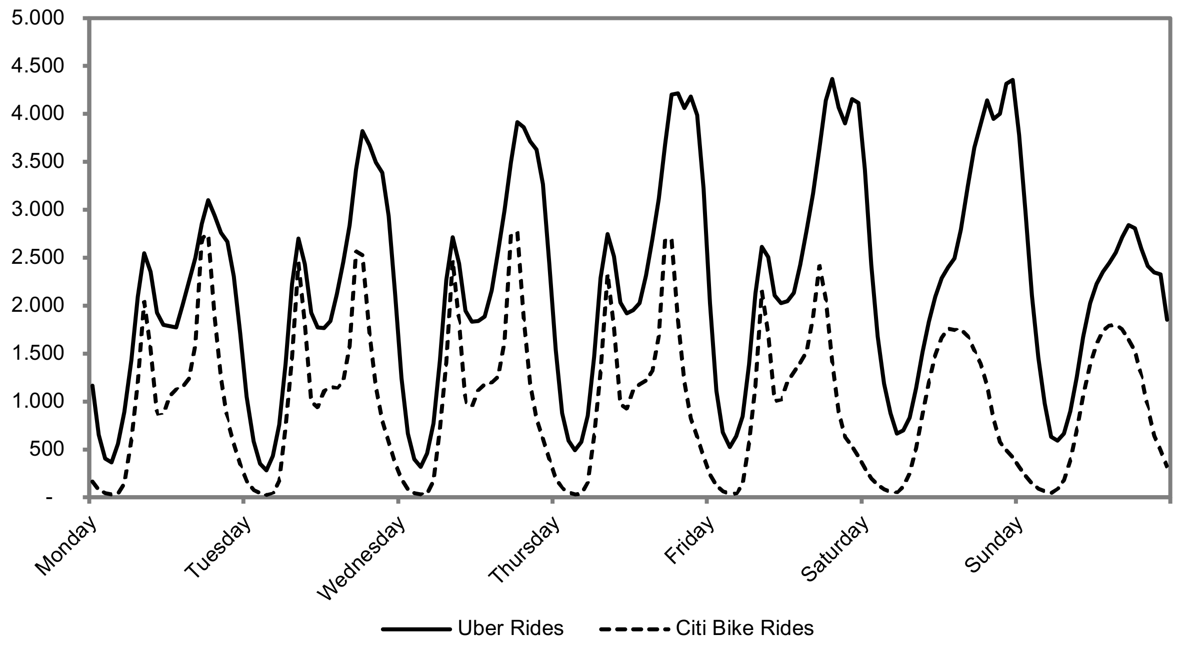
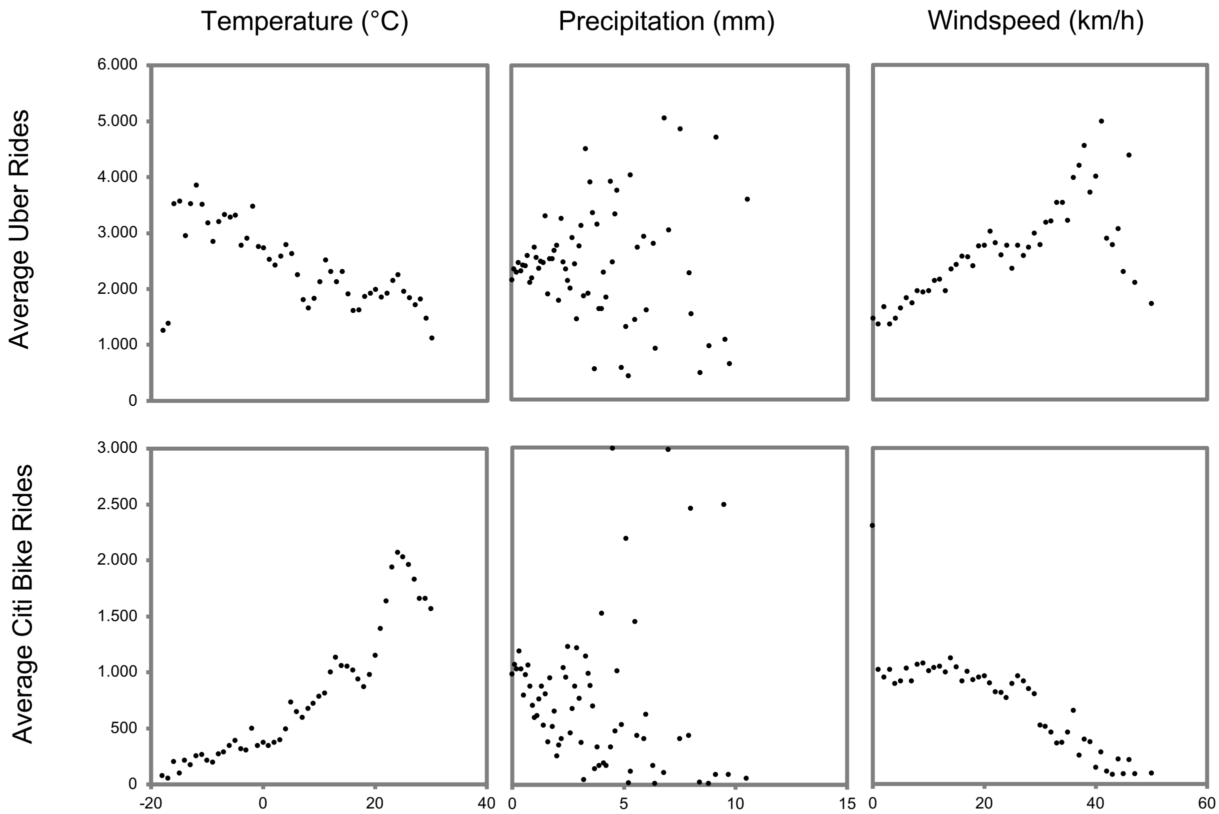
