CHANGES IN THE RIDE-HAILING INDUSTRY DUE TO THE COVID-19 OUTBREAK

by

Mariia Shkulepova

A thesis submitted in partial fulfillment of the requirements
for the degree of

MA in Business and Financial Economics

Kyiv School of Economics

2021

Thesis Supervisor: Volodymyr Vakhitov

Approved by
Head of the KSE Defense Committee, Professor [Type surname, name]

Date _____

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

TNP Transportation Network Provider

US United States

COVID-19 Coronavirus Disease of 2019

CHAPTER 1. INTRODUCTION

The transportation market has changed a lot over the past decade, given the rapid scientific development, urbanization, road construction, and many other factors. Types and means of transportation are changing. No one can imagine that adults will move around the city on scooters that can be rented. At the same time, some of the standard modes of transportation are changing their appearance and companies are establishing communication with the user through a mobile application.

The coronavirus has also made its own adjustments and caught the transport industry by surprise. The pandemic has drastically changed user preferences and priorities. These changes are the reason for the revision of the strategies of companies that work in the field of transport. In a matter of weeks, public transport such as buses, metro or trolleybuses has become less attractive to buyers than their own transport or ride-hailing. Safety and hygiene become much more important for the user than the cost of moving. The ride-sharing industry, in particular the taxi industry, response to this shift in consumer preferences and the global pandemic overall will be the focus of this work.

Ride-hailing is the central part of the shared mobile market, accounting for 90% of it. McKinsey estimated that the total market size of the shared mobility industry for 2019 was up to 140 billion worldwide, with e-hailing accounting for the majority of 130 billion. The field of ride-sharing has attracted a lot of attention in its time. Companies in this area have attracted investments totaling over $ 100 billion[1].

The present paper aims to compare the influence of the main factors on the pricing of a taxi trip. In order to investigate this, data about taxi rides in 2019, 2020, and 2021 years were taken from the official source of the city of Chicago. Thus, it is possible to

[1] https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/shared-mobility-where-it-stands-where-its-headed

compare the main indicators for the industry before the pandemic, during, and after. Also, for a more detailed analysis, this work takes into account the hourly data of temperature and precipitation in the city of Chicago. Additional information about the community areas of the city is used as well.

CHAPTER 2. INDUSTRY OVERVIEW

In this paper, I will use the term TNP to refer to transportation network providers. This includes commonly known companies such as Uber, Lyft, Didi, who are often also called ride-sharing companies. Their main focus of these businesses is on connecting drivers with passengers via an app on the smartphone. The principal difference between ride-sharing companies and taxi companies is their freedom from regulations and the necessary licenses for drivers, which allows them to take the audience away from typical taxi services.

2.1 TNPs overview

The ride-hailing industry has received much attention in the last few years. Not only because of app innovations but also due to the pandemic effect on mobility as a whole.

Given that data that are investigated in this study collected from the Chicago taxi trips market overview is focused on the United States. The total size of the ride-hailing industry in the United States was almost 46 billion $ in 2018 and 57 billion $ in 2019, showing a clear upward trend. However, in 2020, the outbreak of the coronavirus has changed and the numbers have significantly dropped to 50 billion $. The forecast for 2021 is 59.8 billion, indicating that the taxi industry is returning to a positive trend[2]. These numbers confirm the strong impact of the coronavirus on the industry. The leading positions in the market are occupied by such companies as Uber, Lift, Didi. But it's worth noting that the taxi industry is very location-dependent. For example, Didi is the first on

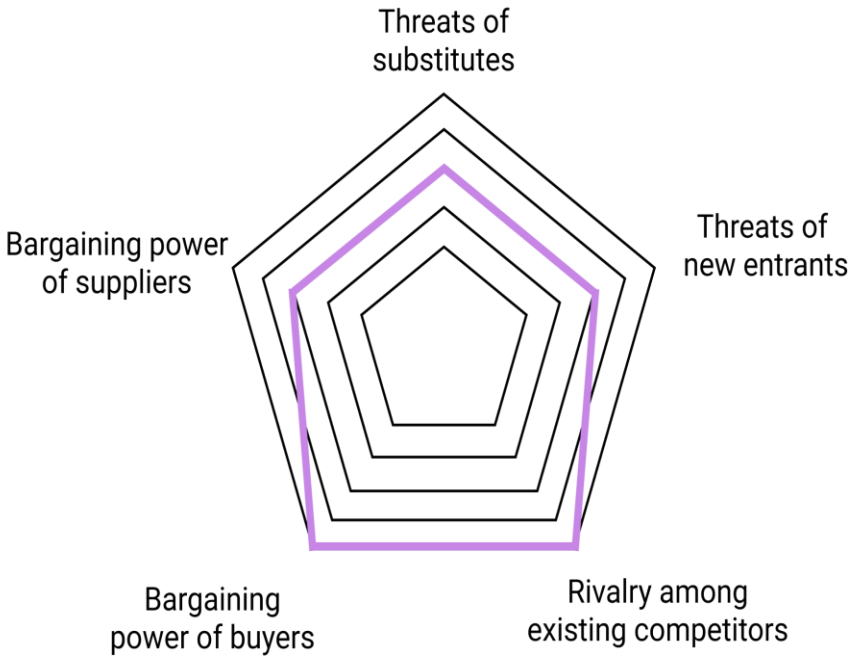
[2] https://www.statista.com/statistics/1174077/taxi-limousine-services-market-size-us/#statisticContainer

the market in China. While Uber and Lyft are ranked first and second respectively in Europe and the United States of America.

