The study of the impact of COVID-19 and related public health policies on taxi drivers' income: a data mining approach

Zixi Qin (qzx.lydia@gmail.com)
University of Melbourne

Mengyan Zhang
University of Melbourne

Research Article

Keywords: Covid-19, taxi drivers, income, public health policy

Posted Date: June 29th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1725237/v1

Abstract

Background: The outbreak of Coronavirus Disease 2019 (COVID-19) has had a significant impact on economic activity and taxi drivers are representative of groups that suffer substantial financial loss. This study aims to examine the impact of this pandemic and related public health policies on taxi drivers' income.

Methods: 350 million records of taxi trip data in New York City (NYC) in 2020 from February to December were used in this study. The taxi trip data was summarised based on the unit of day and combined with the NYC pandemic data and the related policies details. Then we built two-way ANOVA models to assess the significance with relevant factors.

Result: Based on the model comparison, effects from both 7-day averages of count on the total payment per day and case count on the average tip per day are significant, with $p < 2e\text{-}16$, and their corresponding p-value of public health policy is $p < 2e\text{-}16$. Moreover, both total payment per day and average tips per day demonstrate a similar tendency: a significant reduction after the "New York State Pause" followed by a gradual growth after the Phase 1 reopening. The study also discovers that the shrinkage in demand for taxis between 2019 and 2020 is larger during night-time and weekends, compared with daytime on weekdays.

Conclusion: The Covid-19 will lead to a decline in taxi drivers' income, while influence of related public health policies is noteworthy. We recommend that governments should build a database of the impact of public health policies and pandemics on various industries to provide an adequate basis for future decisions under similar circumstances.

Introduction

At the end of 2020, several cases of Novel Coronavirus SARS-CoV-2 were reported in Hubei Province, the People's Republic of China, then the outbreak of COVID-19 started to spread worldwide. Until 12 December 2021, COVID-19 has caused 269,468,311 confirmed cases and 5,304,248 deaths[1]. The data from the World Health Organization (WHO) reveals that the United States of America is the country with the greatest number of confirmed cases and deaths in the world, respectively 49,531,538 and 790,304[1].

In order to ease the rise in positive cases, many countries adopted non-pharmaceutical interventions. The negative effect brought by this pandemic and such non-pharmaceutical interventions on the economic market is also serious. From the perspective of macroeconomics, the economic costs of this disease would enlarge quickly where COVID-19 becomes a global pandemic and the government would play a crucial role against such social and financial stress[2]. From the perspective of microeconomics, three departments committed to the major impacts of COVID-19 and non-pharmaceutical interventions: Restaurant industry[3], Tourism industry and Transport industry[4].

Even before the pandemic, serious physical health issues already existed among taxi drivers[5][6][7]. Taxi drivers also face heavy financial pressures, since they need to work longer to cover the high taxi lease costs[8]. Under the COVID-19, taxi drivers endure worse situations. First, the likelihood of hospitalisations with COVID-19 of taxi drivers is 1-2 times higher than all infected individuals of working age[9]. Also, data from England, Wales and Sweden disclosed that taxi drivers have a higher mortality rate of COVID-19[10]. Furthermore, the following problems lead taxi drivers to face high mental pressure: 1) Risk of transmitting the virus to family members[11]; 2) Pause in work caused by quarantine threaten livelihood for their family[12]; 3) Unable to contact other drivers while waiting for new orders. Such social distancing restrictions had potentially adverse effects on people's mental health[11].

As social distance finalised by the government health committee cannot be maintained in a confined area during this pandemic, people prefer to select private transportation in case a higher risk of getting infected by public transportation[13]. Because of the construction of taxi drivers' work, the vulnerability of this career is sharpened under the pandemic, including losing jobs and financial sources[14][15]. According to the data of the Ningbo Taxi industry, Jingru Yu found that the mean daily income of taxi drivers during the outbreak was higher than the remission period since the number of taxi drivers steeply decreased. Their analysis exposed that the negative effect brought by the COVID-19 on taxi drivers were more serious than ride-sourcing drivers[16].

Nevertheless, the previous research did not examine the impact of distinct levels of government public health policies and different degrees of new confirmed cases on taxi drivers. From the earlier literature review, taxi drivers are a vulnerable group that has a relatively low ability to resist risks during the pandemic[14]. Our research aims to investigate the potential effect of public health policies and coronavirus on taxi drivers' earnings. Further, we would analyse the inclination of citizens to take taxis before and after the outbreak of COVID-19. The result of our research may assist the public health committee to deliver a better compensation mechanism for target different social groups.

Data And Method

1. Coronavirus Disease Data

The health section in NYC official website provides data on COVID-19 in NYC[17]. The Coronavirus disease dataset used in this study includes the temporal information of COVID-19 cases from the date of announcing the first laboratory-confirmed case, February to December 2020. The data are retrieved from the official GitHub repository including 67 factors and only the listing factors have been used in this research: Date, Case count, 7-day average of count (7-day average of count of confirmed cases citywide).