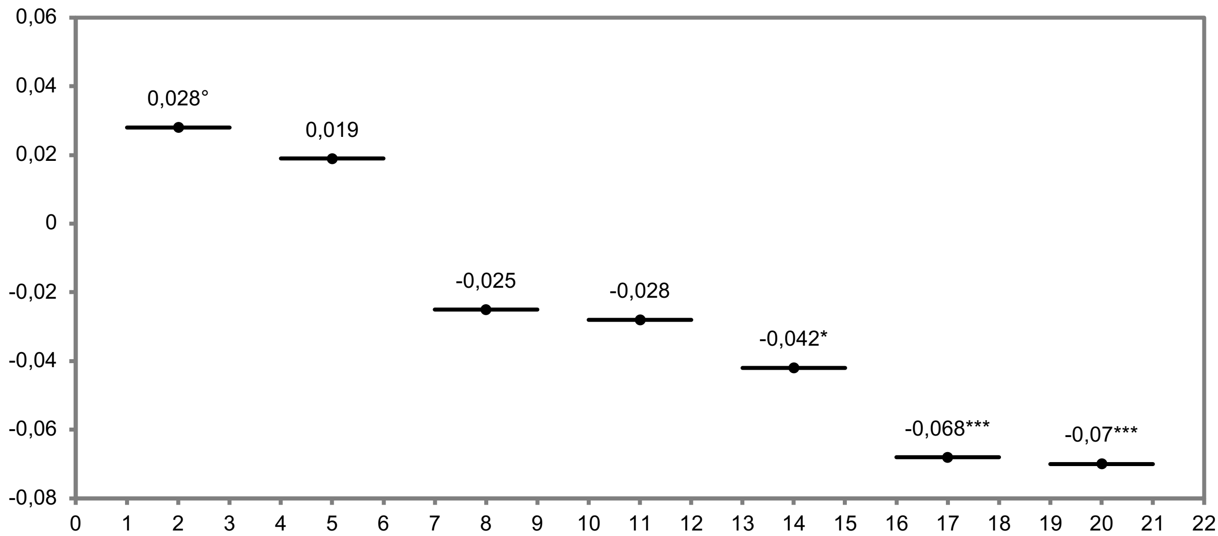