2.2 Poster`s 5 forces analysis

To fully analyze the taxi industry, the Poster`s 5 Foster method was chosen. This approach was proposed by Michael Porter in 1979 and is still often used for business analysis and strategy development. Analysis of the taxi industry from 5 points of view: bargaining power of buyers, bargaining power of suppliers, the threat of new entrants, the threat of substitutes, and rivalry among existing competitors, will help to understand exactly how the pandemic could affect this area.

Figure 1. Five Forces Analysis



1. The threat of new entrants

Let's start by looking at the industry in terms of the threat of new entrants. On the one hand, it is relatively easy to develop an application for calling a taxi. In addition, considering the fact that both for drivers and customers the switching cost is low, new competitors will be able to take part of the audience from already big players on the market.

But on the other hand, an important element here is the ability to scale and increase both the customer base and the driver base. It is complicated enough to take away a significant part of the audience from current players. After all, companies existing on the market can afford to offer their drivers better conditions for maintaining and repairing cars. In addition, it is necessary to control both the demand and supply side of the market, that is, to increase the audience of both drivers and passengers.

The significant point is the brand and the customer's trust. When choosing a taxi service, users indicate safety as the top reason. It is critical to understand here that it is a priority for the user to be confident in the safety of the trip and the driver. Otherwise, they will be more likely to choose another taxi service. Behavioral economics explains the behavior of a person in this situation through the availability heuristic and bandwagon effect. Indeed, users will choose the service which they have already heard about somewhere or which their friends and acquaintances used.

That is, even despite the low cost of developing an application, factors such as customers engaging, maintaining a balance of drivers and users, economies of scale, and brand strategy development will become a big obstacle for investors to put their money in this business. The presence of strong players in the market is likely to slow down the process of financing the startup or even become a bottleneck for acquiring investments at all.

2. The bargaining power of suppliers

For the ride-hailing industry, the bargaining power of suppliers is represented mostly by drivers. To begin with, suppliers, as well as buyers, have small switching costs. Various drivers use not one but few apps at the same time. In addition, it is not so difficult to join the company and become one of the drivers of a business that provides taxi services. At the same time, the concentration of drivers can become a problem as it is necessary to manage the distribution of supply and demand for different locations.

Gas and petrol are also important factors. Many factors affect the price of gasoline, and in this regard, the taxi industry businesses have no leverage over this resource. It is worth noting here that the coronavirus also affected the petrol and gasoline market, and therefore influenced the taxi industry through this factor. At the beginning of the pandemic, there was a sharp drop in gasoline prices, and then a gradual increase in prices is noticeable in early 2021. These facts have led me to conclude that the ride-hailing industry has a medium force of bargaining power of suppliers.

3. The threat of substitutes

From the perspective of the threat of substitutes, it is worth noting that fast and comfortable travel is the central reason for choosing a taxi among all means of transport. Renting a car for a fixed period, public transport services, renting bicycles and scooters, buying your own car - all these are possible substitutes for ordering a taxi. There is often a problem with road congestion in largely populated cities, which is why bike rental is becoming a very popular way of getting around. At the same time, many companies in the ride-hailing industry have foreseen this and are now developing their business in this direction as well. For instance, the Bolt company launched a scooter rental in 2018, and in May 2021 provided an opportunity for car-sharing as well.

Many businesses in the ride-hailing industry study the demand for mobility in even more detail and solve other needs and problems, for example, buying groceries or ordering meals from restaurants. In this case, the business differentiates its ability to gain profit and thus reduces the strength of the threat of substitution. Here we can recall the Uber taxi company, which in 2014 launched the Uber Eats food delivery service.

The offer of ride-hailing businesses is still different enough from its possible substitutes and has a rather unique offer on the market. Therefore, I would conclude that the threat of substitution in the taxi market is moderate. Indeed, on the one hand, there are many possible alternatives. However, some of them have high switching costs, such as buying your own car, while with other substitutes there is a possibility to differentiate products, as in the case with food delivery or scooter rental.

4. The bargaining power of buyers

In the case of the taxi industry, buyers are passengers who use a taxi service via an app. The first factor to consider is the alternative replacement options. The initial need of the user is to transfer from one point to another. Potential buyers have the opportunity to meet this need in other ways, such as using public transport, buying their own car, renting a car, or walking. Scooter and bicycle rental services are developing rapidly as well. Consequently, in this industry, not only direct competition is developed, but also alternative replacement possibilities.

Talking about direct competition, we can state that central players have a fairly similar product with basic functions. Applications for calling a taxi are easy to download, publicly available, and, as usual, are generally known. The registration and identification process for all products is approximately the same. Accordingly, switching barriers are generally low and the user will be more price sensitive. This fact, in turn, leads to the point that there is a high buyer bargaining power of buyers in the taxi industry.