2. New York City Taxi Data

The New York City Taxi Data are collected from the Taxi Limousine Commission section on NYC official website[18]. The records of the taxi data used in this research are from February to December in both 2019 and 2020. Data refers to both green taxis and yellow taxis, so taxi zones could cover the whole

region of NYC. The following information in each record is used in the study: Pick-up and drop-off date-time, Trip distance, Tip amount and Total amount of payment.

Our study assumes the number of taxi drivers in NYC keeps consistent. Under this assumption, the change in the sum of total amount of payment would be a typical factor to denote the change in taxi drivers' income. We also assume the average tip amount per day is all received by the taxi drivers.

3. Public Health Policy Data

The public health policy is adjusted following the changes of the epidemic. There are explicit four stages of reopening in NYC, while after the fourth, the policy is difficult to be classified into unambiguous levels. To summarise the public health policy in 2020, we divided the policies into three levels based on the scope of the policy impact: No policy, Policy that only influences part of the citizen in NYC, Policy that influences almost all the citizen ("New York State on Pause"). The Public Health Policy Data are summarised in Table 1.

Table 1. The three levels of the Public Health Policy and the corresponding period.

Policy	Time frame
No policy = 0	1-2-2020 to 13-3-2020
Policy that only influences part of the citizen in NYC = 1	14-3-2020 to 21-3-2020 8-6-2020 to 31-12-2020
Policy that influences almost all the citizen ("New York State on Pause") = 2	22-3-2020 to 7-6-2020

4. Model

Our study analyses the effect of the COVID-19 and public health policy on taxi drivers' income. Because there is no linear relationship between taxi drivers' earnings and the rest of the independent variables, all features in the Coronavirus Disease Data have been discretised into disparate levels for comparisons in one-way ANOVA. Finally, two two-way ANOVA models are established to estimate the decrease in taxi drivers' earnings by comparing different levels of the COVID-19 and public health policies.

Result

Since the correlation among the factors in the Coronavirus Disease Data is very high and meanwhile considering the influence of collinearity, only one factor will be chosen from the Coronavirus Disease Data and added into the two-way ANOVA model. Based on the One-way ANOVA of each feature in the Coronavirus Disease Data, the effect of the 7-day average of count on total percent in a day is most notable. For the ANOVA of the average tip amount in a day, the case count will be added into the analysis.

1. Total payment per day

The first ANOVA model is conducted to compare the main effect of the 7-day average of count from the Coronavirus Disease Data and Public Health Policy Data on the total payment per day. For the 7-day average of count, it yields an effect size with $\eta p2 = 0.82$ (95% CI: 0.79-0.84), indicating 82% of the variance in the total payment per day can be explained by pandemic data ($F(5, 327) = 846.3$, $p < 2e-16$). And the 14% (95% CI: 0.07-0.20) of the variance in the total payment per day can be explained by the public health policy ($F(2, 327) = 846.3$, $p < 2e-16$).

Figure 1 illustrates the difference of Estimated Marginal Means (EMM) of the sum of total payment per day under different levels of a 7-day average of count. Because of the outbreak of the COVID-19, the EMM of the sum of total payment per day decreased from $4,015, 409.05 to $1,116,858.25. The fluctuated and continuous drop is shown as the 7-day average of count increased from (1,271] to (3820,5782], and finally hit the bottom, with $317,534.0. Figure 2 indicates the EMM of the sum of total payment per day diminishes as the impact area of public health policy expands. When there is no policy, the EMM of the sum of total payment per day was $3,880,781.85. After the policy associated with COVID-19 was announced, the EMM of the sum of total payment per day falls to $627,99.39 and down to the bottom after the 'NYS on Pause', with $147,564.52.

Figure 3 presents the tendency of total payment per day in a day during the COVID-19. It reveals an obvious drop along with the increase in as the new confirmed cases and the whole of New York State is paused meanwhile, from 4.0 million dollars drop to 0.14 million. As phase 1 of reopening started, the total payment per day begins to resume gradually, hitting a high of 1.86 million in November. However, the second outbreak of the COVID-19 resists its slow recovery and the total payment per day shrinks to 0.3 million in December.

2. Average tip amount in a day

A two-way ANOVA is performed to analyse the effect of the Coronavirus Disease Data and Public Health Policy on the average tips in a day. The result demonstrates that there is a statistically-significant difference in the average tips under different degrees of the case count ($f(4)=395.8$, $p < 2e-16$) and distinct levels of public health policies ($f(2)=267.3$, $p < 2e-16$). Additionally, both of case count and public health policies have a high effect size, with respectively $\eta p2$(case count) = 0.39 (95% CI: 0.31-0.46) and $\eta p2$(public health policy) = 0.40 (95% CI: 0.32-0.46).