Figure 4 - Weather Data



Notes: For visualization purposes, each point shows the average number of Uber rides or Citi Bike rides, respectively, in relation to the other variable.

Figure 5 - Beta for Different Trip Durations (min)



° $p < 0.1$ * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Dependent variable: ln Uber Ridership

Other variables are excluded for visual purposes. For complete model see online appendix.

Notes: The figure illustrates the coefficient of ln Citi Bike Trips using Equation 1. Other coefficients are similar to those recorded in Table 2.

A3 Tables and Equations

Table 1 - Surge Pricing

	Mon	Tue	Wed	Thur	Fri	Sat	Sun
12 AM	8,78%	10,67%	9,73%	10,20%	11,14%	18,71%	38,55%
1 AM	7,84%	10,20%	10,20%	8,31%	11,14%	22,96%	36,66%
2 AM	7,84%	9,25%	9,25%	8,31%	11,62%	24,38%	37,61%
3 AM	8,31%	8,31%	7,36%	6,89%	9,25%	15,87%	24,38%
4 AM	17,76%	11,14%	7,36%	9,25%	10,20%	11,14%	17,76%
5 AM	23,43%	10,67%	7,84%	8,78%	11,62%	11,62%	13,98%
6 AM	25,32%	11,62%	9,25%	10,20%	14,92%	12,09%	16,34%
7 AM	30,99%	19,65%	18,71%	22,01%	25,79%	11,62%	12,56%
8 AM	36,19%	30,52%	29,57%	32,88%	32,41%	15,40%	13,98%
9 AM	21,54%	17,29%	17,29%	16,81%	18,71%	17,29%	18,71%
10 AM	8,78%	6,89%	7,36%	6,89%	8,78%	18,23%	25,32%
11 AM	8,31%	5,00%	4,53%	5,95%	10,67%	20,60%	28,16%
12 PM	10,67%	6,42%	6,89%	9,25%	21,54%	25,32%	26,27%
1 PM	8,78%	5,95%	5,95%	9,25%	26,27%	23,90%	23,90%
2 PM	10,20%	7,84%	7,36%	11,14%	26,27%	20,60%	22,96%
3 PM	15,40%	10,20%	9,73%	15,87%	26,27%	19,18%	20,60%
4 PM	14,92%	9,73%	10,67%	17,29%	24,85%	22,96%	23,90%
5 PM	19,65%	17,29%	18,23%	27,68%	29,57%	30,99%	27,21%
6 PM	20,60%	26,74%	23,90%	32,88%	34,77%	37,61%	22,96%
7 PM	15,87%	18,23%	16,81%	21,54%	32,41%	38,55%	16,34%
8 PM	11,62%	13,03%	12,09%	13,03%	16,34%	24,38%	13,03%
9 PM	16,34%	18,23%	15,40%	16,81%	10,20%	20,60%	15,87%
10 PM	22,96%	25,79%	24,85%	29,57%	16,81%	38,08%	20,60%
11 PM	14,92%	15,40%	17,29%	22,49%	23,43%	44,70%	13,98%

Notes: Surge Pricing is less dependent on demand than intuitively assumable. This is due to expected demand increases that are anticipated by drivers. Drivers are more likely to work during peak demand times, such as around 5-6 pm on weekdays, as it gives them a lot of opportunities to provide rides. Surge Pricing is more likely to happen when demand increase is unanticipated or when demand is high during hours that drivers do not like to work (e.g., Saturdays at 11 pm). For more information see Cohen, et al. (2016).

Equation 1 - Hourly Regression Formula

$$\ln \mathit{Uber\ Ridership} = \beta \times \ln \mathit{Citi\ Bike\ Ridership} + \gamma_1 \times \ln \mathit{Yellow\ Cab\ Ridership} + \gamma_2 \times \text{Surge} + \gamma_3 \times \ln \mathit{Air\ Travel} + \gamma_4 \times \mathit{Temperature} + \gamma_5 \times \ln \mathit{Windspeed} + \gamma_6 \times \ln \mathit{Precipitation} + \sum_{x=1}^{6} \delta_x \times \mathit{Day}_x + \sum_{y=1}^{23} \varepsilon_y \times \mathit{Hour}_y$$

Equation 2 - Daily Regression Formula

$$\ln \mathit{Uber\ Ridership} = \beta \times \ln \mathit{Citi\ Bike\ Ridership} + \gamma_1 \times \ln \mathit{Yellow\ Cab\ Ridership} + \gamma_2 \times \ln \mathit{Air\ Travel} + \gamma_3 \times \mathit{Temperature} + \gamma_4 \times \ln \mathit{Windspeed} + \gamma_5 \times \ln \mathit{Precipitation} + \sum_{x=1}^{6} \delta_x \times \mathit{Day}_x$$