5. Rivalry among existing competitors

The market is quite competitive. The specificity of the industry is such that it allows the presence of many strong players but at different locations. Thus, in a particular city or country, there may be a leader, but it is difficult to single out a world leader as such. For example, the DIDI app occupies more than 90 percent of the total ride-hailing market in China. While in the United States of America, Uber is the leader with a market share of about 70 percent[3].

That is, competition does not take place at the global level, but specifically at the local level. Therefore, having a business in several cities with different competitors complicates the development of a business strategy.

2.3 Taxi pricing

Because this study focuses on analyzing the ride-hailing market in the United States and Uber estimates for 70 percent of the market, Uber's pricing strategy is significant. Price formation in this company is formed in real-time and depends on the balance of supply and demand. Depending on many factors such as weather conditions, traffic jams, time of day, or closeness to the metro, users can see diverse numbers in the application throughout the day. It is another distinguishing feature from ordinary taxis in the United States, where the price is formed depending on the cost per minute set by the authorities and the fixed price for getting into the cab. Similar pricing can be seen when renting a scooter or bicycle in many cities. As well as with ordinary cabs, there is a cost of unlocking the vehicle and then there is a charge for every next minute of using the vehicle.

[3] https://www.statista.com/topics/4610/ridesharing-services-in-the-us/#dossierKeyfigures

Although, in Uber, like in many other ride-sharing companies, a dynamic pricing system is used, which is also called surge pricing. This method controls when supply becomes less than demand, and sets the price taking into account a certain multiplier. Thus, it allows companies to balance supply and demand, attracting the attention of drivers to more overcrowded locations and matching them with passengers who are ready to pay the multiplied price.

CHAPTER 3. RELATED STUDIES

McKinsey conducted a study that analyzes how the main reasons for choosing a vehicle have changed. Before COVID-19, respondents for private travel were guided by the following points - time to destination, cost of travel, privacy. While now in the first place is the risk of infection, which previously was already in sixth place. Time to destination has dropped to second place, and the price of a trip has dropped to fifth place in the new top reasons that users pay attention when choosing a transport. This study clearly shows that the pandemic has profoundly changed people's attitudes towards travel and their safety.

On July 20, a survey was conducted that analyzed how the coronavirus affected the ride-hailing industry. 65 percent of respondents answered that the coronavirus has affected their transport habits and now they do not use taxi services. Whereas, 5 percent plan to stop using taxis. And less than a third of respondents said they still use taxis to get around and do not plan to change their preferences.

That is, on the one hand, isolation and hygiene have risen to the top of significant reasons when choosing a means of transportation. Indeed, it has increased taxi demand because of its advantages over other substitutes such as buses or the subway. It is also important to note that in many cities public transport was unavailable during the lockdown. On the other hand, because of the pandemic, demand for transportation has sharply declined. Thus, the distribution among different modes of transport probably shifted in favor of taxis, but overall demand for transport declined more.

In 2020, due to the spread of the coronavirus, many ride-sharing companies introduced additional measures to mitigate the spread of the virus. For example, Uber supplied disinfectants for all network drivers in order to keep the car clean and able to disinfect surfaces. As a result, each additional passenger increases the risk in the way that it

could be a carrier of the virus. This serves as a basis for the hypothesis that the component of the price for the fact of ordering a taxi will increase, which means that the effect of the trip distance on the trip price will decrease. Because no matter how long the trip lasts, the risk of infection remains the same.

Jonas De Vos research in 2020 also focuses on analyzing changes in human behavior and urban movement after the pandemic. Due to the precautionary measures and the request of the governments to keep the social distance, the transition to remote work or online training, the demand for movement around the city has dropped significantly. This suggestion is in line with a survey designed by McKinsey, which claims that before the coronavirus, 37% of the population used public transport at least once a week, while in May 2020 this number decreased to 19%. Furthermore, taking into account the changes not only in the transport preferences of users but also in the daily routine and the general arrangement of life, we can consider the hypothesis of a decrease in the influence of the day of the week on the price of a taxi call. Work and homeschooling, the cancellation of most public events, are likely to be significant factors that affect demand and prices for taxis.

Preliminary work in this field focused originally on taxi trips demand. Several papers, for instance, "Changing Demand for New York Yellow Cabs during the COVID-19 Pandemic" by Ed Manley, have been carried out using an open data source about yellow and green cabs in New York City.

The study "Has Uber Made It Easier to Get a Ride in the Rain?" by Abel Brodeur estimated how rain affects the ability to find a taxi. This study was conducted using 2015 New York taxi trip data. Researches have shown that rainy days lead to an 8 percent increase in travel. This paper gives a reason for taking into account hourly temperature and precipitation data for each area of Chicago.

This paper bears a close resemblance to the one proposed by Guangyue Nian in the article on the impact of COVID-19 on Urban Mobility in Chongqing China. In this study, the spatial OLS regression model is used because of the small number of community areas and a short period selected for analysis. Key findings support the hypothesis that the distribution between days of the week changed after the pandemic. The overall demand for taxis has decreased. An interesting observation was that the average travel speed increased during the pandemic. This can be explained by the decrease in traffic and road congestion.

CHAPTER 4. DATA

The data was taken from the official website of the city of Chicago in the United States (Transportation Network Providers - Trips). Local authorities obliged companies that provide transportation services, particularly ride-sharing companies, to provide trip data. The data has been collected since November 2018. Therefore, it is possible to analyze the industry before the COVID-19 outbreak, during the pandemic, and the recent situation. The data is updated quarterly, which makes it possible to investigate the most relevant data.