Figure 4 discloses the impact of the pandemic on average tips. The EMM of average tips before the outbreak of pandemic is $2.15, while as the daily new cases increases to (1, 200], the EMM of average tips drops to $1.93. The EMM of average tip increases to $1.87, even the daily new cases increase to (573,860], however, it finally hit the bottom with $1.72 when the daily new cases in (860,6353]. Figure 5 detects the change in average tips under three levels of policies, and it clearly indicates the reduction of average tips from $2.18 to $1.55 as the policy change from 0 to 2.

Figure 6 indicates the trend of tip amount under the impact from COVID-19 and public health policy. The average tip amount drops from $2.3 to $1.3 during the period between 5 March 2020 and 4 April 2020. After phase 1 of reopening, the average tip reveals a steady increase to around $2.0. However, as the second wave of the epidemic spread, average tip begins to decline slightly.

3. Change of demand between 2019 and 2020, by hour and day of the week

The demand for taxis between February to December in both 2019 and 2020 are chosen to be analysed. The data exposes that the total demand in 2019 is 79,068,996, while the value falls to 17,628,388, with an 77.7% reduction. With respect to displacement in relative demand by the hour and day of the week, from Figure 7, the fall in demand is slighter during the working hour, 6:00-17:00 from Monday to Friday, with -0.74 on average. And the decrease in demand at weekend from 6:00 to 17:00 is around -0.77. For the late-night, from 19:00 to 5:00, the average decrease in demand is -0.83.

Discussion

The policy is developed in response to the pandemic and to protect the health of all human beings. As the pandemic becomes severer, in order to reduce human movement, the related public health policies that restrict the daily travel of citizens are strengthened. During the exploration of the data, we found that even though public health policies significantly affected taxi drivers' income, it was difficult to consider it as an independent variable, since such policies were made against the spread of COVID-19. Hence, the impact of public health policy and pandemics on the income of taxi drivers is difficult to separate. The negative impact brought by the pandemic and related public health policies on taxi drivers' income and the taxi companies behind it should be seriously considered by the government.

Based on the NYC 11-month-length taxi trip data of 2019 and 2020, the difference in total ridership between these two years has been demonstrated that this epidemic did have a significant impact on residents' daily travelling. More specifically, there is still a demand for taxis during weekdays, while having a lower frequency of travelling during late-night and weekends, as Figure 7 reveals. This overall suggests that COVID-19 has decreased the travel rate, nevertheless, people have to go out (for work) for the purpose of their livelihood. Such a result is similar to what Ed Manley disclosed that the taxi demand showed a downward trend between 2019 and 2020[19]. However, the result of the analysis of displacement of relative demand between 2019 and 2020 has some gaps, because Ed Manley's data was based on the yellow taxi in Manhattan with a short period at the beginning of the pandemic[19]. While the data used in our study includes both yellow taxis and green taxis and spans 11 months in 2019 and 2020. As the impact of COVID-19 has been last for a whole year in 2020, therefore, the overall result demonstrates a significant change in the displacement of relative demand between 2019 and 2020 with a negative side.

The average tip amount has been decreased significantly during the pandemic. Jimin et al. [20] suggest that consumption behaviours during the pandemic have been changed and people would like to spend more money on necessities, like housing (46%), food and beverage (21%), rather than traffic trip (8%).

This could explain that the decline tendency in average travel tips as Figure7 shows is normal, as the travel tip would be naturally saved from basic needs expense.

In previous studies, researchers had pointed out the impact of the pandemic on taxi drivers' income. In this study, we propose and demonstrate that public health policy is also an important factor that cannot be ignored. Figure 3 indicates that the total payment per day tends to decrease after the first positive case appeared. 7 days after the whole of NYC paused, the total payment per day dropped to the bottom on 29 Mar 2020, with $91,605.72, as the average total payment per day in February is $3,996,381.8. Until the first easing of restriction was announced on 8 June 2020, the total payment per day indicates a gradual increasing tendency and peak at $1,863,536 on 7 October 2020. When the second wave of pandemic came to NYC, there were no new restriction policies which were like “the New York State Pause” enacted. The total payment per day has bottomed out, with $293,707.60 on 25 December 2020 faced by such pandemic situations. The average tip amount per day has a similar tendency to the total payment per day. Overall, the reduction in total payment per day and average tips amount per day caused by COVID-19 is less than that caused by relevant public heath polices.

The significant drop in total payment per day and the number of taxi trips discloses a decline in the revenue of taxi drivers. The decrease in income due to the epidemic and policies will be a direct threat to their livelihood. Since the number of taxi drivers is enormous and the importance of health maintenance and other problems is not neglectable, it is necessary for the government providing special support to this industry. We express that government should grant more subsidies for taxi drivers, such as allowances for oil fees and highway fees, supply for hygiene essentials like masks and hand sanitisers. Further, the government could potentially reduce the tax for taxi drivers. This act would relieve the economic pressure for most taxi drivers and keep their savings for longer living maintenance. For the enterprise which suffers a heavy economic loss, the government should permit financial grants to help them pass through the hardship.