Table 2 - Regression Results General Impact

	Ln Uber Ridership (hourly)	Ln Uber Ridership (hourly)	Ln Uber Ridership (hourly)	Ln Uber Ridership (daily)	Ln Uber Ridership (daily)	Ln Uber Ridership (daily)
Ln Citi Bike Ridership Total	-0.051** (-2.739)			0.058 (0.458)		
Ln Citi Bike Ridership Weekdays only		-0.013 (-0.560)			0.361** (2.681)	
Ln Citi Bike Ridership Weekends only			-0.274*** (-6.366)			-0.319 (-1.059)
Ln Yellow Cab Ridership	0.570*** (18.347)	0.020 (0.388)	-0.204 (-1.706)	-2.231*** (-5.420)	-3.498*** (-7.413)	-0.796 (-0.909)
Surge Pricing Possibility	2.221*** (15.384)	0.297 (1.288)	2.672*** (8.849)			
Ln Air Travel	2.018*** (13.896)	2.214*** (13.366)	1.995*** (7.379)	1.733** (2.770)	1.498* (2.259)	2.361 (1.743)
Temperature	-0.038*** (-20.291)	-0.039*** (-18.193)	-0.033*** (-8.076)	-0.044*** (-4.584)	-0.052*** (-5.390)	-0.036 (-1.407)
Ln Windspeed	0.198*** (11.587)	0.219*** (11.403)	0.138*** (4.216)	0.345*** (3.309)	0.372*** (3.486)	0.328 (1.355)
Ln Precipitation	0.089** (2.935)	-0.002 (-0.059)	0.167* (2.356)	0.051 (1.402)	0.076* (2.046)	0.059 (0.695)
N	7,967	5,685	2,282	330	235	95
Adjusted R^2	0.533	0.596	0.473	0.300	0.388	0.245

* $p < 0.05$
** $p < 0.01$
*** $p < 0.001$
Dependent variable: ln Uber Ridership
t-values are shown in parentheses
Method used for missing data: Last Observation Continued Forward (LOCF)
Dummy variables for hour and day specifics are excluded for visual purposes. For complete model see online appendix.

Note: Surge Pricing must be excluded in the daily analysis to avoid problems with multicollinearity.

Table 3 - Regression Results Impact on User Types

	Ln Uber Ridership (hourly)	Ln Uber Ridership (hourly)
Ln Citi Bike Ridership Subscriptions only	-0.064*** (-3.308)	
Ln Citi Bike Ridership Customers only		-0.058*** (-5.464)
Ln Yellow Cab Ridership	0.584*** (18.379)	0.542*** (20.985)
Surge Pricing Possibility	2.235*** (15.486)	2.198*** (15.452)
Ln Air Travel	2.018*** (13.915)	2.139*** (14.528)
Temperature	-0.037*** (-20.384)	-0.035*** (-18.061)
Ln Windspeed	0.198*** (11.577)	0.204*** (11.882)
Ln Precipitation	0.080** (2.672)	0.071* (2.547)
N	7967	7967
Adjusted R^2	0.533	0.534

* $p < 0.05$
** $p < 0.01$
*** $p < 0.001$
Dependent variable: ln Uber Ridership
t-values are shown in parentheses
Method used for missing data: Last Observation Continued Forward (LOCF)
Dummy variables for hour and day specifics are excluded for visual purposes. For complete model see online appendix.

A4 Data and Methodology

All datasets are available to download under:
https://www.icloud.com/iclouddrive/0x-scFWIXNo24JS32WmlDhf8g#Bike-Sharing_and_Uber_-_Online_Appendix
The directory of raw datasets is shown at the end of this appendix.