Rules and city regulations have also been important since the start of the pandemic. In the city of Chicago, on March 20, a stay-home order was introduced. All public places, restaurants, and museums were closed. At the beginning of summer, these limitations were already slowly declining, gatherings of up to 50 people were allowed. Among other things, summer 2020 in Chicago was quite stable with no drastic changes in the rules of using transport.

Since the privacy of data is very important, data in the public domain is in an aggregated or sometimes incomplete format. In the dataset, each line is a unique trip, but at the same time, privacy is protected by the following restrictions:

- travel time is rounded to the nearest 15-minute interval
- fare for the trip, rounded to the nearest $2.50
- the exact location is not specified, instead, the community area of the beginning and end of the trip is indicated

More details about each variable are written in Table 1.

Table 1. Description of taxi trips data

Variable name	**Variable value example**	**Description**
Trip id	fd99c922d8ae3f428	Unique trip identifier
Trip miles	4.6	Distance of the trip in miles.
Trip start timestamp	08/31/2021 11:45:00 PM	When the trip started, rounded to the nearest 15 minutes.
Trip end timestamp	09/01/2021 12:00:00 AM	When the trip ended, rounded to the nearest 15 minutes.
Trip seconds	969	Time of the trip in seconds.
Pickup community area	PORTAGE PARK	The Community Area where the trip began.
Dropoff community area	ARMOUR SQUARE	The Community Area where the trip ended.
Fare	$15	The fare for the trip, rounded to the nearest $2.50.
Shared trip	FALSE	Whether the customer agreed to a shared trip with another customer, regardless of whether the customer was actually matched for a shared trip.

As an additional source of data, there is also hourly information about the weather in the city of Chicago. Furthermore, supplementary information about the average income of each community area in the city is also taken into account. Table 2 and Table 3 describe each dataset in more detail.

Table 2. Description of temperature data

Variable name	Variable value example	Description
Date time	08/23/2021 11:00:00 PM	Time and date of temperature measuring
Temperature	22.2	Average temperature in Celsius
Precipitation	0.71	The amount of precipitation in millimeters
Conditions	Partially cloudy	Possible values are - clear - partially cloud - overcast - rain

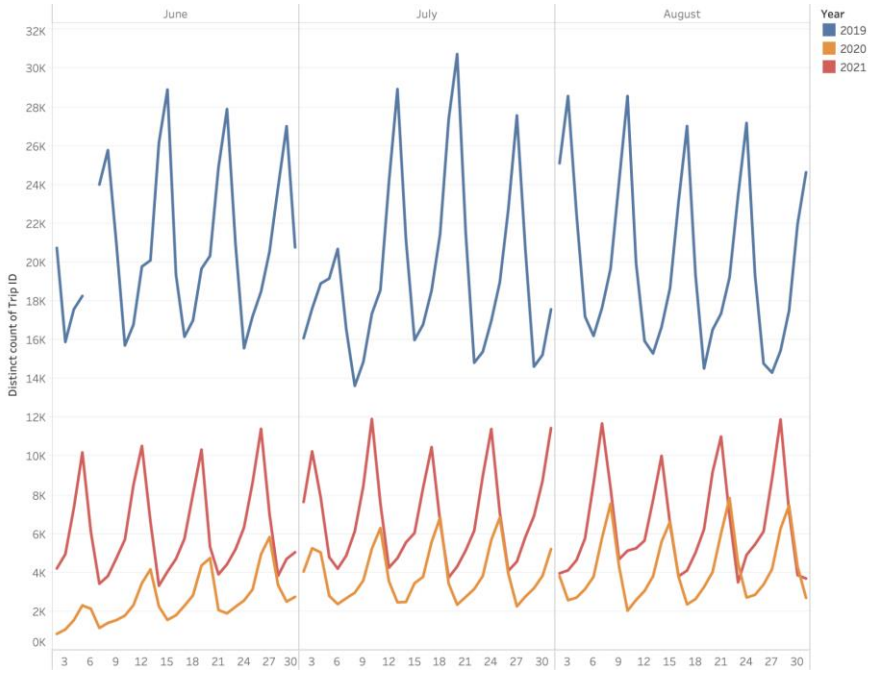
Table 3. Description of community area data

Variable name	Variable value example	Description
Pickup community area	PORTAGE PARK	The Community Area where the trip began.
Income	50 450 $	Median yearly income in this area

A total of 194 millions of rows were collected for this study, where each row is a single taxi trip. Since the data is very voluminous, it was decided to compare only the summer period. This will reduce the time required for data processing, as well as exclude such factors from regression as seasonality, the beginning of the school season, and large temperature drops. Nevertheless, this also turned out to be quite a lot and did not correspond to the computing capability. Therefore, for each summer period, only 10 percent of all trips are randomly taken.

The figure below demonstrates a comparison of taxi trips for summer 2019, 2020, and 2021. This data confirms similar studies, and the demand for ride-sharing services in 2020 fell by almost four times compared to 2019. In 2021, the situation begins to slowly return but is still far from the previous values.