Other pandemics similar to the new coronavirus are the 1918 flu pandemic, which caused 500 million people to become infected with this virus[21]. Since the data in the past is less complete and reliable with the limited channels, the influence of the flu is less quantified. Hence, the government did not have the opportunity to build a complete database to give effective support in case of similar outbreaks in the future. Nevertheless, it is still possible to abstract the little clues on the financial influence on different industries by the 1918 flu pandemic. There was a report in the newspaper at that time that pointed out that retail business dropped to one-third to one-half of usual sales. And the capacity of industrial plants is lower than the time before the flu, as the worker got sick[22].

During the period of 1918 flu pandemic, the understanding of the damage of pandemic on social economy is unquantified and obscure as the world did not enter the era of big data. In the future, an epidemic with the same magnitude could happen again. Before a disaster strikes again, the WHO and national governments should build a complete database to analyse aspects of the impact of infectious diseases and public health policies on different industries from divergent perspectives. Such database

may involve the data from the following fields, including demographic economics, urban economics, information economics and public health. This is an evolution of transferring traditional economics to big data economics, and this transformation arises higher requirements of data storage means, processing equipment, and processing methods[23]. With the help of this database, governments will have quantitative data to support their decisions on targeted subsidy policies if similar epidemics occur.

Conclusion

Our study explores the impacts of different levels of the COVID-19 pandemic and public health policy on the taxi drivers' income in New York City, based on the 11 months taxi datasets with more than 17 million records. We investigate the change in taxi demand between 2019 to 2020, and the tendency of total payment per day and average tips per day, which undoubtfully exposes that taxi drivers' incomes have indeed been influenced by the pandemic and relative public health policies. Since the policies were developed based on the pandemic, it is difficult to distinguish the degree of independent impact of epidemic data and public health strategies. Policymakers should consider subsidising those groups whose interest has been damaged because of the related public health policy. Moreover, we recommend the governments could establish and collect a dataset that includes the influence due to the COVID-19 from public health policies and epidemics on various industries. This would also support the development of public health strategies over infectious disease control in the future from the big data perspective.

Declarations

Ethics approval and consent to participate

The purpose of this research paper is to explore a social phenomenon. The data used in this study, COVID-19: Data and TLC Trip Record Data, were obtained from the public open database on the official website of New York City, and there are no issues of privacy disclosure and no ethical issues of human and animal experimentation.

Consent for publication - Not applicable

Availability of data and materials - The datasets analysed during the current study are available in the TLC Trip Record Data repository(https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page) and COVID-19: Data (https://www1.nyc.gov/site/doh/covid/covid-19-data.page)

Competing interests - The authors declare that they have no competing interests.

Funding - Not applicable

Authors' contributions

Zixi Qin: Conceptualization, Data Curation, Methodology, Formal analysis, Writing - Original Draft, Supervision

Mengyan Zhang: Writing - Data Curation, Validation and Visualization, Review & Editing

All authors read and approved the final manuscript

Acknowledgements -

Throughout the writing of my first dissertation, I have received a great deal of support and assistance.

I would first like to thank my supervisor, Ling Qin, whose expertise was invaluable in formulating the research questions and methodology. In addition, I would like to thank Zhijiang Mo, who encouragement to carry on during the hard times when I was working on the statistical section. Moreover, I would like to express my gratitude to Professor Wanqing Li, for his valuable guidance throughout my studies. I acknowledge the contribution of Mengyan Zhang, who not only provided stimulating discussions and share the academic pressure, as well as happy distractions to rest my mind outside of my research.

References

1. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. (2021, December 12, date last accessed)Retrieved from “https://covid19.who.int"
2. Baldwin R, Di Mauro BW. Economics in the time of COVID-19: A new eBook. VOX CEPR Policy Portal. 2020 Mar:2–3.
3. Song HJ, Yeon J, Lee S. Impact of the COVID-19 pandemic: Evidence from the US restaurant industry. International Journal of Hospitality Management. 2021 Jan 1;92:102702.
4. Haryanto T. COVID-19 pandemic and international tourism demand. JDE (Journal of Developing Economies). 2020 Jun 10;5(1):1–5.
5. Lim SM, Chia SE. The prevalence of fatigue and associated health and safety risk factors among taxi drivers in Singapore. Singapore medical journal. 2015 Feb;56(2):92.
6. Chen JC, Dennerlein JT, Shih TS, Chen CJ, Cheng Y, Chang WP, Ryan LM, Christiani DC. Knee pain and driving duration: a secondary analysis of the Taxi Drivers’ Health Study. American journal of Public health. 2004 Apr;94(4):575–81.
7. Kobayashi F, Watanabe T, Watanabe M, Akamatsu Y, Tomita T, Nakane T, Furui H, Takeuchi K, Okada A, Ohashi R, Hayano J. Blood pressure and heart rate variability in taxi drivers on long duty schedules. Journal of Occupational Health. 2002 Jul;44(4):214–20.
8. Gany FM, Gill PP, Ahmed A, Acharya S, Leng J. “Every disease… man can get can start in this cab”: focus groups to identify south Asian taxi drivers’ knowledge, attitudes and beliefs about cardiovascular disease and its risks. Journal of immigrant and minority health. 2013 Oct;15(5):986–92.