Data Sources

Data	Publisher	Link
Uber System Data	kaggle.com	https://www.kaggle.com/fivethirtyeight/uber-pickups-in-new-york-city
Citi Bike System Data	Citi Bike	https://ride.citibikenyc.com/system-data
Yellow Cab System Data	NYC Taxi and Limousine Commission	https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page
Surge Pricing Possibility	National Bureau of Economic Research	https://www.nber.org/system/files/working_papers/w22627/w22627.pdf
Air Travel Data	The Port Authority of New York and New Jersey	https://data.ny.gov/Transportation/Air-Passenger-Traffic-per-Month-Port-Authority-of-/8pkr-4b7t
Weather Data	World Weather Online	https://www.worldweatheronline.com/hwd/hfw.aspx

Notes on Methodology

The empirical analysis uses several datasets. The raw data includes information on 17,776,670 Uber trips, 7,786,011 Citi Bike rides and 14,6107,173 Yellow Cab rides. The data on Air Travel uses information on passenger arrivals of the airports of John F. Kennedy International Airport, Newark Liberty International Airport, LaGuardia Airport, Stewart International Airport, Atlantic City International Airport. To combine the datasets and compile them to fit the needs of this study, the database management tool SQLiteStudio v3.2.1 and Excel were used. The SQL-code can be found under Sql_Code_HourlyAnalysis and Sql_Code_DailyAnalysis, and the excel-file can be found under Compiled_Data.xlsx.
After compiling the dataset was thoroughly cleaned and tested for mistakes and missing values. The periods of 26.01.2015 17:00 - 27.01.2015 17:00 and 27.03.2015 21:00 - 28.03.2015 21:00 were excluded as there were missing data points on Uber and Citi Bike trips. Consequently, in the daily regression analysis, 26.01.2015, 27.01.2015, 27.03.2015, 28.03.2015 were excluded. The period of 08.03.2015 02:00 - 08.03.2015 03:00 was excluded due to the switch of wintertime to summertime. The Citi Bike dataset excludes any trips under 60 seconds in length as these are potentially false starts or users trying to re-dock their bike to ensure it is locked. The variable measuring surge pricing probability relies on data taken from Cohen, et al. (2016).

The original table illustrates this variable on a colour scale. Explicit values are not given but can be estimated through the legend. To get numerical data, the colour scale was first transformed to greyscale, and the brightness of the fields was measured and approximated through linear trendlines. For more details on this process see the file Surge_Pricing.pdf in the online appendix.
The variable "Temperature" is not transformed into ln as the observation period includes positive, negative temperatures as well as several observations with 0 °C. Taking the ln of this variable would eliminate certain values and consequently would lead to a biased result.
The regression analysis was performed using RStudio version 1.0.136. The code and the regression results can be found in the online data appendix under R_Input_Output.

Raw Data Directory

File name	Data set	Used for
RD_1.1	Uber 2014-04	Uber_Count
RD_1.2	Uber 2014-05	
RD_1.3	Uber 2014-06	
RD_1.4	Uber 2014-07	
RD_1.5	Uber 2014-08	
RD_1.6	Uber 2015-01 to 2015-06	
RD_2.1	Citi Bike 2014-04	CitiBike_Count
RD_2.2	Citi Bike 2014-05	
RD_2.3	Citi Bike 2014-06	
RD_2.4	Citi Bike 2014-07	
RD_2.5	Citi Bike 2014-08	
RD_2.6	Citi Bike 2015-01	
RD_2.7	Citi Bike 2015-02	
RD_2.8	Citi Bike 2015-03	
RD_2.9	Citi Bike 2015-04	
RD_2.10	Citi Bike 2015-05	
RD_2.11	Citi Bike 2015-06	
RD_3.1	Yellow Cab 2014-04	YellowCab_Count
RD_3.2	Yellow Cab 2014-05	
RD_3.3	Yellow Cab 2014-06	
RD_3.4	Yellow Cab 2014-07	
RD_3.5	Yellow Cab 2014-08	
RD_3.6	Yellow Cab 2015-01	
RD_3.7	Yellow Cab 2015-02	
RD_3.8	Yellow Cab 2015-03	
RD_3.9	Yellow Cab 2015-04	
RD_3.10	Yellow Cab 2015-05	
RD_3.11	Yellow Cab 2015-06	
RD_4.1	Air Travel 2014 to 2015	AirTravel_Count
RD_5.1	Weather Data Per Hour	Temperature, Windspeed, Precipitation
RD_5.2	Weather Data Per Day	