Figure 2. Trips amount for different years



As illustrated in Figure 2, the day of the week has a profound effect on the number of trips. Based on this, a variable was additionally created based on the date of the start time of the taxi trip, which indicates what day of the week it was. The highest peaks usually occur on Saturday, while the least rides happen on Monday. To take into account this trend, it was decided to consider the beginning of the working week as the starting point.

By constructing an extra variable, we are able to estimate the network congestion during a taxi call. This additional variable consists of the number of booked taxi trips that were requested within an hour in the current pick-up community area. Thus, to the line that denotes a single taxi trip, an aggregated variable was added that indicates the overall workload in the community area of ordering a taxi.

The average income variable was originally a continuous variable with particular numbers of the average annual income per community area. Although the data is available only for 2019, it was decided to transform this variable into a categorical one. It was done since the values of the average income most likely could change during this time, while the general status of the district is unlikely. Thus, a variable was formed, where 1 means the average annual income is above 80k, and 0 if the average income is below this threshold.

CHAPTER 5. METHODOLOGY

The multivariable OLS linear regression was chosen because it is one of the most practical ways to analyze the influence of various factors on the price. Therefore, having price as a dependent variable will enable us to analyze the effect on price and test the above hypotheses.

Since the price distribution is rather skewed, it was decided to take the logarithm. The same applies to the trip distance (in miles) and precipitation (in mm). This is done so that the distribution is close to normal and there are no problems with heteroscedasticity.

Figure 3. Comparing the variable price and log(price)

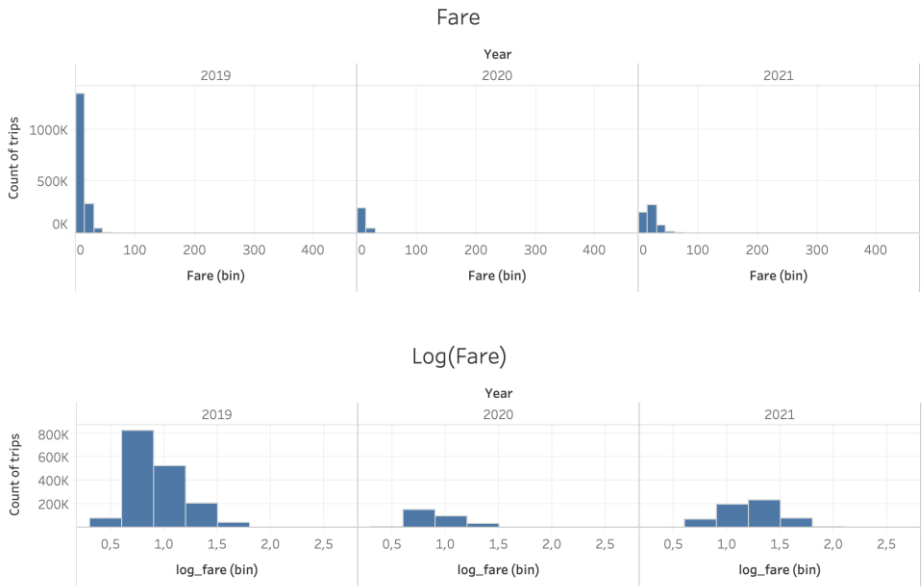


Figure 4. Comparing the variable distance and log(distance)

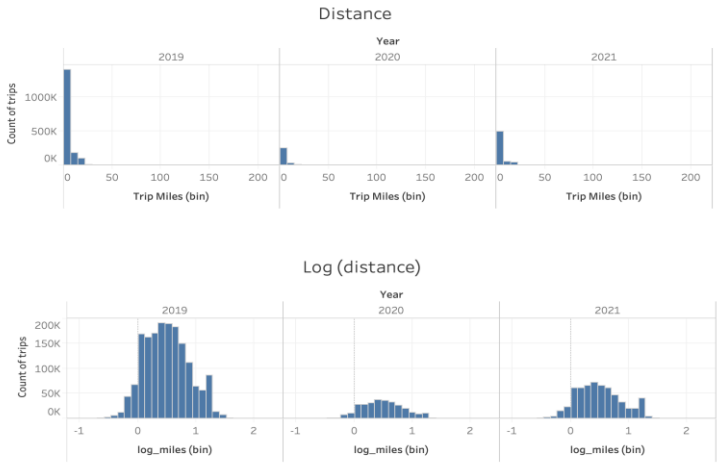
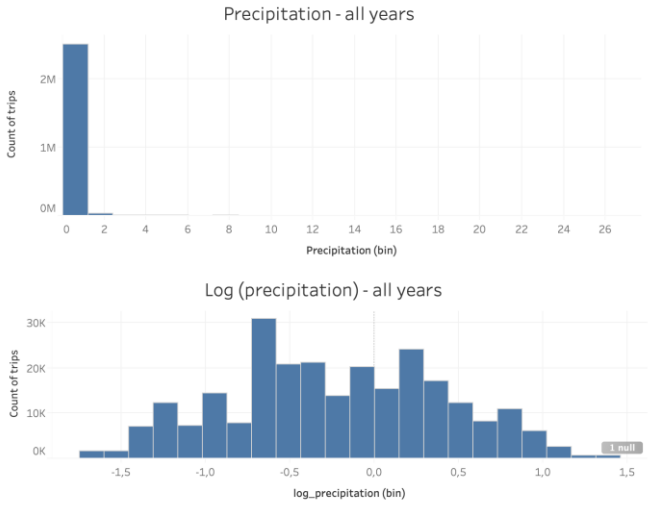


Figure 5. Comparing the variable precipitation and log(precipitation)



First, for each year, correlation matrices were built. Having correlation matrices enabled us to investigate the interdependencies between every two variables in each dataset. We will assume that a correlation value whose absolute value is greater than 0.5 will mean a high correlation and it will be necessary to exclude one of the variables.