9. Magnusson K, Nygård K, Methi F, Vold L, Telle K. Occupational risk of COVID-19 in the first versus second epidemic wave in Norway, 2020. Eurosurveillance. 2021 Oct 7;26(40):2001875.
10. Windsor-Shellard B, Butt A. Coronavirus (COVID-19) related deaths by occupation, England and Wales: deaths registered between 9 March and 25 May 2020.
11. Lizya S, Ningtyas FR, Shellasih NM, Nurhasana R. The impact of COVID-19 on socio-psychological and cultural conditions of online taxi drivers in achieving a sustainable city. InIOP Conference Series: Earth and Environmental Science 2021 Mar 1 (Vol. 716, No. 1, p. 012061). IOP Publishing.
12. Harper CA, Satchell LP, Fido D, Latzman RD. Functional fear predicts public health compliance in the COVID-19 pandemic. International journal of mental health and addiction. 2021 Oct;19(5):1875–88.
13. Bucsky P. Modal share changes due to COVID-19: The case of Budapest. Transportation Research Interdisciplinary Perspectives. 2020 Nov 1;8:100141.
14. Romero YH, Sosa RG. The vulnerability at work of taxi drivers in the Zona Metropolitana del Valle de Mexico in the face of the COVID-19 pandemic. Espacio Abierto. 2020:12–30.
15. Katta S, Badger A, Graham M, Howson K, Ustek-Spilda F, Bertolini A. (Dis) embeddedness and (de) commodification: COVID-19, Uber, and the unravelling logics of the gig economy. Dialogues in Human Geography. 2020 Jul;10(2):203–7.
16. Li Q, Bai Q, Hu A, Yu Z, Yan S. How Does COVID-19 Affect Traffic on Highway Network: Evidence from Yunnan Province, China. Journal of Advanced Transportation. 2022 Feb 28;2022.
17. New York City Department of Health and Mental Hygiene. COVID-19: Data; public use dataset accessed on 5th Jan. 2022
18. “Taxi Fare - TLC.” Welcome to NYC.Gov — City of New York, https://www1.nyc.gov/site/tlc/passengers/taxi-fare.page. Accessed 5th Jan. 2022.
19. Manley E, Ross S, Zhuang M. Changing Demand for New York Yellow Cabs during the COVID-19 Pandemic. Findings. 2021 May 7:22158.
20. Xiong J, Tang Z, Zhu Y, Xu K, Yin Y, Xi Y. Change of Consumption Behaviours in the Pandemic of COVID-19: Examining Residents’ Consumption Expenditure and Driving Determinants. International Journal of Environmental Research and Public Health. 2021 Jan;18(17):9209.
21. Centers for Disease Control and Prevention. (2018, March 21). History of 1918 flu pandemic. Centers for Disease Control and Prevention. from https://www.cdc.gov/flu/pandemic-resources/1918-commemoration/1918-pandemic-history.htm (February 5, 2022, date last accessed).
22. Beach B, Clay K, Saavedra M. The 1918 influenza pandemic and its lessons for COVID-19. Journal of Economic Literature. 2022 Mar;60(1):41–84.
23. Yu, Liping. (2013). Big data and big data economics. China Soft Science, (7), 177–183.

Figures

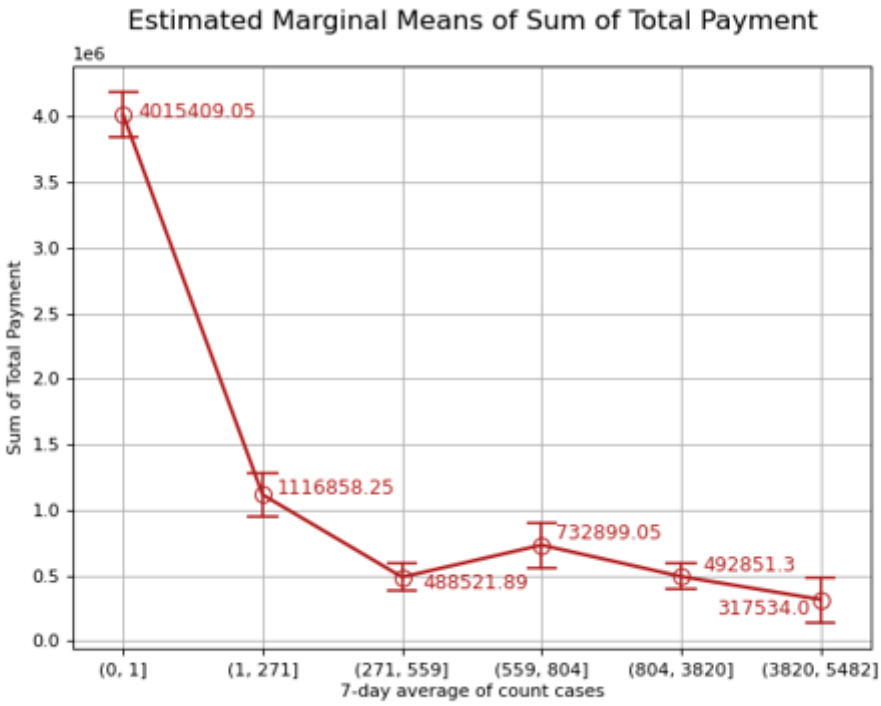

Figure 1

The overall trend of EMM of the sum of total payment per day under different levels of count cases