Table 4. Correlation matrix (2019)

	Trip Miles	Precipitation	Workload	Temperature	Day of week	Is airport trip	Median income
Trip Miles	1.00	-0.03	-0.23	-0.02	-0.02	0.51	-0.29
Precipitation	-0.03	1.00	0.01	-0.06	-0.06	-0.01	0.01
Workload	-0.23	0.01	1.00	0.04	0.11	-0.12	0.51
Temperature	-0.02	-0.06	0.04	1.00	-0.03	-0.02	0.00
Day of week	-0.02	-0.06	0.11	-0.03	1.00	-0.06	0.00
Is airport trip	0.51	-0.01	-0.12	-0.02	-0.06	1.00	-0.17
Median income	-0.29	0.01	0.51	0.00	0.00	-0.17	1.00

The correlation matrix above shows that in most cases the absolute value of correlation does not exceed 0.5 or exceeds insignificantly. The mileage of the trip correlates

slightly with the binary variable that represents whether the trip was from or to the airport. This is due to the moderately long distance of the airport from downtown Chicago. The binary variable about the community area wealth and the network workload are also somewhat correlated. The correlation matrices for 2020 and 2021 are quite related and display approximately the same relationships as in 2019.

The central role of the VIF method is to point out the problem of multicollinearity of independent variables when forming a model. VIF is calculated for each of the variables and indicates how much it correlates with the other variables. There is a general rule for determining the range of acceptable values. If the value of VIF is up to 5 for all variables, then multicollinearity is weak, from 5 to 10 - considered from moderate to high, more than 10 is rated as extremely high. These benchmarks were used to determine which parameters should be left in the model and which should not.

Table 5. Correlation among independent variables - VIF

Variable name	**VIF**		
	2019	**2020**	**2021**
Trip Miles	3.52	3.31	3.85
Precipitation	1.03	1.01	1.02
Workload	2.98	2.48	2.41
Temperature	7.77	8.50	9.79
Day of week (Monday = 0, ..., Sunday = 6)	3.62	4.64	4.78
Is trip to/from airport (True/False)	1.45	1.23	1.59
Community area with high median income (True/False)	4.98	4.35	5.23

These tests revealed that the value of VIF does not rise above 5 for most of the variables in each year. The exception is the temperature variable. The value for VIF can be called quite high for all years. Even though there was no direct correlation with other variables in the covariance matrix, a VIF value above 5 should be considered sufficiently high and consider removing this variable from the regression.

Table 6. Correlation among independent variables (without temperature) - VIF

Variable	**VIF**		
	2019	**2020**	**2021**
Trip Miles	2.65	2.35	2.88
Precipitation	1.03	1.01	1.02
Trips Count	2.96	2.48	2.41
Day of week (Monday = 0, ..., Sunday = 6)	2.90	3.45	3.67
Is trip to/from airport	1.44	1.22	1.58
Community area with high median income	3.83	3.28	3.50

By removing the temperature from the equation, all VIF values for the variables are now within acceptable limits. From the graph, we can note that.

After the changes are done, the final equation looks as follows:

$$log(Fare) \sim \beta_0 + \beta_1 * log(Trip\ Miles) + \beta_2 * log(Precipitation)$$

$$+ \beta_3 * Workload + \beta_4 * dayOfWeek$$

$$+ \beta_5 * isAirportTrip + \beta_6 * isHighIncome + \varepsilon$$

where

- fare - continuous variable,
- trip miles - continuous variable,
- precipitation - continuous variable,
- workload - continuous variable,
- dayOfWeek - categorical variable with 0 as Monday, 1 as Tuesday and so on,
- isAirportTrip - binary variable that indicates whether a pick-up or dropoff location was an airport (0 - False, 1 - True),
- isHighIncome - binary variable that shows whether.

The expected effect of independent variables on taxi prices:

- the longer the distance between the peak-up or drop-off point (that is, the trip in miles), the higher the cost: a positive coefficient
- More precipitation, the higher the cost of a taxi ride: a positive coefficient

- More network congestion, the higher the cost of the trip: a positive coefficient

- Coefficient next to a binary variable that tells whether a trip was to/from the airport is expected to be positive; closure of borders due to the pandemic and, therefore, a decrease in demand for flights: a positive coefficient

- Coefficient next to a binary variable, which is responsible for the average income of the community area, is also expected to be positive due to the fact that a taxi can be considered as a more luxurious service: a positive coefficient

Hypotheses for comparing price influencers for 2019, 2020, and 2021:

1. Influence of the day of the week has become less on the price of a taxi trip
2. Duration of the trip (in miles) has less influence on the fare of a taxi ride
3. Parameter for the variable about travel to the airport has changed, taxi trips to the airport have become cheaper

CHAPTER 6. RESULTS

Three OLS linear regressions were done for each year. The formula of the model is the same for each year, this will enable us to then compare the results.

All three models showed more or less similar results. The expected results on the direction of influence of factors on the trips price, which was described in the previous section, were confirmed. For all three models, the coefficients have a positive effect on the fare of the taxi trip. More details about each year and their comparison will be discussed later in this chapter.

Moreover, it is crucial to remember that such variables as the cost of the trip, trip miles, and precipitation are expressed through the logarithm. Therefore, its interpretation will differ from standard linear regression. Besides, further interpretations of the coefficients will always mean a requirement for the steadiness of the remaining variables (ceteris paribus).

6.1 Pre-COVID – 2019

For 2019, R2 is 70 percent. This means that 70 percent of the data can be explained by this regression, which can be considered as a high result. Moreover, all parameters were statistically significant. From the table, you may notice in more detail the standard deviation and confidence intervals for each factor.

An increase of one percent in the mileage of a trip causes an increase in the cost of a trip by 0.53 percent. Each additional percent in mm of precipitation will incur a price increase of 0.01 percent. Every 100 additional trips in the current hour in the specific community area where a taxi is ordered increases trip fare by 4%.

As indicated in the figure in the previous chapter, the cost of travel increases by the end of the week.

A significant factor is a trip to the airport or from it: it is on average 30% more expensive. The average income of the community area is also a statistically significant factor. Taxis from a more profitable area will cost about 1.7 percent more.

Table 7. Coefficients of OLS regression model (2019)

	coef	**std err**	**t**	**[0.025**	**0.975]**
const	1.452**	0.001	1828.789	1.450	1.453
Trip Miles	0.534**	0.000	1617.521	0.533	0.535
Precipitation	0.010**	0.000	33.326	0.009	0.010
Trips Count	0.0004**	0.000	173.286	0.000	0.000
Day of week (Monday = 0, ..., Sunday = 6)	0.006*	0.000	45.131	0.005	0.006
Is trip to/from airport	0.313**	0.001	274.499	0.311	0.315
Community area with high median income	0.017**	0.001	25.735	0.016	0.018
R2 = 70%, *p<0.1; **p<0.05; ***p<0.01					

6.2 Pandemic – 2020

As well as for 2019 an increase of one percent in the mileage of a trip causes an increase in the cost of a trip by 0.56 percent. Thus, the influence of distance on the cost of travel increased slightly, which unfortunately does not agree with the hypotheses indicated in this work. Each additional percent in mm of precipitation will incur a price increase of 0.004 percent. Every 100 additional trips in the current hour in the specific community area where a taxi is ordered increases trip fare by 2%.

During the pandemic, the influence of the day of the week on taxi fare has increased, which also does not keep up with the initial assumption. This can be explained by the fact that the share of taxi trips to the workplace has decreased to a greater extent than taxi trips to events on weekends.

The hypothesis of reducing the cost of travel to or from the airport was confirmed. In 2020, the fare of a trip to the airport or from the airport dropped by 15 pp. That is, such a trip will no longer cost 30 percent more than regular trips, but 15 percent. This is logical in connection with the closure of borders and the decrease in demand for air travel.

Table 8. Coefficients of OLS regression model (2020)

	coef	**std err**	**t**	**[0.025**	**0.975]**
const	1.409**	0.002	883.207	1.406	1.412
Trip Miles	0.565**	0.001	856.733	0.564	0.567
Precipitation	0.004**	0.000	7.995	0.003	0.004
Trips Count	0.002**	0.000	94.748	0.001	0.002
Day of week (Monday = 0, ..., Sunday = 6)	0.018**	0.000	64.191	0.017	0.018
Is trip to/from airport	0.156**	0.003	54.539	0.150	0.161
Community area with high median income	0.0003**	0.001	-0.249	-0.003	0.002
R2 = 74%, *p<0.1; **p<0.05; ***p<0.01					

6.3 Post pandemic – 20210

In 2021, not all trends from the previous year have been preserved. There is also the possibility to track changes in factors after exiting lockdown and strict restrictions.

In the case of this year, the trip distance already has less influence on the cost of the trip. With an increase of 1 percent in the trip distance, the trip rises in price by 0.41 percent. The influence of the day of the week has become even greater. As the end of the week arrivals, the trip increases in price by 6 percent.

Table 9. Coefficients of OLS regression model (2021)

	coef	**std err**	**t**	**[0.025**	**0.975]**
const	1.978**	0.002	1010.146	1.974	1.981
Trip Miles	0.417**	0.001	533.485	0.416	0.419
Precipitation	0.030**	0.001	55.642	0.029	0.031
Trips Count	0.001**	0.000	130.481	0.001	0.001
Day of week (Monday = 0, ..., Sunday = 6)	0.062**	0.000	197.502	0.061	0.062
Is trip to/from airport	0.237**	0.002	96.740	0.232	0.242
Community area with high median income	0.098**	0.002	64.076	0.095	0.101
R2 = 47%, *p<0.1; **p<0.05; ***p<0.01					

CHAPTER 7. CONCLUSIONS AND RECOMMENDATIONS

The coronavirus, global pandemic, lockdown have brought many changes to our daily lives and the operation of many businesses. The taxi industry is no exception. Considerable research has already been conducted on changes in behavioral factors that affect the demand for different modes of transport, including taxis. And indeed, even though the main advantage of taxis is safety and convenience, the pandemic still significantly reduced the demand for all modes of transportation.