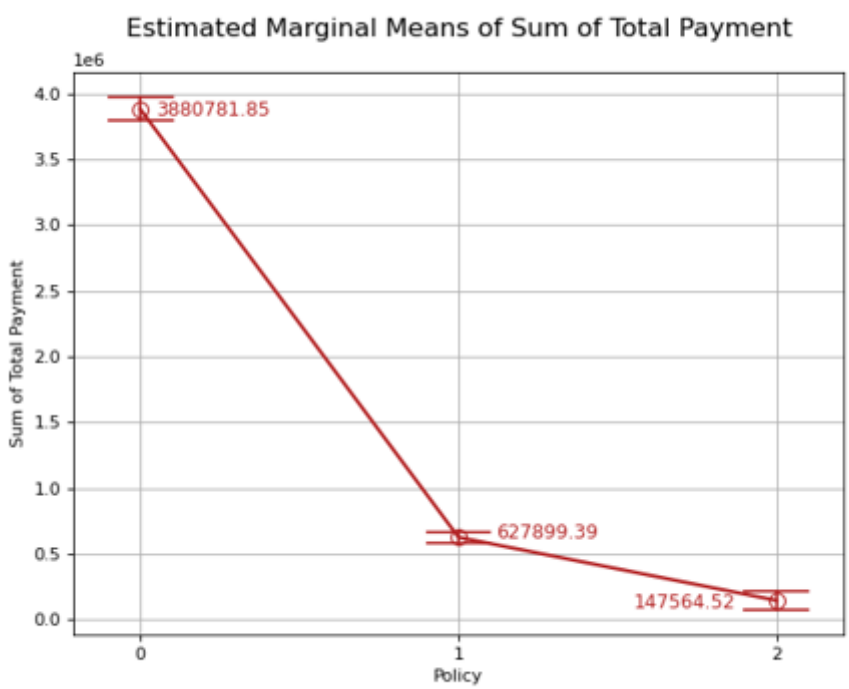

Figure 2

The overall trend of EMM of the sum of total payment per day as the impact area of public health policy enlarges is decreasing.

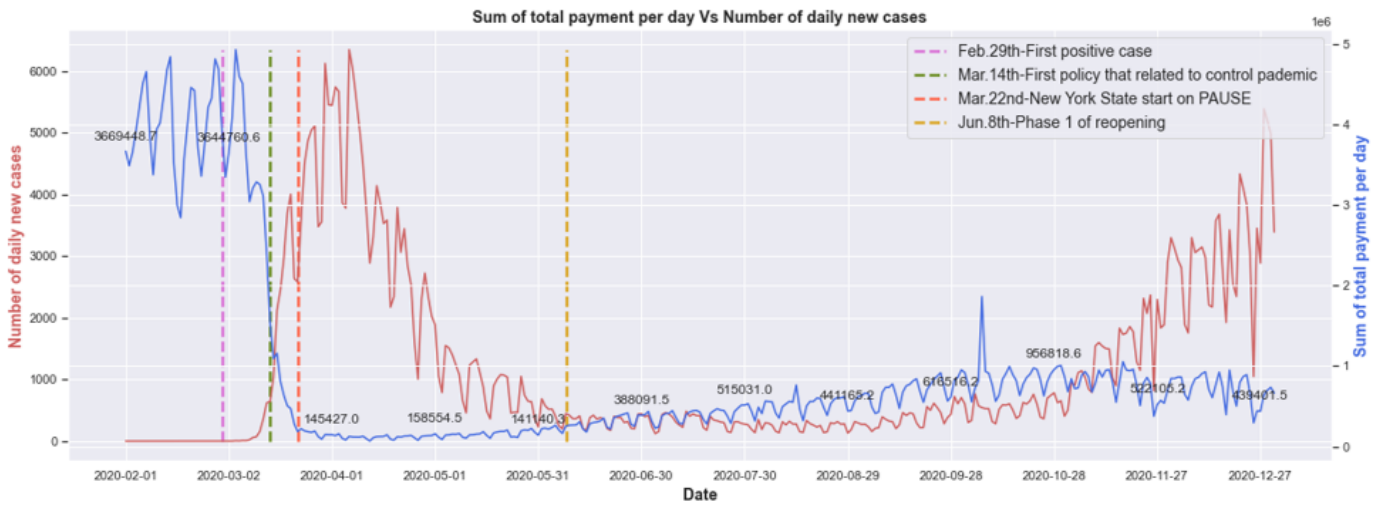

Figure 3

The line chart illustrates the trend of sum of total payment per day during the pandemic and key date of relevant policy are marked.

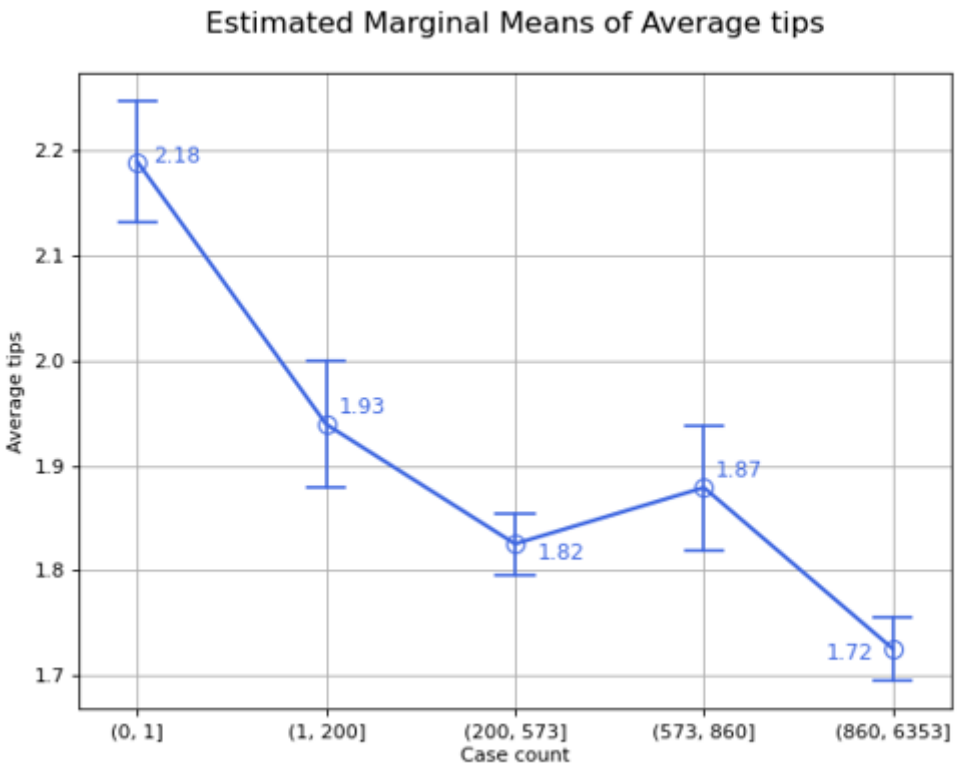

Figure 4

The overall tendency of EMM of average tips is decreasing as the epidemic become severer.

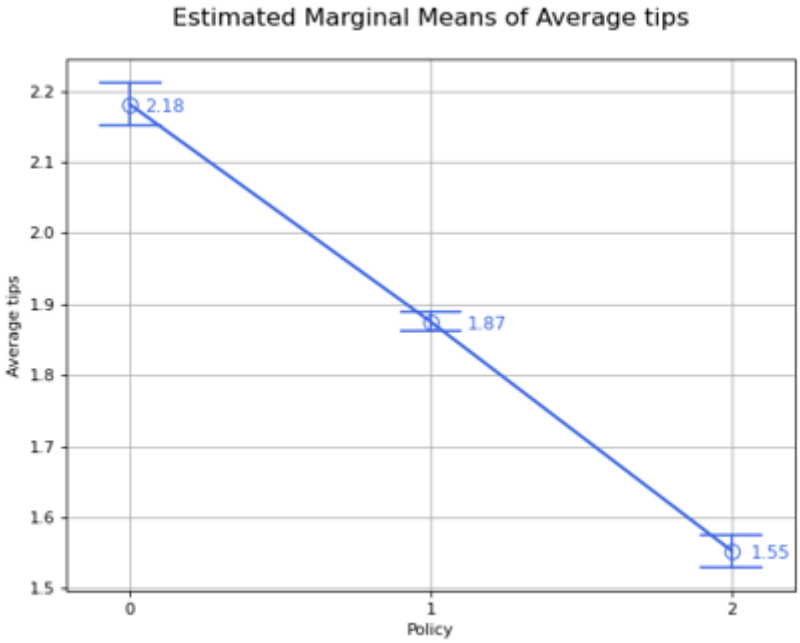

Figure 5

The EMM of average tips continuously drops from policy = 0 to policy = 2.