This paper took a closer look at the taxi industry and the main risks associated with this market. The general market analysis was carried out using the Poster`s 5 Foster method. Each of the 5 powers was fully researched and it turned out that for the ride-hailing industry, all forces have a moderate or high impact on the business.

Due to the sharp decline in demand, the strategies of ride-hailing companies have changed a lot. Therefore, the primary focus of this study was to investigate what affects the price of the trip and how the influence of these factors changed during the pandemic.

In an attempt to do this, we took the taxi ride data for the summer for 2019, 2020, and 2021 for the city of Chicago. Thus, over 2.5 million taxi trips were analyzed in this work. Moreover, data on hourly temperature and precipitation and data on average income from selected areas of Chicago were used as well.

After collecting the data and transforming the variables, it was decided to use linear regression for each period to determine the influence of factors on the price. Answering the question of what affects the pricing, all the factors that were used in the model turned out to be statistically significant. All of them had a positive effect on the cost of the trip. That is, with their increase, the taxi trip fare also increased. The list of these factors includes the trip distance in miles, the amount of precipitation, the day of

the week, whether this trip is to the airport or not, the average income of the pick-up community area, and the workload.

The second research question concerned the change in the influence of factors on the price. Three hypotheses were developed that dealt with variables about travel distance, day of the week, and travel to or from the airport. The last hypothesis about a reduction in the fare of a trip to or from the airport during a pandemic was approved. Indeed, in 2019, a taxi trip to the airport was 30 percent more expensive than the rest. While in 2020 this number dropped to 15 percent. That is the influence of the drop-off or pick-up area as an airport has become less significant.

Unfortunately, hypotheses about reducing the impact of travel distance were not confirmed in 2020 but showed a significant decline in 2021. Also, the day of the week turned out to be more vital in the models for 2020 and 2021 compared to the period before the pandemic. This can be explained by the fact that the demand for work travel has decreased more than the demand for travel to events on weekends.

Rejection for several hypotheses makes it clear that the data for each city is different and that data plays a large role in the investigation of this industry. As stated in the analysis of the ride-hailing business, the taxi market is quite various for different locations and therefore we may conclude that the strategies and pricing will also differ. In the case of this report for the city of Chicago, we got the before-mentioned figures. Of course, most of the factors will have the same influence on the price, but the degree of their influence may differ over time, as was analyzed in this work.

REFERENCES

Jonas De Vos Transportation Research Interdisciplinary Perspectives. 2020 May. *The effect of COVID-19 and subsequent social distancing on travel behavior*

Nikhil Menon, Ph.D (2020). Research Reports. 254. *"Impact of COVID-19 on Travel Behavior and Shared Mobility Systems"* https://scholarcommons.usf.edu/cutr_nctr/254

Ci Yang, Eric J. Gonzales, 2014. *"Modeling Taxi Trip Demand by Time of Day in New York City"*

McKinsey Center, December 15, 2020 "*From no mobility to future mobility: Where COVID-19 has accelerated change"*
https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/from-no-mobility-to-future-mobility-where-covid-19-has-accelerated-change

McKinsey Center, August 11, 2021*, "Shared mobility: Where it stands, where it's headed"*
https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/shared-mobility-where-it-stands-where-its-headed

Guangyue Nian, 2020, "*Impact of COVID-19 on Urban Mobility during Post-Epidemic Period in Megacities: From the Perspectives of Taxi Travel and Social Vitality"*

Kristin Mammen, Hyoung Suk Shim, 2017 *"New York City Taxis: Demand and Revenue in an Uber World"*

Jianhe Du, Hesham A. Rakha, 2020, "*COVID-19 Impact on Ride-hailing: The Chicago Case Study"*

Ed Manley, Stuart Ross, Mengdie Zhuang, May 2021, *"Changing Demand for New York Yellow Cabs during the COVID-19 Pandemic"*

Charitha Dias, November 2020, "*Exploring the impacts of COVID-19 on travel behavior and mode preferences"*

Abel Brodeur, Kerry Nield, June 2016, *"Has Uber Made It Easier to Get a Ride in the Rain?"*

Peter Cohen, Robert Hahn, Jonathan Hall, September 2016, *"Using Big Data to Estimate Consumer Surplus: The Case of Uber"*

Ci Yang, Eric J. Gonzales, January 2014, "*Modeling Taxi Trip Demand by Time of Day in New York City"*

Jabbari, Parasto, and Don MacKenzie. 2020. *"Ride Sharing Attitudes Before and During the COVID-19 Pandemic in the United States."*

Zhang, Jiameng; Li, Jing; Lu, Shengyun, 2016, *"Factors Affecting the Demand for the Taxi – Evidence from Zhejiang, China"*

Haleh Ale-Ahmad, Hani Mahmassani, "*Impact of COVID-19 on Taxi Operation in Chic From no mobility to future mobility: Where COVID-19 has accelerated change*