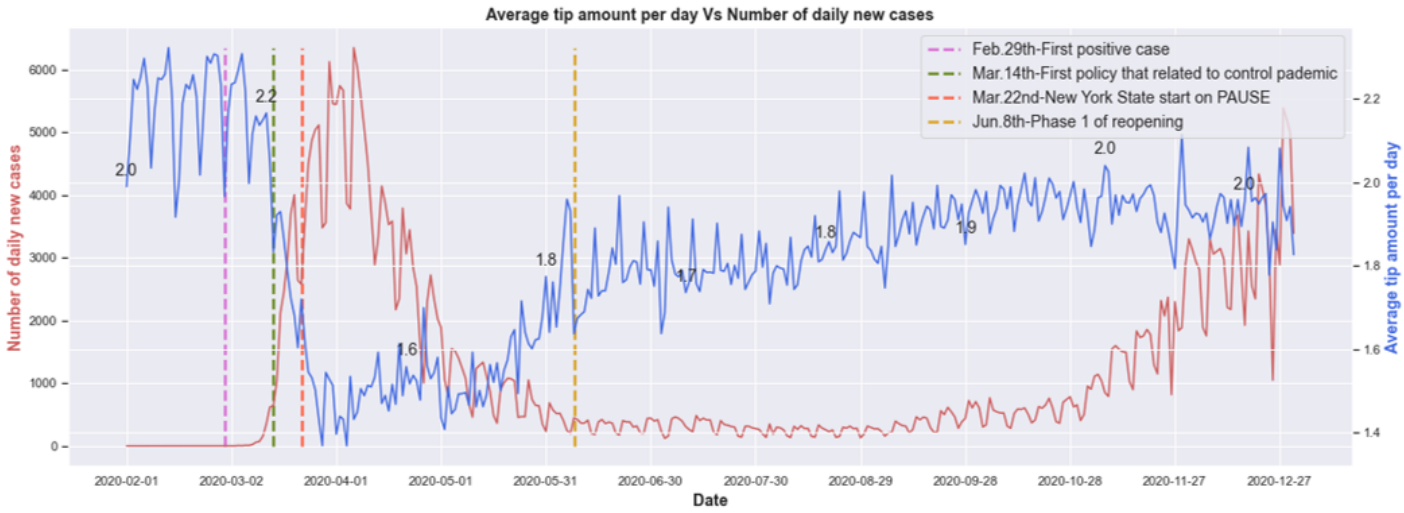

Figure 6

The line chart illustrates the trend of tip amount per day during the pandemic and key date of relevant policy are marked.

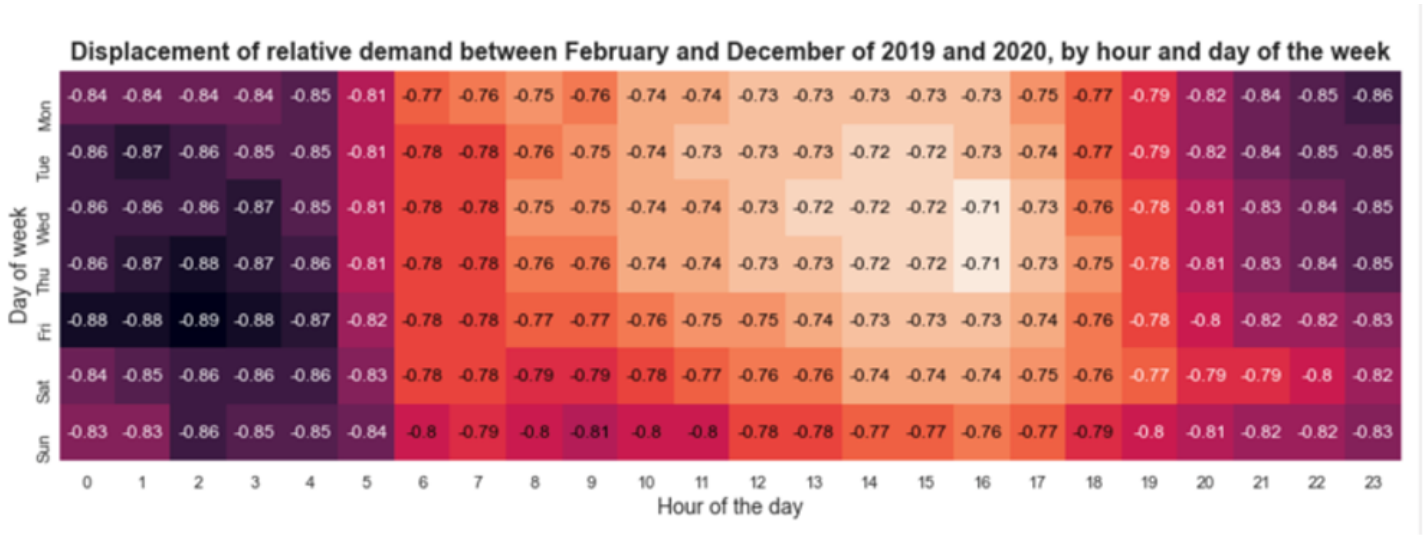

Figure 7

The confusion matrix plot demonstrates the difference in demand for taxis between 2019 and 2020, in units of hours and days.